

TecnOlivo. Entregables E3.1-E3.3

Modelado de datos agronómicos

v [1.1]
Emisor: [UHU]
Fecha: [17/12/2020]

0. Introducción

La actividad 3 se centra en el desarrollo de modelos que permitan la correcta interpretación de los datos obtenidos por los sensores, tanto terrestres como los equipados en la plataforma aérea. El objetivo es ofrecer al olivicultor información precisa acerca del desarrollo de la planta y el fruto, monitorizando parámetros clave para la toma de decisiones como el estatus hídrico y nutricional del cultivo o el estado de maduración y calidad de la cosecha. Para la creación de los modelos agronómicos se analizarán en primera instancia los datos obtenidos por los distintos sensores de manera individual. Adicionalmente, en los casos en los que más de un sensor aporte información acerca de un mismo parámetro, estos serán también analizados de forma conjunta para maximizar su capacidad descriptiva.

Para la generación y validación de modelos agronómicos a partir de datos obtenidos por sensores, es necesario un proceso de calibración que se realiza tomando como referencia mediciones obtenidas en el terreno (estos conjuntos de referencia se denominan “gold standard”). Para cada parámetro a modelizar se determinó un conjunto “gold standard” usando variables agronómicas ampliamente aceptadas de uso estandarizado.

Durante la fase de modelización de los parámetros objetivo se analizó el potencial descriptivo de los datos obtenidos con los sensores enfrentados a los conjuntos “gold standard”. Los algoritmos y técnicas empleados para la modelización en cada caso dependerán de la naturaleza de los sensores utilizados, que podrán incluir termografía, análisis de imagen y espectroscopia a partir de sensores multi o hiperespectrales para la aplicación de técnicas de extracción de conocimiento como "data mining". Concretamente, el uso de un enfoque basado en "data mining" consiste en el uso técnicas de análisis de imagen o espectroscopía para la extracción de características procedentes de varios sensores para el entrenamiento de los modelos predictivos. Se identificó y estableció el conjunto de las mejores características medidas con los sensores disponibles para la consecución de cada uno de los problemas agronómicos descritos. Los modelos fueron generados y validados de forma iterativa con los datos obtenidos.

El último hito se trata de evaluar la eficiencia y precisión de los modelos generados en la Acción A3.2 en condiciones de uso real en olivares comerciales. Para ello se realizarán una serie de test interactivos que permitirán la evaluación y ajuste fino de los parámetros finales en un entorno de uso comercial.

Modelado de datos agronómicos

Índice

0. Introducción	1
Índice	2
1. Adquisición de datos mediante sensores y generación de conjuntos “gold standard”	1
1.1 Estado hídrico.....	1
1.2 Estado nutricional	4
1.3 Indicadores de calidad de la cosecha.....	5
2. Desarrollo de modelos de evaluación a partir de datos procedentes de sensores	8
2.1 Estrés hídrico.....	8
2.2 Estrés nutricional	9
2.3 Indicadores de calidad de la cosecha.....	9
3. Validación de los modelos y algoritmos mediante pruebas de campo.....	10

1. Adquisición de datos mediante sensores y generación de conjuntos “gold standard”

La modelización de datos adquiridos con sensores para la estimación de parámetros agronómicos requiere la adquisición previa de un conjunto de datos de referencia (gold standard), el cual debe ser de un volumen suficiente y un rango amplio. Esto se consigue mediante el establecimiento previo de un diseño experimental que dé lugar a una variación del parámetro objetivo en campo. En el caso del estado hídrico esto se consigue aplicando programas de riego diferenciales a lo largo de la superficie de cultivo, que vayan desde riegos normales a deficitarios. Por otro lado, si el parámetro objetivo es el estado nutricional se debe variar la dosis de fertilizantes aplicados.

Una vez generada la mencionada variabilidad muestral, se debe seleccionar una variable de referencia que sirva para prospectar las condiciones reales del cultivo, la cual en última instancia será sustituida por la respuesta de los modelos alimentados por las medidas del sensor.

