

# INFORME TÉCNICO FINAL

PROGRAMA ASTURIAS

## DATOS DE IDENTIFICACIÓN DEL PROYECTO DE I+D+i

Referencia proyecto	ID/2016/000180		
Acrónimo	TIC-ALI		
Título del Proyecto			
Periodo de Justificación	Desde :	01/01/2017	Hasta: 31/12/2017

En Noreña , a 14 de marzo de 2018

## 1. MEMORIA TÉCNICA

Explique **DETALLADAMENTE** las actividades realizadas, incluyendo la descripción de la metodología empleada, con referencia expresa a los hitos, tareas y plan de trabajo de la solicitud.

### **ACTIVIDAD 1: Adecuación de ASINCAR a los objetivos de Industria 4.0:**

#### **Tarea 1.1. Desarrollo vigilancia tecnológica.**

La información y el conocimiento son dos de los factores de mayor relevancia en relación al crecimiento de las empresas. En el caso de ASINCAR, dado su fuerte vinculación con la innovación, estos dos factores son clave para lograr una transferencia efectiva desde el mundo de la investigación a las empresas. En un mundo moderno donde los desarrollos en investigación son cada vez más complejos y con altos niveles de especialización, es difícil que las empresas, especialmente pymes, puedan estar al tanto y menos aún implementar estas nuevas tecnologías en sus líneas productivas. Es por ello que desde ASINCAR se considera primordial un profundo conocimiento de ambos sectores (investigación y empresa), para estar en disposición de identificar necesidades y posibilidades, y así, poder establecer los vínculos necesarios para formalizar el trasvase de conocimiento que conlleve no solo la satisfacción de necesidades, sino nuevos escenarios de mejora.

Dentro de las actividades de vigilancia tecnológica realizadas se debe comentar que en las temáticas donde se ha dedicado más tiempo son:

- Habilitadores digitales para la hibridación del mundo físico y digital.
- Tratamiento masivo de datos.

Dentro de las labores de vigilancia tecnológica realizadas dentro de los habilitadores digitales cabe destacar:

- Espectrofotometría NIR: Se han evaluado diferentes productos del mercado en función de las necesidades de la industria agroalimentaria. Esta búsqueda fue especialmente activa en lo concerniente a dispositivos con posibilidad de portabilidad. La miniaturización de la tecnología NIR avanza a pasos de gigante, con dispositivos cada vez más pequeños. Aun así, en función de la aplicación y las longitudes de onda necesarias existe una barrera tecnológica que impide crear dispositivos baratos de pequeño tamaño. Siendo una solución habitual presentar dispositivos con prestaciones inferiores a los espectrómetros de sobremesa o con necesidades adicionales para su transporte (mochila o soporte). Siendo lo más habitual encontrarse espectrómetros manuales en el rango Visible-NIR (350-1200). En la búsqueda se han encontrado ejemplos con interesantes prestaciones cabe destacar el espectrómetro “MicroNIR” de la marca Viavi con rango de trabajo NIR (950-1650), además este espectrofotómetro tiene un peso muy bajo (64 g.) y una amplia variedad de gadgets en función del tipo de medida que se quiere realizar. Otros ejemplos notables son el “QualitySpec Trek Portable Spectrometer” de ASD, el “ASP-NIR-M-Reflect” de Nirvascan o “Aura Handheld NIR” de Zeiss. De estos tres cabe destacar el primero con un amplio rango de longitudes de onda de 350-2500, pero con un peso de 2.5 kg. como contrapartida.

Con el fin de recabar más información de primera mano sobre los nuevos dispositivos NIR portátiles de pequeño tamaño, se ha asistido a una reunión comercial con la marca israelí "VIAMI". En esta reunión fueron presentados algunos de los dispositivos NIR portátiles de tamaño más pequeño que hay en el mercado. Se realizaron las correspondientes demostraciones. Las impresiones fueron positivas, dispositivos muy ligeros y suficientemente rápidos con la única desventaja de trabajar solo a longitudes de onda bajas del espectro NIR.

También se ha adquirido conocimiento relativo a los dispositivos de toma de muestras, tanto para sólidos como para líquidos. Así, se ha recopilado información del procedimiento de medida en función de la muestra, de los modos de medida (reflectancia y transmitancia) necesarios y de las ventajas y limitaciones de cada uno.

- Imagen Hiperespectral: La vigilancia tecnológica de la imagen hiperespectral se desarrollará en profundidad en la actividad 3.
- Luces LED y posibles estrategias de configuración: En los últimos años las luces LED emergen como una tecnología de fácil implementación en un marco productivo 4.0 por su gran versatilidad en su integración en sistemas electrónicos con la finalidad de automatizar tareas. ASINCAR ha estudiado su viabilidad en dos tipos aplicaciones: Mediante iluminación artificial mejorar la calidad y propiedades de hortalizas y frutas, y la utilización de luces LED en seguridad alimentaria y conservación de los alimentos, concretamente, el control y/o reducción de los microorganismos patógenos.

En el primer caso a diferencia de las tecnologías tradicionales de iluminación, las LED son capaces de producir distintos espectros de luz y emitir a diferentes longitudes de onda siendo posible su modulación con el fin de escoger la configuración óptima para el crecimiento de la planta específica tratada. Además, las LED son capaces de emitir combinaciones específicas de longitudes de onda, permitiendo la aplicación de distintas estrategias de iluminación que conduzcan a mejores y más rápidos resultados para agricultores e investigadores. Asimismo, hemos de destacar la gran reducción de consumo energético con respecto a las tecnologías de iluminación tradicionales, la emisión de menor calor radiante sobre las plantas y calor en general, así como la mayor vida útil de las LED. Finalmente, destacar que los sistemas LED son capaces de producir iluminación bajo un espectro amplio de temperaturas, lo que es interesante dados los cambios climatológicos entre las diferentes estaciones del año. Todas estas ventajas y propiedades han permitido que el uso de sistemas LED se esté desplegando a buena velocidad dentro del sector agrícola. Implican un consumo mínimo de energía adicional que contrarrestan con creces con los aumentos de productividad que catalizan. También hay que sumarle a todo esto sus precios asequibles.

