

Grupo Operativo FRUIT FORECAST

Desarrollo de modelos de predicción de producción y calidad de fruto en Cerezo (*Prunus avium*) y Melocotonero (*Prunus persica*)

El objetivo de este Grupo Operativo ha sido desarrollar modelos predictivos, mediante el uso de datos productivos y de calidad de fruto y de tecnología Big Data, que anticipen la información sobre volúmenes de cosecha y calidad de fruto, con el fin de obtener una planificación de cosecha fiable, tanto para cereza como para melocotón.

Gemma Reig^a, Miguel Ros^b, Miriam Hernández^c, Albert Duaigües^b, David Olmo^b, Elisabeth Molina^d, Enric Pedrós^d, Dídac Masip^e, Marc Nart^c

^a IRTA-Fruitcentre, 25003 Lleida, ^b Raw Data, 46024 València,

^c Fruits de Ponent, 25180 Alcarràs, ^d Clúster FEMAC, 25003 Lleida,

^e Producció Cirera Ribera d'Ebre, 43746 Tivissa

INTRODUCCIÓN

Este Grupo Operativo ha sido liderado por dos empresas muy importantes del sector de fruta de hueso, principalmente del sector del cerezo y del melocotón. La primera empresa es Producció Cirera Ribera d'Ebre SL (Cerima Cherries), creada en 2012 y que actualmente cuenta con 450 hectáreas (ha) repartidas por toda la Ribera d'Ebre y parte de la Terra Alta, Segrià y Baix Ebre. La segunda es el Grupo Cooperativo Fruits de Ponent SCCL, constituido en 1992 y que actualmente cuenta con 2.546 ha de las cuales el 75% están destinadas al cultivo de fruta de hueso, principalmente melocotón.

Para llevarlo a cabo, estas dos empresas han contado con la participación de la empresa tecnológica RawData SL, creada en 2018 con la intención de proporcionar todas las soluciones tecnológicas que necesita el sector agrario mediante una misma plataforma, y del Instituto de Investigación y Tecnología Agroalimentaria (IRTA) de la Generalitat de Catalunya, en particular el centro de investigación de Lleida IRTA-Fruitcentre, focalizado, prin-



principalmente, en dar respuesta al sector de fruta dulce y de hueso del Valle del Ebro. Toda la coordinación del Grupo Operativo se ha llevado a cabo a través del Clúster FEMAC, quien agrupa a las principales empresas y centros tecnológicos de Catalunya en el sector de los medios de producción agrícola.

OBJETIVO

El objetivo de este Grupo Operativo ha sido desarrollar modelos predictivos, mediante el uso de datos productivos y de calidad de fruto y de tecnología Big Data, que anticipen la información sobre volúmenes de cosecha y calidad de fruto, con el fin de obtener una planificación de cosecha fiable, tanto para cereza como para melocotón. De esta forma, poder ganar una posición estratégica a la hora de tomar decisiones comerciales estratégicas y de negociar contratos de venta (precios, fechas de entrega y volúmenes) con clientes. Pero también, poder reducir los costes de producción, gracias a una mejor eficiencia en la gestión de los recursos, tanto humanos como materiales.

Como consecuencia, reducir la incertidumbre actual en la planificación de cosechas y mejorar la fiabilidad de las predicciones de volumen de cosecha de calidad del fruto con las tecnologías Big Data y Machine Learning. Cuando hablamos de Big Data nos referimos a la rama tecnológica que se encarga de analizar y estudiar grandes cantidades de datos y que se complementa perfectamente con la tecnología Machine Learning, rama en la que se hace uso de la capacidad de las máquinas de aprender a través de grandes cantidades de datos aportados por el Big Data.

