



**Universidad del País Vasco/Euskal Herriko
Unibertsitatea**

**Facultad de Ciencias Químicas/Kimika Zientzien
Fakultatea**

Grado en Química

TRABAJO FIN DE GRADO

**Cuantificación de colorantes alimentarios por análisis de imagen.
Determinación de rojo allura en muestras en presencia de tartrazina.**

Autor/a: Gorka Albizu Martínez

Dirigido por: Mainer Vidal Postigo

San Sebastián, Julio de 2017

Índice

| | |
|---|----|
| 1. Resumen/Abstract..... | 4 |
| 2. Introducción | 6 |
| 2.1. Colorantes alimentarios | 6 |
| 2.1.1. Tartrazina (E 102) | 9 |
| 2.1.2. Rojo Allura (E 129) | 9 |
| 2.2. Análisis de imagen digital | 11 |
| 2.3. Quimiometría | 14 |
| 2.3.1. Análisis de componentes principales | 16 |
| 2.3.2. Calibración multivariable..... | 17 |
| 2.4. Espectrofotometría ultravioleta-visible | 17 |
| 3. Objetivo | 19 |
| 4. Experimental | 20 |
| 4.1. Preparación de muestras y optimización del pH..... | 20 |
| 4.2. Análisis de imagen..... | 21 |
| 4.2.1. Aparatos y softwares..... | 21 |
| 4.2.2. Muestras empleadas | 22 |
| 4.2.3. Matriz de calibración y de validación | 22 |
| 4.2.4. Toma de las imágenes | 23 |
| 4.2.5. Procesamiento de la imagen | 25 |
| 4.3. Espectrofotometría UV-Vis..... | 27 |
| 4.3.1. Aparatos y softwares..... | 27 |
| 4.3.2. Muestras empleadas | 28 |
| 4.3.3. Matriz de calibración y de validación | 28 |
| 4.4. Cálculo de los errores de los modelos de calibración multivariable | 29 |
| 4.5. Cálculo de los límites de detección | 29 |

| | |
|--|----|
| 5. Resultados y discusión..... | 30 |
| 5.1. Optimización del pH | 30 |
| 5.2. Análisis de imagen..... | 31 |
| 5.2.1. Patrón de color, medida de la estabilidad de la imagen | 31 |
| 5.2.2. Influencia de los pocillos con muestra sobre pocillos vecinos | 33 |
| 5.2.3. Modelos de calibración para los colorantes a partir de las imágenes | 35 |
| 5.3. Espectroscopía ultravioleta-visible..... | 44 |
| 5.4. Comparación del método del análisis de imagen con el de referencia | 48 |
| | |
| 6. Conclusiones/Conclusions | 50 |
| | |
| 7. Referencias | 52 |

1. Resumen

En este trabajo de fin de grado se ha determinado el colorante rojo allura (E 129) en presencia de tartrazina (E 102) mediante análisis de imagen. Las técnicas que normalmente se emplean en la determinación de colorantes son el HPLC o la espectrofotometría. Estas técnicas necesitan de más tiempo y son más caras que la que aquí se describe. Para la toma de imágenes se hace uso de un escáner de oficina, y para su posterior tratamiento se emplea el software *MatLab*. En el escáner se colocan placas de microtitulación, cuyos pocillos se rellenan con tan solo 400 μ L de muestra, y que permiten medir más de una muestra en la misma imagen.

Una vez obtenidos los parámetros deseados de la imagen se ha realizado un modelo PLS que ha relacionado estos parámetros con las concentraciones de E 129. Los errores obtenidos han sido del 5,7%, 10,2% y 12,4% para la calibración, la *cross-validation* y la validación externa respectivamente. La precisión se ha evaluado por medio de la reproducibilidad y la repetibilidad, obteniéndose valores de RSD% menores al 14,2%. En cuanto a la exactitud, esta se ha medido por medio de las recuperaciones en muestras reales del E 129, lográndose valores entre 90,6% y 106%. El límite de detección obtenido es de 1,8 mg/L.

Como técnica de referencia se ha empleado la espectrofotometría UV-Vis. Los resultados con esta técnica han sido mejores a los obtenidos por medio del análisis de imagen. A pesar de ello, con el test de la elipse, en un 95% de confianza, se demuestra que los resultados con ambas técnicas son comparables.