1.1 Estado hídrico

En el seno del proyecto TecnOlivo se llevó a cabo un experimento con el objetivo de desarrollar y validar un sensor IR de bajo costo para la monitorización del estado hídrico del olivar. El prototipo desarrollado es capaz de estimar la temperatura de un objeto, en este caso el dosel arbóreo, a partir de las emisiones electromagnéticas de este en el infrarrojo. Este dispositivo se desarrolló sobre una plataforma de bajo costo, específicamente un ArduinoTM MEGA 2560. Además, incluye una pantalla que posibilita la supervisión del correcto funcionamiento del dispositivo durante su uso en campo, una tarjeta SD para el almacenamiento de datos, un botón activador, una placa de interconexiones y una batería. Todo ello contenido en una carcasa realizada mediante impresión 3D. El sensor IR integrado opera en un rango de temperatura apropiado para aplicaciones de agricultura (-40°C - +85°C), con una resolución de 16x4 píxeles y un ángulo de apertura de 40°. El mencionado sensor IR ofrece un valor de temperatura por cada píxel y además está configurado para determinar la temperatura media del total de píxeles. De este modo, con cada imagen se almacena el valor de temperatura medido por cada píxel (64), así como la temperatura media general de la imagen. La Fig. 1 muestra una representación 3D del prototipo desarrollado. Destacar que, pese a que la configuración evaluada hasta el momento está diseñada para su uso manual, su reducido costo abre la posibilidad a su instalación en una red de sensores donde el número de nodos es determinante de cara a la resolución espacial de la monitorización.

El fundamento teórico que sustenta la estimación del estado hídrico del olivar a partir de la temperatura de su dosel se basa en su comportamiento isohídrico. El comportamiento isohídrico es una característica propia de especias adaptadas a condiciones de sequía. Este consiste en la regulación de la tasa de transpiración en función de la disponibilidad de

agua, mediante el control del grado de apertura estomático. De este modo, ante una situación de escasez la planta responde ocluyendo sus estomas, lo cual reduce la evapotranspiración limitando la pérdida de agua. En condiciones normales parte de la energía térmica que recibe la planta por la radiación solar se disipa a través de la evaporación. Por lo tanto, un árbol bajo condiciones de buena disponibilidad de agua presentará una temperatura de dosel igual o menor a la ambiental, mientras que un árbol bajo escasez hídrica presentará una temperatura de dosel superior a la ambiental.



Figura 1. Representación 3D de la cámara IR desarrollada.

Tradicionalmente, la determinación del estado hídrico del olivar se ha realizado midiendo su potencial hídrico haciendo uso de cámaras de presión o cámaras Schölander. Otro parámetro altamente influenciado por el estado hídrico del olivo es su conductancia estomática, la cual se mide con el porómetro. El potencial hídrico es la presión requerida para la movilización de la sabia a través de la planta, la cual depende del balance entre el agua captada por la raíz mediante absorción y la pérdida mediante transpiración por los estomas de las hojas. La conductancia estomática por su parte es la cantidad de aire que transpira a través de los estomas de las hojas en un momento determinado, lo cual va a estar determinada por el número de estomas y el grado de apertura de estos. Ambas variables son indicadores de estrés hídrico ampliamente estudiados y validados por la comunidad científica. Sin embargo, su uso para la monitorización a gran escala es inviable debido a que suponen procesos tediosos que requieren personal cualificado y una elevada inversión de tiempo, por lo cual su uso queda restringido al ámbito de la investigación.

El experimento mencionado se llevo a cabo en un olivar comercial superintensivo de la variedad arbequina localizado cerca de la localidad de Elvas (Portugal) y administrado por el socio del proyecto Elaia TM. Tras varios meses de exposición a diferentes tratamientos

de riego se realizó una prospección del estado hídrico de los olivos sometidos a cada uno de los tratamientos con el objetivo de generar un conjunto de datos de referencia (gold standard). Para esto se usó el potencial hídrico medido usando una cámara de presión o cámara Schölander (Solfranc TecnologíasTM) (Fig. 2a) y la conductancia estomática medida usando un porómetro de hojas modelo SC-1 (Decagon DevicesTM) (Fig. 2b). Simultáneamente a la medición de ambos parámetros de referencia se determinó la temperatura del dosel arbóreo en cada punto de muestreo usando la cámara IR propuesta.