Respecto al segundo punto decir que encontrar estrategias de conservación que extiendan la vida media de los productos tanto como sea posible, sin comprometer la seguridad y calidad es un factor crítico para las empresas del sector agroalimentario. Esto se debe a que la extensión de la vida media de los alimentos posibilita la reducción de los residuos generados cuando los alimentos sobrepasan la fecha de caducidad, incrementando el tiempo en el mercado del producto, así como poder llegar a destinos de distribución más lejanos. Aquí es donde entra en

juego la radiación ultravioleta (UV) con sus grandes ventajas como que no requiera la adición de químicos o calor y no es una tecnología cara. Además, los productos irradiados con la justa dosis de radiación UV no presentan cambios a nivel organoléptico, a diferencia de los métodos tradicionales. Como resultado de la búsqueda de transferencia de conocimiento entre la investigación y la empresa en este punto, ASINCAR participa en un proyecto sobre la aplicación de luces LED ultravioleta para aumentar la vida útil de carne de potro y ternera, utilizando espectroscopía Vis-NIR para verificar el estado de la carne.

- Sensores. También se ha estudiado la viabilidad de la implementación de sensores de temperatura, humedad, color, ... en diferentes partes de procesos productivos.

El otro ámbito de vigilancia tecnológica es el tratamiento masivo de datos. Las búsquedas y adquisición de conocimiento realizadas están altamente relacionadas con la espectroscopía NIR y la quimiometría. Esto se debe a que ASINCAR apuesta por contribuir a generalizar su uso en la industria agroalimentaria. El tratamiento de datos, la quimiometría, la espectroscopía NIR y el software informático usado se analizarán con detalle en la actividad 2.

La persona contratada es conocedora de técnicas quimiométricas y de aprendizaje automático. En el caso de la quimiometría y la tecnología NIR, la relación fue fácil de establecer, ya que su desarrollo histórico ha ido casi a la par, y conseguir información y programas informáticos que cubran una necesidad es relativamente rápido. El problema surge cuando se trata de aplicar las modernas técnicas de aprendizaje automático que debido a su novedad resulta más difícil encontrar información, especialmente aplicada a la tecnología NIR. Por eso se ha desarrollado una laboriosa revisión bibliográfica en la búsqueda de vínculos entre la espectroscopía NIR y técnicas de aprendizaje automático e inteligencia computacional como redes neuronales, máquinas de vectores de soporte o algoritmos genéticos. Siempre teniendo en mente la búsqueda de una mejora respecto a las técnicas de la quimiometría más clásica. Para lograr este objetivo se ha recurrido a métodos de difusión de conocimiento como libros y revistas como a otros métodos más en boga hoy en día como páginas web y blogs (muy usados por los investigadores de aprendizaje automático para difundir sus resultados y opiniones de la forma más rápida posible).

Como fruto de las capacidades del nuevo personal incorporado y de la búsqueda de las sinergias entre las técnicas de análisis de tratamiento de datos y las necesidades de la industria. ASINCAR ha sido subcontratado como colaborador externo para un proyecto sobre digitalización de una planta de elaboración de café. ASINCAR participa en el proyecto apoyando a las empresas participantes en aquellas tareas que requieren el concurso de conocimientos técnicos específicos en áreas que, a priori, no pertenecen a su dominio de negocio. En concreto, Análisis del proceso productivo y definición de puntos de control, análisis de elementos de sensorización y control, análisis de datos y desarrollo de modelos predictivos, validación de modelos predictivos en entorno operativo. Otro proyecto en el que participa ASINCAR subcontratado es un proyecto de digitalización de la planta piloto de ASINCAR. El objetivo fundamental de este proyecto es la construcción de una solución vertical de industria 4.0 para el sector alimentario que permita mejorar los procesos de elaboración mediante la captación, análisis y explotación de datos provenientes de un sistema automatizado con el triple objetivo de garantizar la seguridad alimentaria, mejorar la calidad sensorial y nutricional del producto de cara al consumidor y reducir los costes de producción.

Además todo este conocimiento adquirido en esta Actividad se ha utilizado para el desarrollo del “III Plan Estratégico 2017-2020 de ASINCAR”, especialmente para el “OBJETIVO Cualitativo 3.- Industria 4.0. Automatización, análisis de datos, monitorización on-line, seguridad alimentaria automatizada”.

Tal y como nos trasladó Dña. Begoña Cristeto Blasco, Secretaria General de Industria y Pyme, en el encuentro de Agrupaciones empresariales innovadora Celebrado en la sede del MINECO, la industria 4.0 no es una moda ni una sugerencia. Se trata de un cambio de paradigma obligado para la industria española, y las que queden fuera de él no podrán ser sostenibles. El término Industria 4.0 fue acuñado por el gobierno alemán para describir la fábrica inteligente, una visión de la fabricación informatizada con todos los procesos interconectados por Internet de las Cosas (IOT). El término Industria 4.0 conlleva muchos significados, pero los primeros avances en este ámbito han implicado la incorporación de una mayor flexibilidad e individualización de los procesos de fabricación. Se espera que, junto con los fabricantes de electrónica, la industria de alimentos y bebidas será pionera en la adopción de procesos de fabricación flexibles e individualizados. Son seis tecnologías las indispensables para la transición a la Industria 4.0:

- IloT y Sistemas Ciberfísicos – El concepto de IloT (Industrial Internet of Things) se refiere al uso de las tecnologías IoT en los procesos industriales. Los sistemas Ciberfísicos son todos aquellos dispositivos que integran capacidades de procesamiento, almacenamiento y comunicación con el fin de poder controlar uno o varios procesos físicos. Los sistemas Ciberfísicos están conectados entre sí y a su vez conectados con la red global gracias al paradigma IoT. Ver más.
- Fabricación aditiva, impresión 3D – Permite, entre otras cosas, la hiper personalización inherente a la Industria 4.0 y al concepto de servitización- no encarece el proceso porque permite fabricar productos, sin penalizar el coste, independientemente de si se tiene que fabricar un determinado número de piezas iguales o todas distintas. Además, hace mucho más sencillo producir lotes pequeños de productos, desde pequeñas piezas de maquinaria hasta prototipos. Big Data, Data Mining y Data Analytics – La cantidad de información que actualmente se almacena en relación a diferentes procesos y sistemas (tanto industriales como logísticos), servicios (ventas, conexiones entre usuarios, consumo eléctrico, etc.) o tráfico de datos (logs en routers y equipos, entre otros) resulta ingente e inmanejable de forma manual. El análisis de estos datos puede proporcionar información muy valiosa acerca del comportamiento de estos procesos; se pueden prevenir problemas en un determinado proceso industrial a través de la detección de resultados o medidas anómalas (sin la necesidad de haber definido previamente qué medida es o no es anómala) o determinar qué eventos están relacionados dentro de un proceso más complejo facilitando su gestión a través de la predicción, sabiendo de antemano que un evento desencadenará otro con cierta probabilidad. A partir de toda esta información se pueden realizar simulaciones que, además, permiten predecir qué recursos van a ser necesarios, pudiendo optimizar su uso de forma automática y proactiva anticipando los acontecimientos futuros.
- Inteligencia Artificial – Son necesarias herramientas y tecnologías que sean capaces de procesar en tiempo real grandes volúmenes de información que extraemos de las tecnologías Big Data, así como algoritmos capaces de aprender de forma autónoma a partir de la información que

reciben, con independencia de las Página 120 de 150 fuentes, y de la reacción de los usuarios y operadores (técnicas de Machine Learning, Deep Learning y Artificial Intelligence).