RECOGIDA DE DATOS

Para ello, se han elaborado protocolos específicos de recogida de datos (meteorológicos, productivos, índices de vegetación y calidad del fruto), durante dos campañas seguidas, para cada una de las dos especies frutales,

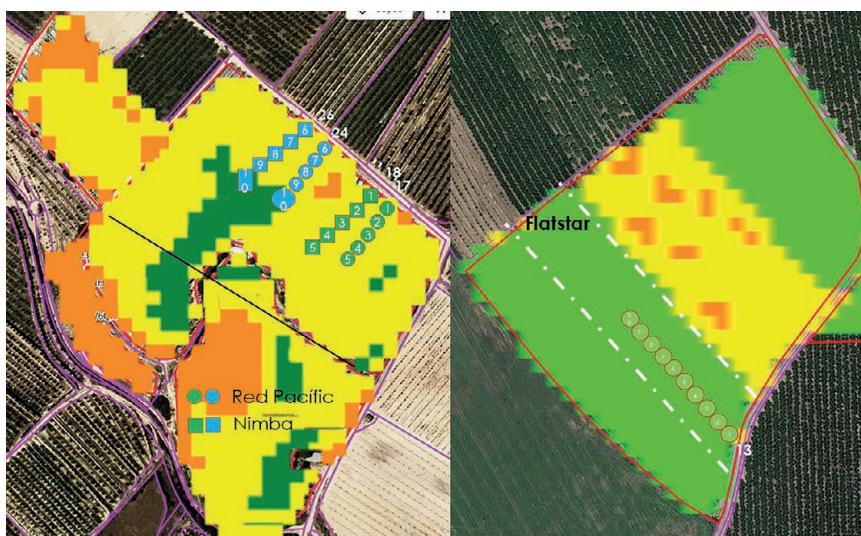


Figura 1. Ejemplo de dos parcelas de estudio (cerezo, foto de la izquierda, y melocotonero, foto de la derecha) con zonas de distintos vigores (diferentes colores en función del NDVI, índice de vegetación de diferencia normalizada).



Figura 2. Marcaje de ramas fructíferas en cerezo durante la plena floración.

cerezo y melocotonero. La recogida de datos (productivos, índices de vegetación y calidad de fruta) se ha centrado en aquellas variedades más representativas, en cuanto a número de hectáreas plantadas, de cada cultivo. Para poder tener representatividad de cada variedad, se han seleccionado 2 parcelas por variedad, de distinta localización geográfica. Para melocotón se han evaluado 12 variedades distintas, mientras que para cerezo han sido 6 variedades. A todo esto, hay que añadir que, paralelamente,

se han recopilado el máximo de datos históricos posibles, principalmente datos de producción, de índices de vegetación y meteorológicos.

Cada parcela se ha geolocalizado en el SigPac, y posteriormente estos datos se han cruzado con imágenes satelitales con el fin de obtener los índices de vigor (Figura 1) de los árboles de cada parcela de estudio, y así determinar con mayor eficacia los puntos de control, que en este estudio han sido 10 árboles por parcela.

Los datos productivos para determinar el aforamiento en campo y los de calidad de fruto de cada variedad, evaluados durante dos años consecutivos, varían en función de la especie frutal. Para el cultivo de cerezo se ha determinado la fecha de floración, la de cuajado y la del endurecimiento del hueso, el número de frutos en dos ramas fructíferas a dos y una semana antes de la cosecha comercial, y la producción real a central, mientras que, para el melocotonero, además de las mismas fechas mencionadas anteriormente y de la producción real a central, se han contado los frutos de una rama secundaria justo después del aclareo y antes de la cosecha comercial (Figura 2).

En cuanto a la calidad del fruto, para los frutos de cerezo se ha evaluado el diámetro, el color y el contenido de azúcares a dos y una semana antes de la cosecha comercial, mientras que para el fruto de melocotón se ha determinado el índice de degradación de clorofila también en los mismos períodos de tiempo (Figura 3).

Los datos de meteorología se obtuvieron tanto de todos los años históricos de producción aportados como de los dos años de duración del Grupo Operativo. La franja de cada año ha sido desde el momento de floración hasta la fecha de recolección y las principales variables que se han obtenido de los datos meteorológicos han sido la temperatura, pluviometría, velocidad del viento, humedad, radiación solar, etc.

Por otro lado, los índices de vegetación, tanto de los años históricos como de los dos años del Grupo Operativo, se han obtenido a través de las imágenes hiperespectrales procedentes del satélite europeo Sentinel 2. Como índices principales usados en este proyecto cabe mencionar el NDVI (Índice de vegetación de diferencia normalizada) y el EVI (índice de vegetación mejorado), obtenidos para la misma franja temporal que los datos meteorológicos.