Los resultados de las mediciones se correspondieron con lo esperado en base al diseño experimental planteado y nos permitieron adquirir un conjunto de datos “gold estándar” con el rango y la extensión apropiados.

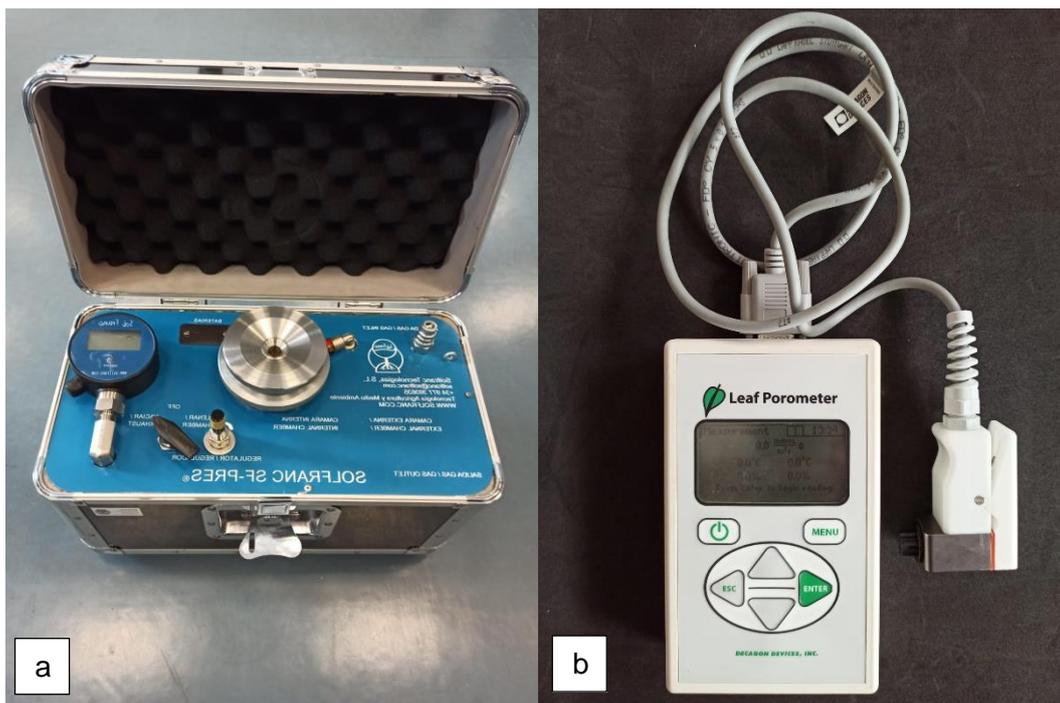


Figura 2. Dispositivos empleados para la adquisición de indicadores de estrés hídrico: cámara Scholander (a) y porómetro (b).

1.2 Estado nutricional

En el seno del proyecto TecnOlivo se llevó a cabo un experimento con el objetivo de desarrollar una metodología que permitiera cartografiar de manera eficiente el estado nutricional del olivar mediante imágenes espectrales captadas con una cámara multispectral comercial montada sobre una base aérea pilotada de forma remota. Tradicionalmente, la caracterización del estado nutricional del olivar se ha hecho mediante análisis químico. Esto requiere un proceso previo de recolección de muestras, el transporte de estas a laboratorios apropiadamente habilitados y el procesado posterior de las mismas mediante técnicas que requieren personal cualificado, una elevada inversión de tiempo y reactivos y maquinaria de un elevado coste. Por todo ello, en el sector olivarero la prospección del estado nutricional de las fincas se realiza por lo general en una única ocasión a lo largo de la campaña y usando como referencia pocos puntos de muestreo, lo que da una idea alejada de la variabilidad real presente en las fincas y por lo general redundante en un uso excesivo de fertilizantes.