- Robótica Colaborativa (Cobot) – Este término define a una nueva generación de robots industriales que coopera con los humanos de manera estrecha, sin las características restricciones de seguridad requeridas en aplicaciones típicas de robótica industrial. Se caracteriza, entre otras cosas, por su flexibilidad, accesibilidad, y relativa facilidad de programación.
- Realidad virtual y Realidad aumentada – La mayor accesibilidad de estas tecnologías en los últimos años las ha hecho situarse como una herramienta útil para la optimización de los diseños, la automatización de los procesos, el control de la fabricación y la construcción, el entrenamiento y la formación de los trabajadores, y los trabajos de mantenimiento y de seguimiento.

La clave para el sector agroalimentario en general y para el sector cárnico asturiano en particular está ahora en diagnosticar el estado en que todas estas tecnologías se encuentran implementadas en las industrias y en que puede ayudar su incorporación. En el caso de las empresas miembro de ASINCAR AEI, parece evidente que aún queda mucho trabajo de dinamización y concienciación, incluso de hacer entender a las empresas a que se refiere este nuevo termino. En nuestra opinión, el punto de arranque de la implementación y desarrollo de estas tecnologías en nuestras industrias debería partir de la mano de dos conceptos:

- Optimización de procesos y mejora de la eficiencia y de la productividad. (Automatización de procesos)
- Obtención de un valor añadido en la fabricación de productos, por mejoras en los estándares de producción, control de calidad y de seguridad alimentaria. En este sentido la monitorización on-line de parámetro de fabricación, su captación de datos y el análisis de los mismos para la toma de decisiones debe ser el camino a seguir.

Abordar estos aspectos debe hacerse desde la cooperación entre ASINCAR AEI, las empresas miembro de la agrupación, y los organismos de conocimiento especializados en estos aspectos, como son Fundación PRODINTEC, y Fundación CTIC, por citar dos ejemplos. Sin lugar a dudas la consecución de este objetivo derivará en mayor oportunidad para acceder a mercados más exigentes y a nuevos nichos de mercado con productos diferenciados.



Figura 1. Relación entre los pilares y los objetivos clave definidos en la estrategia de ASINCAR AEI.

### Tarea 1.2. Evaluación y definición de la tecnología NIR como habilitador para Industria 4.0.

El uso de la espectroscopia de infrarrojo cercano (NIR) para la medición de características o parámetros de interés en la industria alimentaria se encuentra en una etapa crítica de su evolución ya que se está dando el paso de la escala de laboratorio/planta piloto a su uso industrial.

Los estudios científicos tanto nivel de laboratorio como a escala semi-industrial o de planta piloto han demostrado la capacidad del NIR en la predicción de los componentes principales de los alimentos (humedad, proteína, grasa, contenido en sólidos totales); parámetros relacionados con las cualidades tecnológicas de las materias primas o de alimentos (textura, peso, capacidad de retención de agua) o para la clasificación o caracterización de las materias primas (frescura en pescado, ciclos de congelación/descongelación). Estos estudios se han hecho en un amplio rango de alimentos tales como frutas, vegetales, carne y productos cárnicos, leche y derivados lácteos, cereales, etc. La gran mayoría de esta literatura científica comparte la conclusión de la gran potencialidad de esta tecnología para el control de la calidad de los alimentos.

Es importante señalar que los estudios respecto a la capacidad de predicción de parámetros relacionados con la calidad microbiológica (ej vida útil) o de seguridad alimentaria son mucho menores y sólo a nivel de laboratorio.

El paso a una escala industrial conlleva una serie de retos debido a la necesidad de cumplir una serie de necesidades, según el tipo de aplicación, en relación a su diseño (seguro frente explosiones, resistencia

al agua o la facilidad de limpieza). Además de ello, es conocido que el ambiente industrial es más agresivo que el ambiente controlado de un laboratorio: humedad, vibración, temperaturas altas o bajas y polvo pueden estar presentes en la industria y que pueden afectar a la capacidad analítica del equipo (sensibilidad, reproducibilidad, repetitividad, etc). Esto no ha impedido que ciertas aplicaciones en carne y leche (control de componentes como proteína, grasa), vino (grado de maduración de las uvas) y sobre todo a nivel de control de parámetros fisicoquímicos de pienso y grano se hayan ido aplicando en la industria.

Durante este proceso AEI ASINCAR detectó dos graves carencias en este proceso incipiente de introducción de la tecnología NIR en la industria agroalimentaria:

- No se contempla la monitorización de la calidad microbiológica mediante esta técnica.
- Es una tecnología alejada de la PYME, principal tipo empresa que forma la ASINCAR AEI y el tejido industrial agroalimentario español.

Respecto al primer punto AEI ASINCAR ha detectado que el foco de desarrollo de esta tecnología se centraba en aspectos fisicoquímicos dejando los aspectos microbiológicos, tan importantes en este tipo de industria, en un segundo plano. Así, y en consonancia con su Plan Estratégico 2012-2016, se comenzó a trabajar en esta línea consiguiendo desarrollar, junto a la colaboración una empresa de base tecnológica, aplicaciones para el control de la calidad microbiológica que ha permitido estimar la vida útil del producto adquiriendo una amplia experiencia en el desarrollo de la metodología necesaria de toma de muestra y del tratamiento de los datos. También, comentar que la detección de fraudes es una línea de trabajo especialmente activa para AEI ASINCAR, debido a la demanda actual de una sociedad concienciada de la importancia de conocer la composición de los alimentos que forman parte de su dieta, además de la alarma social provocada recientemente por los fraudes alimenticios derivados de hacer pasar unos productos por otros de menor calidad y precio.

Respecto al segundo punto existen dos limitaciones para la asunción de esta tecnología por parte de las PYMES. Por un lado, el coste del equipo y relacionado con esto, su tamaño.

Los fabricantes de NIR están comenzando a desarrollar equipos de sobremesa más pequeños y en algunos casos han desarrollado equipos de mano para otras industrias (minería). El gran hándicap de estos equipos es que sus características técnicas no se acoplan a las necesidades de la industria técnica tanto por exceso como por defecto. Así, algunos equipos presentan características técnicas excesivamente buenas lo que encarece y hace grande el equipo o en cambio existen equipos que sus características quedan muy cortas y los rangos espectrales de trabajo no son válidos para determinar la calidad microbiológica o físico-química de los alimentos (equipos de mano de la minería).