MODELOS PRELIMINARES

Para las dos especies frutales del estudio, la recopilación de datos productivos y meteorológicos durante

el período del Grupo Operativo para determinar el aforamiento de cada parcela juntamente con los datos históricos (productivos, meteorológicos e índices de vegetación), propiedad de cada una de las empresas líderes, han dado como resultado unos modelos preliminares.

En primer lugar, se han elaborado unos modelos usando solamente datos históricos de productividad, meteorología e índices de vegetación, para cada especie frutal. Éstos no han representado con fiabilidad la realidad respecto a la producción. Para afinar estos modelos, se han añadido los datos de aforamiento recopilados durante el período del Grupo Operativo. Con esto, los resultados obtenidos mejoraron, ya que se aproximaban un 30% más a los datos de producciones de años pasados. Es decir, para ambas especies frutales se ha obtenido una fiabilidad a la realidad del 60-65%.

A esta fiabilidad obtenida, hay que tener en cuenta que durante las dos campañas del Grupo Operativo se han producido heladas primavera-



Figura 3. Medida del índice de clorofila a una semana antes de la cosecha.



les que han afectado con mayor o menor medida a muchas de las parcelas del estudio. Esto ha conllevado que algunas parcelas de estudio no se han evaluado durante dos años consecutivos. Pero aún con todo, se ha determinado que los datos de aforamiento obtenidos durante el Grupo Operativo, principalmente el conteo de frutos la semana antes de cosecha para ambas especies, son realmente importantes en la predicción de productividad de una campaña.

En el caso de la calidad de fruto, tanto para cerezo como para melocotonero, debido a la falta de históricos por parte de las empresas líderes, los modelos preliminares obtenidos en este Grupo Operativo no han aportado predicciones claras y próximas a la realidad. Los datos históricos es información clave para el entrenamiento del modelo, es decir, para obtener predicciones con mayor fiabilidad. Aun así, se ha podido observar que de los datos de maduración (índice de degradación de la clorofila para el melocotón y el diámetro, el color y el contenido de azúcares para la cereza) recolectados en distintos períodos de tiempo durante el Grupo Operativo han aportado gran cantidad de información. Cabe destacar que los modelos inteligentes han dado más relevancia a las medidas realizadas una semana antes de cosecha para ambas especies, dando un acierto del modelo entorno al 50%, en

el caso del melocotonero, y al 55-60% en el caso del cerezo. Además, cabe destacar que en esta última especie frutal la variable azúcares es mucho más determinante en el modelo de predicción de la maduración que el color o el diámetro del fruto.

CONCLUSIONES

Finalmente, y gracias a la duración del Grupo Operativo, se han conseguido mejoras en la optimización de la recogida de datos, tanto de aforamiento como de calidad del fruto. Se ha hecho un análisis exhaustivo de los datos obtenidos el primer año, con la intención de determinar cuál de todos los procesos son totalmente necesarios y cuál de ellos son mejorables. Se ha determinado que no es necesario realizar conteos de dos ramas por árbol en el caso del cerezo, ya que el conteo de una sola rama aporta prácticamente la misma información. También se ha descubierto que el número de parcelas por cada variedad de fruto es realmente importante, ya que la situación geográfica es un factor relevante dentro de la variedad de los datos. Se ha concluido

que hay que reducir el número de conteos en campo a uno solo, para ambas especies. Cuanto más cerca se realiza el conteo de frutos a la fecha de recolección mejor, ya que el modelo de aforamiento ha reflejado con más realismo los datos que habría en el momento de recolección. No obstante, esta predicción se contradice un poco con el "modus operandi" de las centrales frutícolas, que tienen que predecir la producción de ese año a 2-3 meses vista del inicio de campaña.

Como conclusión general de este Grupo Operativo, hay que destacar que hace falta más años de experimentación para tener una base de información lo suficientemente amplia como para que un modelo de predicción pueda contar con el mayor número de escenarios posibles y así predecir los posibles comportamientos en cada uno de ellos. Con esto, cabe reflejar que durante los años de recogida de datos tuvieron lugar heladas que obligaron a cambiar alguna toma de datos en parcela y variedad, lo cual desvirtuó la trazabilidad de los datos de una campaña a otra. ■



Proyecto financiado a través de la Operación 16.01.01 de Cooperación para la innovación del Programa de desarrollo rural de Catalunya 2014-2020