Los tres procesos de absorción, reflexión y transmisión describen las interacciones entre la radiación incidente, los componentes bioquímicos de las hojas y los rasgos biofísicos del dosel. Con lo cual, caracterizando esta interacción en el rango espectral adecuado podemos modelizar parámetros agronómicos clave para la toma de decisiones. Esto unido a la capacidad de los RPAS para cubrir grandes extensiones de terreno, nos daría un enfoque capaz de cartografiar la extensión total de fincas con la suficiente resolución espacial y temporal. Ajustando el uso de fertilizantes a las necesidades específicas de cada parcela del cultivo y por lo tanto favoreciendo la economía del sector a través del ahorro de recursos y contribuyendo a una agricultura más sostenible.

Los parámetros objetivo en el mencionado experimento fueron el contenido foliar de nitrógeno (N), fósforo (P) y potasio (K), los cuales son los fertilizantes mayoritariamente aplicados en los programas de fertilización. El experimento se llevó a cabo en Elvas, Portugal, en un olivar superintensivo cedido por el socio del proyecto Elaia TM. El área de cultivo objeto de ensayo, comprendía dos zonas alternativas con olivos adultos de la variedad Arbequina y Arbosana respectivamente. Sobre dicha área, se dispuso un esquema de fertirriego diferencial comprendido por tres tratamientos definidos con el objetivo de generar variabilidad en cuanto al contenido foliar de N, P y K.

Tras varios meses de exposición a diferentes tratamientos de fertilización se realizó una prospección del estado nutricional de los olivos sometidos a cada uno de los tratamientos para obtener un conjunto de datos de referencia (gold standard). La captura de datos espectrales se realizó con una plataforma aérea autónoma multisensor (Fig. 3) desarrollada sobre la base de un vehículo aéreo pilotado de forma remota (RPAS) comercial. En concreto, se utilizó el RPAS DJI Matrice 100 como base, al que se le integró una cámara multispectral MicaSense RedEdge-M, que permite la captura en 5 bandas espectrales distintas: azul, verde, rojo, rojo frontera, e infrarrojo cercano. Para la calibración de las

mediciones espectrales, y hacerlas por tanto robustas ante distintas condiciones de iluminación, la plataforma fue también equipada con un sensor de medición de luz incidente. Adicionalmente, la cámara multispectral fue dotada con un sistema de georreferenciación propio. Así, la plataforma desarrollada permite recopilar datos espectrales de manera eficiente y con una altísima resolución espacial.



Figura 3. Micasense Rededge-M instalada en la plataforma aérea (DJI Matrice 100)

El día del experimento se sobrevoló el cultivo, capturando imágenes multispectrales de toda la superficie de estudio. A continuación, se seleccionaron y georreferenciaron 70 puntos de muestreo, de los cuales se recogieron muestras foliares para someterlas a un análisis químico de laboratorio. Los análisis foliares revelaron una evidente variabilidad en cuanto al contenido foliar de los macronutrientes objeto de ensayo (N, P y K). Posteriormente, se trataron las imágenes multispectrales y, mediante técnicas de visión artificial, se calcularon los valores medios de intensidad de la radiación reflejada por el dosel en cada una de las cinco bandas del espectro captadas, para todos los puntos de muestreo. Obteniendo así un conjunto de datos de referencia “gold estándar” que posteriormente nos permitiría modelizar los parámetros objetivo.

1.3 Indicadores de calidad de la cosecha

En el seno del proyecto TecnOlivo se llevó a cabo un experimento con el objetivo de desarrollar y validar un sensor multispectral de bajo costo para la caracterización de parámetros clave para el control de calidad de aceitunas. Las almazaras usan de manera rutinaria parámetros como el contenido graso, la acidez o la humedad de los frutos para determinar el valor económico de la cosecha recibida. Los agricultores por su parte deciden el momento de la recogida en base a criterios subjetivos como el momento de la campaña

o la opinión experta. Sin embargo, la realidad es que los mencionados parámetros no se desarrollan de manera homogénea en toda la extensión de cultivo, ya que existen diferentes factores como el tiempo de exposición a la radiación solar, la composición y estructura del suelo, la incidencia parcial de patologías etc, que se distribuyen heterogéneamente y afectan a procesos fisiológicos de la planta implicados en la maduración del fruto. Por lo tanto, la manera óptima de proceder sería contar con medios que permitiesen monitorizar de manera continua el proceso de maduración a lo largo de la extensión del cultivo y adaptar de esta manera la recolección a las necesidades de cada parcela, esto permitirá mejorar la calidad del producto final y desde el punto de vista de los agricultores obtener un precio óptimo.