## **ACTIVIDAD 2: Desarrollo de aplicaciones para sensores VIS UV: Tratamiento masivo de datos generados por los sensores de NIR y análisis multivariante (quimiometría):**

### **Tarea 2.1. Definición de programas quimiométricos.**

Existen una gran cantidad de programas quimiométricos, tanto libres como de pago. A parte del dinero a desembolsar, la otra gran diferencia entre ambos es la facilidad de uso. Generalmente, mientras que los programas de pago suelen estar diseñados para un uso más directo que conlleve una curva de aprendizaje rápida, los programas libres suelen depender en gran medida de las habilidades y

conocimientos del usuario. Así, en un programa quimiométrico comercial están disponibles una serie de opciones prefijadas entre las que el usuario elige la que crea más conveniente. En el caso del software libre, el escenario suele ser diferente. Habitualmente se parte de un entorno de programación en el que existen paquetes informáticos de quimiometría que el usuario puede descargar e instalar. A partir de ahí el usuario configura la aplicación en función de sus habilidades, necesidades, conocimientos o preferencias personales. A modo de titulares se puede decir que generalmente los programas quimiométricos comerciales son de uso más directo y los programas quimiométricos libres son más personalizables. Una diferencia entre el software comercial y el libre es que las compañías de software comercial organizan habitualmente cursos para el manejo de su software, mientras que el uso del software libre suele estar asociado a hábitos más autodidactas. Entre los programas comerciales cabe destacar:

- Unscrambler de Camo.
- PLSToolbox (paquete de Matlab) de EVRI.
- WinISI de FOSS.
- Grams de Thermo Fisher Scientific.
- NIRCal de Büchi.
- OPUS Spectroscopy Software de Bruker.
- Pirouette de Infometrics.

Entre los programas libres cabe destacar:

- R.
- Octave.
- Python.
- Scilab.

En principio se ha optado por software libre por las siguientes razones aparte de las vistas anteriormente:

- Facilitar el acceso al código informático, visibilizando los datos y algoritmos y evitando que el software sea una caja negra como es habitual en los paquetes de software comercial.
- El uso de técnicas de machine learning. El software comercial viene con unos modelos prefijados a aplicar. Esto dificulta el uso de las modernas técnicas de machine learning. Un modelo como las redes neuronales, de uso habitual para ajuste funcional no lineal, es muy difícil de encontrar en el software comercial. Por eso a modo de anticipación de tendencias dentro del campo, el software libre permite estar a la par de desarrollos propios del campo de la investigación.

### **Tarea 2.2. Definición de técnicas de tratamiento masivo de datos.**

En todo modelo quimiométrico de espectroscopia NIR se pueden diferenciar dos etapas (Figura 2). Una primera etapa de creación del modelo quimiométrico y una segunda de utilización y trabajo con el modelo quimiométrico. En la primera etapa, se utiliza un conjunto de datos (longitudes de onda vs

propiedad física) lo suficientemente extenso y mediante técnicas de optimización se obtienen los parámetros del modelo. Una vez se tiene un modelo optimizado se puede pasar a la segunda etapa. En esta fase, cuando se le presentan al modelo espectros de infrarrojo cercano obtenidos con el espectrofotómetro y este devuelve los correspondientes valores predichos de propiedad químico-física o microbiológica de las muestras. Comentar que mientras la etapa de optimización puede llegar a ser costosa desde el punto de vista de recursos computacionales y de tiempo, en la segunda etapa ocurre lo contrario, ejecución rápida y sin gran necesidad de recursos computacionales. De esta forma, la primera etapa será la más compleja con diferencia de las dos. Esta etapa de optimización de parámetros del modelo consta generalmente de tres partes que dado su importancia se procederá a explicar en detalle:

- Pretratamiento de los espectros.
- Selección de longitudes de onda.
- Calibraciones.

El primer paso es aplicar a los datos de los espectros diversas técnicas de pretratamiento con la finalidad de acondicionar los espectrogramas para obtener un mejor resultado cuando se le apliquen las técnicas de calibrado. Aunque una clasificación de estos métodos es complicada dado que muchos de ellos se podrían incluir en varias categorías, se pueden definir las siguientes categorías de métodos utilizados:

- Suavizado: “Movil average”, “Median Filter”, “Savitzky Golay Smoother” (2 parámetros),...
- Normalización: “Area Normalization”, MSC (“Multiplicative Signal Correction”), EMSC (“Extended Multiplicative Signal Correction”), SNV (“Standad Normal Variate”), “Autoscale Normalization”,...
- Corrección de línea base y derivación: “Savitzky-Golay” (3 parámetros), “Gap-Derivative”, “Norris Gap” y “Detrend”,...

La segunda etapa de la creación del modelo quimiométrico está justificada porque la información útil relativa a los enlaces químicos implicados en el modelo quimiométrico, generalmente se encuentra localizada en determinadas zonas del espectro. Restringir los cálculos quimiométricos a determinadas zonas (selección de longitudes de onda) es una estrategia realmente útil, ya que eliminar las longitudes de onda donde no se encuentra información de utilidad ayuda a mejorar los resultados obtenidos con el modelo quimiométrico. Esto es especialmente palpable cuando se presentan nuevos datos (no utilizados en el cálculo de la calibración) al modelo, puesto que el modelo tiene una mejor generalización que en el caso de usar todas las longitudes de onda en el cálculo. Entre los métodos usados para la selección de longitudes de onda se encuentran:

- VIP (“Variable importance in projection”).
- Algoritmos genéticos.
- i-PLS (“interval-Partial Least Squares”).
- Gi-PLS (“Genetic interval-Partial Least Squares”), ...

Una vez se ejecutan los pasos de pretratamiento y selección de longitudes de onda, llega el turno de las calibraciones propiamente dichas. Lo que se logra mediante una calibración es la relación entre unas variables de entrada o dependientes (valores para las diferentes longitudes de onda) y otras de salida o

dependientes (propiedad químico-física o microbiológica). Las técnicas de calibración multivariante más populares son:

- MLR (“Multiple Linear Regression”).
- PCR (“principal component regression”).
- PLS (“partial least squares”).