1. Abstract

In this degree final project the allura red (E 129) dye is determined in presence of tartrazine (E 102) by using image analysis. Techniques that are normally used in the determination of food dyes are HPLC or spectrophotometry. These techniques need long time of analysis and are more expensive than the image analysis technique. An office scanner is used for the image acquisition, and *MatLab* software for image processing. Microtitration plates are placed on the scanner window, and wells are filled up with 400 μL of sample. This allows to measure more than one sample in the same image.

After the image parameters are obtained, a PLS model is done to relate these parameters to E 129 concentrations. This model shows errors of 5,7%, 10,2% and 12,4% for calibration, cross-validation and external validation respectively. The precision is evaluated by means of reproducibility and repeatability, and RSD% values lower than 14,2% were obtained. Recovery of E 129 in real samples is measured in order to evaluate the accuracy of the method. Results between 90,6% and 106% were observed. The LOD obtained was 1,8 mg/L.

UV-Vis spectrophotometry is used as a reference technique. The results with this technique are better than the ones obtained with image analysis. However it is demonstrated that the results by both techniques can be comparable if a 95% confidence ellipse test is accomplished.

2. Introducción

2.1. Colorantes alimentarios

Según la Agencia Española de Consumo, Seguridad Alimentaria y Nutrición (ACOESAN) *“los aditivos son sustancias que se añaden a los alimentos con un propósito tecnológico (para mejorar su aspecto, textura, resistencia a los microorganismos, etc.) en distintas etapas de su fabricación, transporte o almacenamiento”*^[1].

Los aditivos alimentarios se dividen en 27 grupos, y los dos más importantes son los que corresponden a los colorantes y edulcorantes. El color es la primera característica por la que se juzga un alimento, y su calidad y sabor tienen una estrecha relación con él^[2]. Por lo tanto los colorantes se añaden con la intención de obtener alimentos con colores que atraigan a los consumidores, o para que el alimento tenga un color relacionado con su sabor. Un ejemplo de esta aplicación podría ser una chuchería con sabor a fresa, en cuyo caso se añadiría un colorante que le diera un color rojo.

Todo aditivo que se añada a un alimento debe de estar señalado en la lista de los ingredientes, y para ello en la Unión Europea (UE) existe un código, llamado código E. En él se pone la letra E, de Europa (lo que garantiza que el aditivo ha pasado los controles de seguridad de la UE), seguida de tres o cuatro dígitos (código INS), los cuales hacen referencia a uno de los muchos aditivos que existen y normalmente se agrupan según al grupo al que pertenecen (Ej. Manitol, un edulcorante que potencia el sabor dulce de los alimentos: E 421)^[3]. En el caso de los colorantes alimentarios, su uso es muy común en muchos alimentos como por ejemplo en dulces (Figura 1), todos ellos se nombran como E 1XX, donde el número 1 indica que pertenecen al grupo de los colorantes, la segunda cifra correspondería al color que proporciona ese aditivo, y finalmente, con el último dígito se especifica cual es:

E 104 → 1. Colorante; 0. Color amarillo; 4. Amarillo de quinoleína

E 132 → 1. Colorante; 3. Color azul; 2. Indigotina



Figura 1. Uso de colorantes con colorantes en dulces^[4].

Como cualquier otro producto que se añade a los alimentos, los colorantes están sometidos a un control muy estricto de los efectos secundarios que puedan causar al consumidor y en muchas ocasiones es tema de debate el potencial peligro de su empleo. Este es el caso de los colorantes azoicos, que se caracterizan por tener un mismo grupo funcional (Figura 2) derivado del diazeno (HN=NH), donde se sustituyen los H por grupos hidrocarbilo, en su mayoría sustituyentes aromáticos [5].

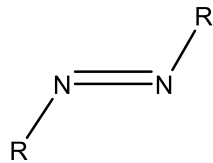


Figura 2. Estructura general de los compuestos azo.

El uso de estos colorantes azoicos es tema de debate especialmente porque muchos estudios demuestran que pueden ser cancerígenos al llevar a cabo, la microbiota intestinal, la azorreducción de estos a metabolitos cancerígenos, lo que depende de la cantidad ingerida. Es por esto que se busca disminuir la ingesta de este tipo de colorantes a pesar de que la absorción de estos en nuestro metabolismo es muy baja y por lo tanto no se metabolizan, haciendo que su toxicidad generalmente sea baja [3]. En cambio, existen estudios que confirman que los colorantes azoicos forman naftilaminas sulfonadas como productos metabólicos, los cuales son mutágenos [6].