El objetivo final de la presente línea de investigación es la implementación de este tipo de dispositivos en condiciones de campo. Como un primer acercamiento a este objetivo se realizó un experimento preliminar para verificar la viabilidad del sensor multiespectral desarrollado en condiciones controladas de laboratorio. El dispositivo desarrollado incluye un sensor multiespectral y una fuente de luz, ambos componentes dispuestos sobre una estructura metálica con un ángulo de 90° entre ellos, como se muestra en la Fig. 4. Esta disposición pretende facilitar la captación de luz por el sensor después de ser reflejada por la muestra. Además, el dispositivo estaba provisto de una cubierta sólida que lo aislaba de la iluminación exterior. El sensor multiespectral se basa en la tarjeta de desarrollo AS7265x, que consta de tres dispositivos sensores: AS72651, AS72652 y AS72653. El AS7265x se conectó a un Arduino MKR 1200, que controla el sensor y almacena las mediciones en una tarjeta microSD extraíble. Se diseñó una carcasa a medida e imprimió en 3D para instalar toda la electrónica. Se colocó una capa redonda y fina de PTFE delante del sensor para actuar como difusor y homogeneizar la iluminación. El dispositivo multiespectral desarrollado presenta un rango de captación espectral que va desde el visible hasta el NIR. Cada uno de los tres dispositivos sensores tiene 6 filtros ópticos independientes. La combinación de los tres sensores da como resultado un sensor multiespectral de 18 canales (410-940 nm) con un ancho total a la mitad de un máximo de 20 nm. La fuente de luz seleccionada fue una bombilla halógena porque su rango espectral de emisión es más amplio que el de las bombillas led, por lo que permite explotar las capacidades del sensor en el dominio NIR. La captura de espectros se controla con un ordenador utilizando un software desarrollado a medida con interfaz con la placa Arduino.



Figura 4. Imagen del prototipo desarrollado con el sensor multiespectral a la izquierda y la fuente de luz a la derecha

El experimento de validación se llevó a cabo en los laboratorios de la almazara de Nuestra Señora de la Oliva en Gibraleón, Huelva, perteneciente al socio del proyecto Oleodiel. Durante una semana, todas las muestras de aceitunas ($n = 50$) entregadas por varios agricultores para su tasación, fueron analizadas simultáneamente con el dispositivo desarrollado y con los medios usados de manera rutinaria en la almazara. Para este devenir, Oleodiel utiliza un analizador automático disponible en el mercado (OliveScan™), que ha sido ampliamente probado, el cual determina el contenido graso, la humedad y la acidez de las muestras de aceitunas.

De este modo, se adquirió un conjunto de datos “gold standard” compuesto por la reflectancia de las muestras en cada una de las 18 bandas espectrales captadas por el sensor como variables de entrada y la acidez, contenido graso y humedad de las muestras determinada con el analizador automático como variables de referencia.

2. Desarrollo de modelos de evaluación a partir de datos procedentes de sensores

Una vez adquiridos los conjuntos “gold standard” se abre la fase de modelización matemática de los parámetros objetivo. Los algoritmos y técnicas empleados para la modelización en cada caso dependen de la naturaleza de los sensores utilizados.