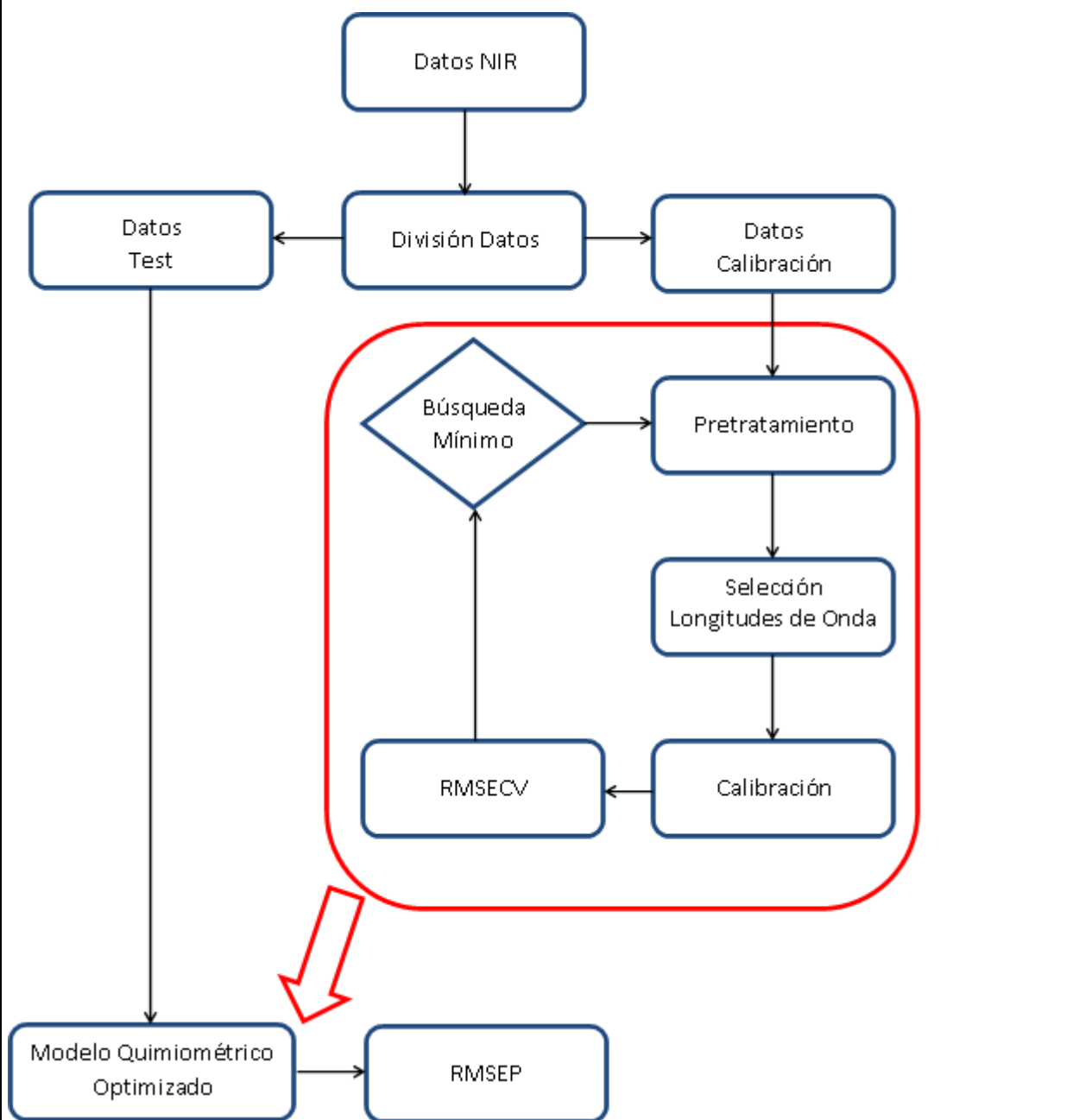


Figura 2. Diagrama del proceso de optimización del modelo quimiométrico.

Las dos últimas técnicas son especialmente útiles ya que conllevan una reducción de la dimensionalidad de los datos, siendo generalmente necesarias unas pocas variables latentes donde se agrupa la mayor parte de la varianza de los datos. Estos tres métodos se engloban dentro de los llamados métodos lineales. Las técnicas cualitativas de análisis multivariante (clasificación) más populares son:

- Análisis linear discriminante.
- Regresión Logística.
- SIMCA.
- Clasificación k-NN.

También existen los métodos no lineales. Estos métodos son computacionalmente más complejos y el tiempo requerido es sustancialmente superior al empleado en los métodos lineales. Generalmente con los métodos lineales se logran resultados satisfactorios, pero en ocasiones debido a la no linealidad de los datos, conviene usar métodos no-lineales que son capaces de abordar estos casos. Entre los métodos no lineales, los dos más ampliamente usados son:

- NN ("Neural Networks").
- SVM ("Support Vector Machines").

Estas dos técnicas son dos técnicas básicas dentro del "machine learning". Son aplicables tanto a análisis cualitativo como cuantitativo. Respecto a estas técnicas comentar que su manejo es sensiblemente más complicado que los métodos lineales, ya que por ejemplo para los métodos lineales se para un número dado de parámetros se presenta una solución para los parámetros unimodal (un único mínimo global), mientras que en el caso de los métodos no-lineales, lo habitual es que existan varios mínimos locales a parte del global. Estos métodos serán implementados mediante paquetes de software libre, ya que los paquetes comerciales como Unscrambler o Grams tienen ciertas limitaciones para estos métodos no lineales o no los tienen implementados (redes neuronales).

A parte de los pasos de pretratamiento, selección de longitudes de onda y calibración, También se deben tener en cuenta la existencia de datos espúreos, de forma que se implementen estrategias para su localización y criterios de eliminación.

Habitualmente se presentan dos problemas en las calibraciones, sobreajuste y sobregeneralización. El sobreajuste se presenta cuando el modelo comienza a memorizar los datos y con ello su ruido según transcurre la optimización. Esto es un grave problema ya que el modelo presenta una mala generalización con datos nuevos no usados en la creación del modelo. Este comportamiento es propio de modelos con demasiados parámetros, el modelo es demasiado complejo. En el lado contrario de la balanza se puede encontrar el subajuste o sobregeneralización, en este caso el modelo no es capaz de captar la tendencia subyacente. Esto suele ser debido a que el modelo no tiene suficientes parámetros, el modelo es demasiado simple. Debido a estos problemas conviene establecer estrategias para cuantificar el desempeño del modelo. Para ello se procede a dividir los datos en conjuntos para la validación del modelo quimiométrico. El procedimiento habitual consiste en dividir los datos en dos conjuntos de datos, entrenamiento y test. Una opción es que el conjunto de entrenamiento tenga entre un 70-80 % de los datos y el de test un 30-20 %, aunque eso depende en gran medida de los datos de que se disponga. Con los datos de entrenamiento se construirá el modelo quimiométrico. Para evaluar los resultados de la calibración se van a usar técnicas de validación cruzada con los datos de entrenamiento. Para ello se divide el conjunto de entrenamiento en "N" particiones y se realizan "N" optimizaciones de parámetros. Para ello en cada una de estas optimizaciones se utilizan todas las

particiones de datos menos una en la optimización de los parámetros del modelo quimiométrico, de forma que al final todas las particiones son utilizadas en validación. En este caso se considerarán dos tipos de validación cruzada, “K-fold cross validation” y LOOCV (“Leave-One-Out-Cross-Validation”), en la primera se crean “K” particiones del conjunto de entrenamiento iguales y en el segundo se crean tantas particiones como datos de forma que de cada vez solo se usa un dato en la validación. En ambos modelos se realiza la media aritmética de los resultados para obtener un único resultado.