Otros estudios también reflejan la capacidad que tienen estos colorantes para causar otros problemas; como son los trastornos de conducta en los niños o las infecciones respiratorias y alergias. En concreto, hay indicios de que este tipo de colorantes pueden desarrollar trastornos de déficit de atención e hiperactividad (TDAH) en los niños. En un estudio en el año 2004, llevado a cabo por Schab y Trinh [7], no se pudo observar la relación entre esta enfermedad y los colorantes azoicos, pero a los pocos años, en el 2007, McCann et. al. [8] hicieron un gran ensayo clínico con grupos de niños de tres y ocho años para comprobar la sospecha. En los resultados se pudo ver que había una relación estadísticamente significativa entre la ingesta de colorante y el desarrollo de TDAH, y por lo tanto la EFSA (Autoridad Europea de Seguridad Alimentaria) realizó un análisis detallado de este estudio, llegando a la conclusión de que los resultados no eran concluyentes por diferentes limitaciones del estudio, dejando sin resolver la relación entre colorantes azoicos y el TDAH. Finalmente, en el año 2012, un nuevo estudio, esta vez de Nigg et. al. [9], reveló que el 8% de los niños que sufren esta enfermedad mejoran al disminuir el consumo de alimentos y bebidas con este tipo de colorantes [3]. Es por ello que hoy en día todo alimento que contenga estas sustancias debe señalarlo y avisar de que “pueden tener un efecto adverso sobre la actividad y la atención de los niños”. Estudios posteriores llevados a cabo en animales demuestran que estos aditivos también pueden tener efectos en el

desarrollo neurológico, y que además su efecto puede verse aumentado en presencia de otro colorante (sinergismo) [3].

Como se ha dicho anteriormente, también se han llevado a cabo investigaciones del posible efecto que pueden tener estos colorantes sobre el sistema respiratorio y la aparición de alergias [3], alergias que pueden desencadenar episodios de urticaria. Estos síntomas se ven mejorados en un 30% al evitar consumir alimentos que puedan llevar estos aditivos. El 15% de la población que es sensible a la aspirina y el 8% de los pacientes que sufren de urticaria crónica intratable son sensibles también a la tartrazina [10].

Por todo esto, existe una legislación y una normativa que se tiene que cumplir a la hora de utilizar estos aditivos, y un control estricto. Desde el año 1962 existen normas y organizaciones destinadas al control de los alimentos con la creación del Codex Alimentarius (Código Alimentario) por parte de la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, *Food and Agriculture Organization of the United Nations*) y la Organización Mundial de la Salud (OMS). A partir de entonces, se ha creado el Comité del Codex sobre Aditivos Alimentarios y Contaminantes de los Alimentos (CCFAC, *Codex Committee on Food Additives and Contaminants*), que se centra en el control de los aditivos alimentarios y los contaminantes, para que cumplan la regulación. En lo que a la legislación alimentaria se refiere, en todos los países miembros de la Unión Europea se tiene que cumplir el REGLAMENTO (CE) N° 178/2002 [11] del Parlamento Europeo y del Consejo, del 28 de enero de 2002, donde se hace referencia a los aditivos, indicando que los que se utilicen deben ser revisados por primera vez por la EFSA, autoridad que se basa en las recomendaciones del CCFAC [3]. EFSA recoge los métodos oficiales a la hora de determinar los aditivos y siempre y cuando se tengan los datos suficientes se da también el valor de la ingesta diaria aceptable (ADI *Acceptable Daily Intake*), que es la cantidad máxima que una persona puede consumir de dicho aditivo diariamente durante toda su vida sin sufrir ningún riesgo para su salud, y se expresa en mg por kg de peso corporal (pc) y día [12]. Posteriormente, en el año 2008 en el REGLAMENTO (CE) N° 1333/2008 [13] del Parlamento Europeo y del Consejo se estableció una lista de aditivos alimentarios, indicando su toxicidad, y con la idea de que la EFSA los volviera a evaluar. Finalmente se elaboró el REGLAMENTO (UE) N° 1129/2011 DE LA COMISIÓN [14], que es con el que se regula actualmente el uso de los aditivos, y que modifica el anexo II del reglamento 1333/2008.

A modo de resumen, se puede decir que existe una gran controversia sobre los efectos que estos colorantes pueden producir en nuestro organismo, ya que hay muchos trabajos que encuentran indicios de que estos aditivos causan los efectos adversos ya nombrados, mientras que otros afirman no encontrar una relación entre estos efectos y el consumo de colorantes azoicos.