2.1 Estrés hídrico

El valor de temperatura del dosel per se, no se considera un buen indicador del estado hídrico de cultivos, ya que esta depende de factores ambientales. De este modo, un mismo valor de temperatura de dosel puede significar que el árbol medido se encuentre bajo condiciones de una insuficiente disponibilidad de agua o no en función de las condiciones ambientales. Por ello, los datos ofrecidos por el sensor IR propuesto requieren ser transformados de manera que se estandaricen las mediciones haciéndolas válidas independientemente de las condiciones ambientales. Para ello usamos el índice CWSI (*Crop Wáter Stress Index*), el cual es un índice de vegetación ampliamente aceptado como indicador de estrés hídrico. Este consiste en una normalización de la temperatura del dosel en base a unos límites teóricos, T_{dry} y T_{wet} , que corresponden a la temperatura del dosel en ausencia de transpiración y transpiración óptima, respectivamente.

$$CWSI = \frac{T_{dosel} - T_{wet}}{T_{dry} - T_{wet}}$$

La mayor dificultad de este método reside en la determinación de estos límites. En el estado del arte encontramos diferentes aproximaciones que van desde el uso de superficies artificiales de referencia a la estimación a partir de fórmulas matemáticas basadas en variables ambientales. En el presente experimento se planteó un enfoque novedoso, que consiste en el establecimiento de estos límites mediante una aproximación adaptativa basada en los histogramas de temperatura derivados del total de las imágenes térmicas. De este modo, T_{wet} correspondería a la parte más fría del histograma de temperaturas, mientras que T_{dry} se extrae del extremo más caliente. Este método se secunda en el amplio volumen de mediciones obtenidas, ya que cada imagen ofrece un total de 64 valores de temperatura (uno por píxel) y en la distribución de los mismos en forma de campana de Gauss. Así, el modelo ofrece un valor normalizado entre 0 y 1 que indica el estado hídrico del olivo a partir del valor de temperatura medido con el sensor IR propuesto.

2.2 Estrés nutricional

En este caso el proceso de modelización es más complejo ya que como se comenta en el apartado 1.2 la cámara multiespectral empleada capta la reflectancia del dosel en cinco bandas espectrales. De modo que en este caso contamos con un espacio de característica pentadimensional. Por ello, se optó por un método no lineal y no paramétrico impulsado por aprendizaje automático, como es la red neuronal artificial. Dicho modelo se desarrolló usando el software Matlab 2020a. Para cada uno de los parámetros objetivo (contenido foliar de N, P y K respectivamente) se diseñó una red con una arquitectura sencilla comprendida por una única capa oculta con una neurona, cinco variables de entrada que serían la intensidad de la luz reflejada por el dosel en las cinco longitudes de onda captadas por el sensor y una única salida que sería el parámetro de referencia en cada caso.

2.3 Indicadores de calidad de la cosecha

El sensor multiespectral empleado en este experimento determina la reflectancia de la luz en 18 bandas espectrales. Por lo tanto, la no linealidad existente entre estas variables en la determinación de los parámetros objetivo hace necesario de nuevo el uso de métodos no paramétricos. Una vez más se optó por la red neuronal artificial. Para cada uno de los parámetros objetivo (contenido graso, humedad y acidez de los frutos) se definió una red con una arquitectura comprendida por una única capa oculta con 6 neuronas, 18 datos de entrada que serían la reflectancia en cada una de las 18 bandas espectrales captadas por el sensor y una única salida que sería el parámetro de referencia en cada caso.

3. Validación de los modelos y algoritmos mediante pruebas de campo

El rendimiento de los modelos se mide mediante el coeficiente de correlación (R^2) y el error cuadrático medio de la raíz (RMSE), entre la respuesta de cada modelo en la estimación de su respectivo parámetro objetivo y los datos de referencia para ese parámetro en concreto. Valores bajos de RMSE y valores próximos a 1 de R^2 indican el mejor rendimiento de los modelos.

En todos los casos, el objetivo inicial era el de recolectar, en la campaña del 2020, nuevos datos con los que validar los modelos de manera independiente a los datos con los que se entrenaron. Debido a las dificultades impuestas por la pandemia del COVID19, ha sido imposible la ejecución de estos nuevos experimentos de campo. No obstante, se ha realizado una segmentación de los datos recogidos en las campañas anteriores para conseguir este fin. En el informe final del proyecto se describirán con todo detalle los resultados alcanzados.