Para cuantificar los diferentes tipos de errores que se pueden producir, es habitual encontrar en bibliografía tres estadísticos que describen el comportamiento del modelo (error) y están íntimamente relacionados con la división de datos en diferentes subconjuntos:

- RMESC (“Root Mean Square of Calibration”): Error cuadrático medio que presentan los datos de entrenamiento cuando son utilizados para optimizar los parámetros del modelo
- RMESCV (“Root Mean Square of Cross Validation”): Es el error cuadrático medio que presentan los datos de entrenamiento cuando son utilizados para validar los parámetros del modelo.
- RMESP (“Root Mean Square of Prediction”): Es el error cuadrático medio que presentan los datos de test que no son utilizados en ningún momento en la optimización del modelo.

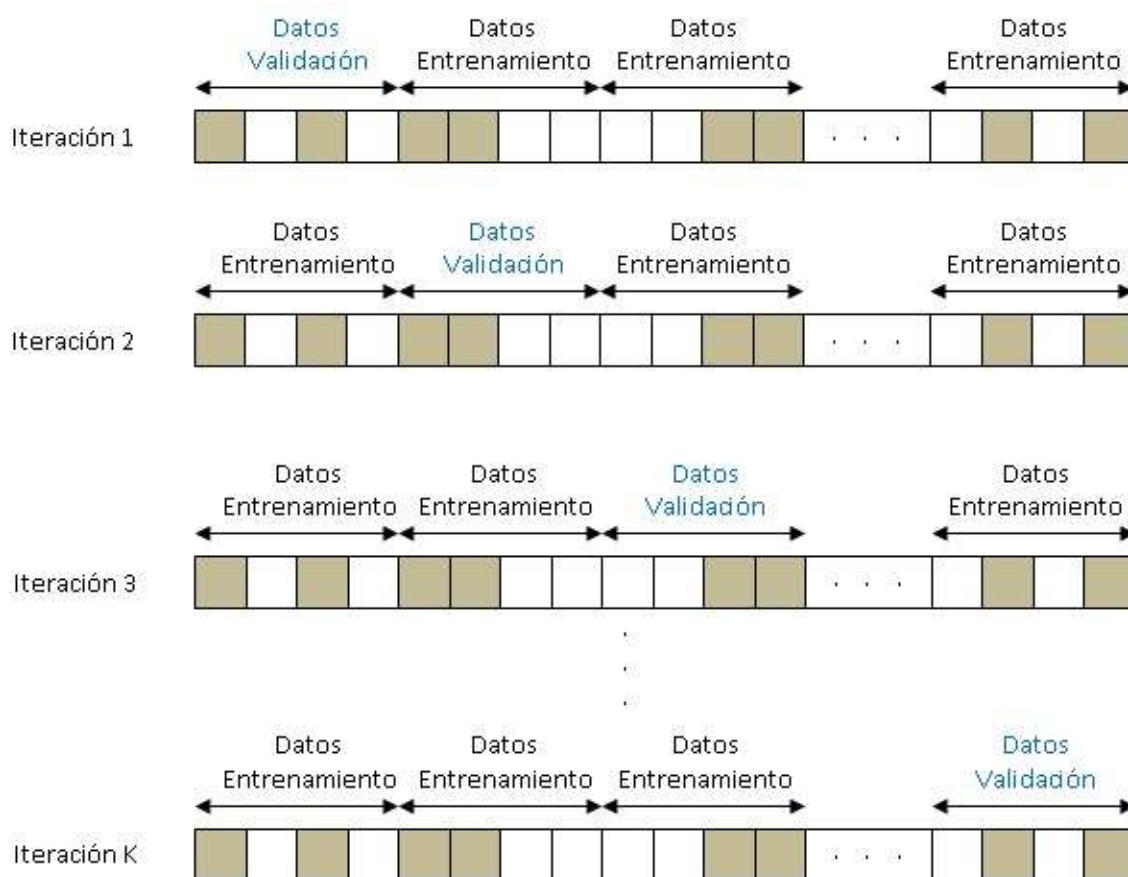


Figura 3. Diagrama del método de validación k-fold CV, el método LOOCV es igual solo que cada dato es una partición en si misma.

En estos estadísticos se pueden observar diferentes comportamientos del modelo quimiométrico. Por

ejemplo, tener un RMSEC bajo quiere decir que el modelo ajusta bien los datos de entrenamiento cuando se usan en la optimización de parámetros. Con esto hay que tener especial cuidado porque si el RMSECV (datos de entrenamiento en validación) es alto quiere decir que tenemos sobreajuste en los datos. La situación óptima es que estos dos valores sean bajos y similares. El RMSEP es una validación del modelo totalmente externa (aquí es donde se utilizan los datos del conjunto de test). Conviene hacer esta validación, ya que con la validación cruzada se produce cierto sesgo en el RMSECV del modelo. Así con datos totalmente externos se tiene una idea más aproximada del error de predicción del modelo. Así mediante estos estadísticos se pueden elegir los modelos quimiométricos con mejor desempeño entre la gran variedad de opciones que se presentan para los diferentes métodos de pretratamiento, selección de longitudes de onda y calibraciones.

Una vez se tiene el modelo quimiométrico (valores longitudes de onda vs propiedad física), se puede utilizar junto con espectros conseguidos con el espectrofotómetro para predecir propiedades físicas de la muestra, de forma que se pueden monitorizar/analizar la calidad (parámetros físico-químicos y microbiológicos). Para ello primero se ejecuta el pretratamiento escogido, luego se eligen las longitudes de onda y se procede a calcular el valor de propiedad física a partir de la calibración ya optimizada.