En este trabajo en concreto se han determinado dos de los muchos colorantes azoicos cuyo uso está permitido. Estos son la Tartrazina (E 102) y el Rojo Allura (E 129).

2.1.1. Tartrazina (E 102)

La tartrazina (5-hidroxi-1-(4-sulfonatofenil)-4(4-sulfonato-fenilazo)-3-carboxilato-1H-pirazol trisódico), cuya estructura se puede ver en la figura 3, es uno de los colorantes más utilizados en la industria alimentaria, y como consecuencia de ello, también uno de los aditivos más estudiados, investigados y controlados. Se utiliza para darle a los alimentos un color amarillo, y se emplea en postres y dulces, como pueden ser las gominolas o las gelatinas, en bebidas, en snacks (cuyo ejemplo podrían ser las patatas fritas), y como condimento, como por ejemplo en la paella, donde se utiliza para darle ese color amarillo tan característico ^[15].

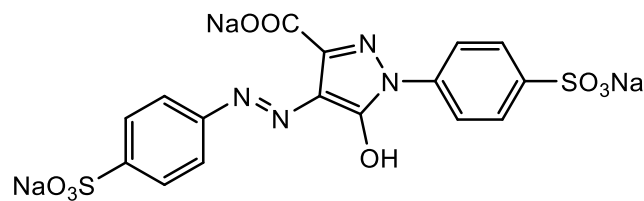


Figura 3. Estructura química de la tartrazina.

La tartrazina, además del nombre IUPAC y su nombre común, también se nombra de acuerdo al código E con el nombre de E 102, ya que es un colorante que proporciona color amarillo, y en EEUU recibe el nombre de FD&C Amarillo N°5. Es un sólido muy soluble en agua, y actualmente su valor de ADI se encuentra entre los 0 y los 10 mg/kg pc/día. Su toxicidad viene estudiándose desde la década de los 60, y en estudios realizados en el año 2009 se decidió que la ADI para este aditivo estaba entre 0 y 7 mg/kg pc/día, y posteriores estudios así lo confirmaron. Pero finalmente, en el año 2016, la dosis máxima diaria fue aumentada, hasta alcanzar el valor actual debido a que no había evidencias claras de los efectos adversos que podía provocar la tartrazina ^[6]. Existen muchos estudios e informes que hablan sobre la peligrosidad de este colorante, causando diferentes problemas, como pueden ser la activación de receptores de estrógeno que pueden provocar cirrosis biliar primaria en mujeres posmenopáusicas, o la capacidad de dañar los linfocitos humanos y unirse directamente al ADN, aunque en dosis altas, además de los supuestos efectos que se han explicado anteriormente ^[3].

2.1.2. Rojo Allura (E 129)

Este colorante, cuyo nombre común es rojo allura (2-hidroxi-1-(2-metoxi-5-metil-4-sulfonato-fenilazo)-naftaleno-6-sulfonato disódico), en la UE se nombra sistemáticamente con el nombre de E 129, mientras que en EEUU se conoce como FD&C Rojo N°40, y como este último nombre indica, se utiliza para dar colores rojos a los alimentos. En la figura 4 podemos ver su estructura química.

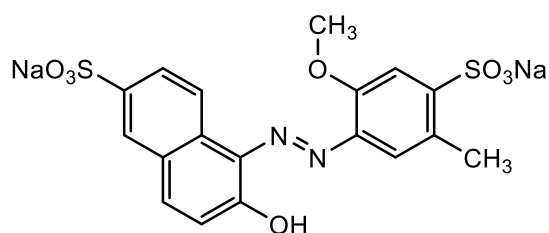


Figura 4. Estructura química del rojo allura.

Este también es un sólido soluble en agua, pero en este caso tiene un color muy oscuro. Es muy empleado en la industria de la repostería y en la fabricación de gominolas, alimentos muy relacionados con los niños. Este colorante fue estudiado por primera vez en la década de los 80, y más adelante en diferentes ensayos, obteniéndose unos resultados que se tradujeron en un valor de ADI de entre 0 y 7 mg/kg pc/día por parte de la EFSA, valor que en este caso si se ha mantenido en la última revisión de este organismo en el año 2016 [16]. A diferencia de la tartrazina, el rojo allura está prohibido en países como EEUU o India, mientras que su uso está permitido en la UE y en países como China, Japón, Australia o Nueva Zelanda [17].