Con esta hibridación entre lo físico y digital, con toda esta información exhaustiva y estos algoritmos de cálculo, se busca obtener información con valor y dotar a los procesos de inteligencia autónoma, de forma que se logre una mayor eficiencia y tiempos de respuesta menores.

### **Tarea 2.3. Desarrollo de calibraciones y validaciones de las aplicaciones.**

En la duración del contrato de la persona contratada se han realizado diversas calibraciones. Para todas ellas se ha realizado un diseño de experimento, se ha establecido como se debe realizar la toma de muestras, se han realizado las calibraciones pertinentes y se han validado los resultados.

Dado que la industria cafetera tiene una gran importancia en el Principado de Asturias, se ha empezado a estudiar posibilidades a este respecto. Se ha desarrollado una calibración para obtener el grado de Café Natural/torrefacto a partir de medidas NIR. Se han distribuido las mezclas de café natural/torrefacto por todo el intervalo que va desde café natural puro hasta café torrefacto puro. Como resultado se obtiene un error inferior al 4% en la estimación del porcentaje de mezcla.

También, se han realizado calibraciones de espectros NIR de huevos para obtener propiedades físicas como la altura del albumen, color de la Yema, grosor de la cascara, resistencia de ruptura, peso de la cáscara. Se realizaron tres medidas en el ecuador del huevo y dos en los polos. Se han buscado cuales proporcionaban los mejores resultados para cada caso y los resultados obtenidos son parecidos a la bibliografía.

Otra línea donde se ha aplicado la espectroscopía NIR con las correspondientes calibraciones es en la determinación de la existencia de biofilms en plásticos. Se incubaron cepas de 3 tipos de microorganismos (salmonella, listeria y pseudomonas) en placas de plástico de forma que se escanea la zona central de la placa, que es donde se forma la mayor cantidad de biofilm y la superior donde generalmente no se forma. Después se tinte la placa y se puede diferenciar a simple vista la presencia o ausencia de biofilm. Por métodos de clasificación quimiométrica se busca un método que diga si existe o no biofilm a partir de espectroscopía NIR. El número de aciertos ronda el 90%.

También se ha empezado el trabajo en otras aplicaciones de la tecnología NIR y la quimiometría, en donde a pesar de estar en fase de realización se han obtenido resultados esperanzadores. Algunos casos son la determinación de los grados Brix (Sólidos solubles) en kiwis, parámetro muy relacionado con la maduración de los kiwis. Determinación del pH y actividad de agua en chorizos. Se ha empezado a realizar análisis de mieles del Principado de Asturias para en un futuro evitar posibles fraudes. También se ha empezado a analizar atunes para evitar fraudes derivados de la especie.

### **Actividad 3: Transición a la imagen hiperespectral (HSI).**

#### **Tarea 3.1 Desarrollo vigilancia tecnológica.**

En ASINCAR, el 2017 supuso un punto de partida hacia la imagen hiperespectral. Por eso se han realizado grandes esfuerzos en esa dirección. Se han realizado minuciosas búsquedas bibliográficas, se ha buscado en internet información de todo tipo como fundamentos, aplicaciones, empresas suministradoras, empresas desarrolladoras de software, seguimiento de investigadores de vanguardia, ... También se ha contactado por teléfono y skype con los principales fabricantes de cámaras de imagen hiperespectral con la finalidad de conocer sus productos, con sus posibilidades y limitaciones, así como posibles aplicaciones. También se les ha transmitido nuestras inquietudes en lo referente al sector agroalimentario y explorar sinergias beneficiosas para ambas partes.

Otras actuaciones concretas es la asistencia al congreso "*CHII2017: Conference on Hyperspectral Imaging in Industry*" que tuvo lugar en Gratz (Austria) de 7 al 8 de junio de 2017. Este es el congreso más importante desarrollado en Europa cuando se habla de imagen hiperespectral aplicada a la industria en el que se dan cita los principales fabricantes de imagen hiperespectral así como empresas con desarrollos funcionales. Esta conferencia ha servido como punto de referencia para conocer en detalle el punto en que se encuentra la industria y hacia dónde se dirige. Además, ha servido para hacer contactos con las empresas punteras del sector como Specim, Hypex, Senop Oy, Xenics, ESPROS o EVK.

De estas empresas se ha mantenido contacto con Specim y Hypex, para en un futuro poder presentar proyectos conjuntos. En concreto con Specim se busca una cámara hiperespectral móvil. Para ello a parte de los correspondientes contactos telefónicos se ha mantenido una reunión en Madrid con el fin de buscar sinergias e intereses comunes. Los contactos con Hypex se mantienen con la finalidad de la búsqueda de una solución fija. Siempre a todos estos contactos se les ha transmitido nuestras inquietudes en lo referente al sector agroalimentario y la necesidad de la búsqueda de soluciones beneficiosas para ambas partes.

Siempre se debe tener en mente que la imagen hiperespectral es una tecnología altamente novedosa que se encuentra en los albores de su uso industrial. Por eso la búsqueda de vínculos entre los diferentes actores es tan importante. Así un problema común que comentan todos los fabricantes del sector de la imagen hiperespectral es la barrera que supone el desconocimiento de los sectores de aplicación de su tecnología. Generalmente estos fabricantes conocen su producto a la perfección, como es lógico. Pero desconocen cómo aplicarlo a los diferentes procesos productivos, debido a las especificidades y limitaciones del día a día que se presentan.

#### **Tarea 3.2 Evaluación y definición de la Tecnología HSI como habilitadora para la industria 4.0.**

La imagen hiperespectral es la evolución natural de los espectrómetros clásicos frente a los problemas

surgidos por las características de la medida. Cuando se realiza una medida NIR, únicamente se obtiene información de un punto o un promedio de una pequeña zona. Esto es un problema cuando la muestra de medida presenta heterogeneidad, ya que una única medida no es representativa de toda la muestra. Como ejemplo en el sector agroalimentario tenemos el caso de la carne con zonas donde hay solo magro, otras solo con grasa y todo tipo de escenarios donde hay magro y grasa.

Otro aspecto a tener en cuenta es que se pueden encontrar dispositivos de imagen hiperspectral desde el UV hasta el NIR, abriendo un abanico realmente grande de posibles aplicaciones. En concreto en el caso del NIR, los fabricantes hablan de imagen química. Esto se debe a que es habitual que la imagen hiperspectral después de los correspondientes análisis quimiométricos proporcione “mapas” de concentración de diversos compuestos químicos o propiedades físicas en la muestra. Por ejemplo, se puede utilizar para ver distribución de agua en una amplia variedad de alimentos.

También se han encontrado aplicaciones en clasificación de alimentos. Un caso paradigmático es la detección de fraudes alimentarios, como por ejemplo se podría utilizar para detectar tipo de especie para evitar fraudes en la venta de pescado. Otra utilización sería detectar impurezas en la materia prima a la entrada de un proceso productivo. Así, por ejemplo, se podrían identificar para luego separar piedras que puedan venir dentro de un lote de cereal. Por último, otra de las aplicaciones donde la imagen hiperspectral se vuelve especialmente útil es cuando se trata de detectar imperfecciones. Esto es especialmente útil en el caso de la fruta, ya que es habitual que con la utilización de longitudes de onda NIR se puedan diferenciar más claramente las zonas de defectos.

Además, debido a su inherente facilidad de conectividad y que la inclusión en un entorno altamente digitalizado está siempre contemplada por los fabricantes, convierte a estos dispositivos en firmes candidatos a asentarse en las líneas de producción de las fábricas involucradas en la revolución 4.0. Otro aspecto en línea con la industria 4.0 es la necesidad de tratamiento de una gran cantidad de datos. Las cámaras de imagen hiperspectral generan gran cantidad de datos debido a que generan una imagen de la muestra analizada para cada longitud de onda analizada. Así es habitual encontrar archivos de imagen con 512x512 y 400 longitudes de onda, con lo cual este archivo tiene más de 100 millones de vóxeles (pixels en 3D). De ahí la necesidad y la vinculación con el tratamiento masivo de datos.

De esta forma, la imagen hiperspectral presenta una fuerte vinculación con varios aspectos de la industria 4.0 como tratamiento masivo de datos, fácil implementación en un entorno digital y el ser un habilitador tecnológico que aporta nuevas soluciones a la industria agroalimentaria.

Como conclusión global a toda la memoria, la persona contratada ha realizado los siguientes trabajos:

- Participar en la elaboración del Plan Estratégico ASINCAR 2017-2020, y verificación de concordancia con las actividades relacionadas con la Industria 4.0.
- Confección del Informe viabilidad tecnología Hiperspectral, donde se analiza la viabilidad de la imagen hiperspectral a la industria agroalimentaria. También, definió la estrategia de incorporación de la imagen hiperspectral a la empresa, a través de contactos con fabricantes de imagen hiperspectral y definición de posibles aplicaciones de la tecnología al sector.
- Desarrollo de aplicaciones para sensores Vis-NIR para la monitorización y medida de parámetros relacionados con la calidad microbiológica, calidad fisicoquímica y detección de fraudes. Para

ello:

- Se desarrollaron herramientas informáticas usando software libre para poder realizar las calibraciones quimiométricas,
- Se realizaron diseños de experimento para realizar posteriormente las calibraciones, se realizaron las medidas espectrométricas,
- Se realizaron las calibraciones con su correspondiente análisis de datos y se
- Se verificaron las calibraciones con datos externos que no intervinieron en el cálculo de las calibraciones.

## 2. ACTIVIDADES SUBCONTRATADAS

- 2.1 Especifique las actividades concretas realizadas por cada una de las entidades subcontratadas (si las hubiere), indicando si se corresponden con las previstas en la solicitud o si se han producido desviaciones.
- 2.2 Valore el grado de cumplimiento de las entidades subcontratadas en la ejecución de sus actividades.

No procede.

## 3. HITOS ALCANZADOS. GRADO DE CUMPLIMIENTO

- 3.1 Indique los diferentes hitos alcanzados, relacionándolos con la planificación inicial del proyecto. Detalle los indicadores de seguimiento y entregables obtenidos en el periodo justificado.
- 3.2 Declare el % estimado de cumplimiento en cada uno de ellos.

### **Hito 1. Adecuación de ASINCAR a los objetivos de industria 4.0**

Grado de cumplimiento respecto a lo planificado en 2017: 100 %

Grado de cumplimiento respecto a lo planificado en la ejecución total de proyecto: 100 %

**Indicadores de seguimiento obtenidos:** Búsqueda de Sensores compatibles con la industria 4.0., priorizando la búsqueda de dispositivos NIR comerciales compatibles con la industria agroalimentaria (especialmente dispositivos de pequeño tamaño). Establecer procedimientos de cálculo de métodos quimiométricos. Potencialidades y problemas del binomio entre la tecnología NIR y la industria agroalimentaria.

**Entregable obtenido:** Plan Estratégico ASINCAR 2017-2020.

### **Hito 2. Desarrollo de aplicaciones para sensores VIS UV: Tratamiento masivo de datos generados por los sensores de NIR y análisis multivariante (quimiometría)**

Grado de cumplimiento respecto a lo planificado en 2017: 100 %

Grado de cumplimiento respecto a lo planificado en la ejecución total de proyecto: 100 %

**Indicadores de seguimiento obtenidos:** Búsqueda y establecimiento de métodos quimiométricos. Pros y contras del software quimiométrico, tanto comercial como libre. Desarrollo de diferentes calibraciones relacionadas con la industria agroalimentaria.

**Entregable obtenido:** Aplicación NIR para industria alimentaria.

**Hito 3. Transición a Imagen Hiperespectral (HSI)**

Grado de cumplimiento respecto a lo planificado en 2017: 100 %

Grado de cumplimiento respecto a lo planificado en la ejecución total de proyecto: 100 %

**Indicadores de seguimiento obtenidos:** Búsqueda de información de imagen hiperespectral, especialmente la relacionada con el sector agroalimentario. Establecer contactos con principales fabricantes de cámaras hiperespectrales. Comprender y evaluar las posibilidades que ofrece la imagen hiperespectral.

**Entregable obtenido:** Informe viabilidad HSI.

**4. RESULTADOS CONSEGUIDOS**

Enumerar brevemente, pero con claridad y precisión, los resultados obtenidos, haciendo referencia expresa a los resultados inicialmente previstos.

RESULTADOS INICIALMENTE PLANTEADOS	RESULTADOS FINALES OBTENIDOS
Plan Estratégico ASINCAR 2017-2020, y verificación de concordancia con las actividades relacionadas con la Industria 4.0	Plan estratégico finalizado y acorde con las actividades relacionadas con la industria 4.0.
Desarrollo de aplicaciones para industria agroalimentaria.	5 calibraciones realizadas (diseño de experimento, medida NIR, cálculos quimiométricos y validación) relacionadas con la industria agroalimentaria. Dos más en curso.
Informe viabilidad tecnología Hiperespectral.	Entregado. Posibilidades y limitaciones de la imagen hiperespectral comprendidas. Se está en disposición de introducir la imagen hiperespectral en la industria agroalimentaria.