Como se ha podido observar, el control y determinación de los aditivos alimentarios es muy importante, y en concreto el de los colorantes azoicos. Existen por lo tanto una gran cantidad de técnicas que son capaces de analizar los colorantes. Hay que tener en cuenta que los alimentos normalmente son muestras muy complejas, y que en muchos casos, previo al análisis, hay que llevar a cabo algún tipo de extracción del colorante.

Una de las técnicas más empleadas en la determinación de estos colorantes es la de la espectrofotometría, ya que al igual que la mayoría de los colorantes, tanto la tartrazina como el rojo allura absorben luz en el rango del visible. Por medio de la ley de Lambert-Beer, se puede relacionar la radiación absorbida con la concentración de colorante y por ello, se ha utilizado en muchas ocasiones en la determinación de colorantes alimentarios [18][19]. Es una técnica muy útil y sencilla para alimentos que contengan un solo colorante, o en caso de tener más de uno, que estos no absorban en la misma zona del espectro, ya que si hay solapamiento espectral habría que hacer uso de técnicas más complejas como la quimiometría. Si hay además muy poca discriminación espectral (como puede ser por la presencia de dos colorantes rojos), entonces habría que recurrir a otras técnicas de análisis.

Para solucionar este problema, también es muy utilizada la cromatografía líquida de alta resolución (HPLC, *High Performance Liquid Chromatography*), pero para ello se necesitan equipos más sofisticados y el análisis es más costoso. En la mayoría de los casos se acopla a detectores UV-Vis, PDA o MS. Este último detector se emplea en los casos en los que los espectros de los colorantes de la mezcla son muy parecidos y no se puede designar cada pico a cada colorante

[17]. Esta técnica se ha empleado en la determinación de rojo allura en artículos como el de K. Rovina et. al. [20], o de la tartrazina junto con otros colorantes, como en el trabajo de M. Šuleková et. al. [21].

Otras técnicas que también se emplean con el mismo fin son las de la cromatografía en capa fina (TLC) o la electroforesis capilar (CE). El primero de los métodos se utiliza para realizar un análisis cualitativo sencillo y económico, y con el segundo se puede obtener información cuantitativa [17].

2.2. Análisis de imagen digital

El análisis de imagen es una técnica cuya importancia en el campo de la química analítica ha aumentado mucho en los últimos años, gracias en parte al desarrollo de nuevas y más potentes técnicas del análisis del color. Esto se refleja en que entre los años 2005 y 2015 se publicaron el 87,5% de los artículos que había hasta entonces relacionados con este tipo de análisis [22].

La principal característica de los colorantes alimentarios es su capacidad de dar color a los alimentos, lo que se puede apreciar de manera visual, incluso llegar a afirmar que un alimento tiene mucho o poco colorante en función de su tonalidad e intensidad, pero en ningún momento podremos decir la cantidad exacta de colorante que hay. En cierto modo, el color, apreciado visualmente, si ha sido una propiedad que se ha utilizado en el análisis químico, pero sobre todo para en análisis cualitativo o para encontrar el punto final de una reacción [22].

El análisis de imagen es un conjunto de técnicas destinadas a obtener datos relativos a un sistema objeto de estudio a partir de imágenes de dicho sistema [23], y para poder llevarlo a cabo, se necesita un dispositivo de imagen (una cámara o un escáner) conectado a un ordenador, un sistema de iluminación y un software de control y procesamiento. Una vez tomada la imagen, esta se divide en cuadrículas bidimensionales que se conocen como píxeles, que son la unidad mínima de información en una imagen digital [22].

Para poder cuantificar los colores se utiliza la colorimetría. Esta es la ciencia que estudia los procesos relacionados con el color desarrollando métodos para cuantificar la información y así poder procesarla y analizarla, empleando el color como una propiedad más en la detección química y física. Para ello lo que se hace es utilizar los que se conoce como espacio de color, por medio de los cuales se representa un color empleando tres colores primarios, de acuerdo con la afirmación de Helmholtz de que cualquier color puede ser formado combinando tres colores primarios correctamente elegidos. Existen diferentes espacios de color, y la mayoría emplean tres componentes para describir con precisión el sistema de visión humano [22].

Un ejemplo es el espacio de color RGB, que es el más empleado en los dispositivos de imágenes. Funciona de manera similar al sistema visual humano, empleando el rojo (R, red), el verde (G, green) y el azul (B, blue) como los tres componentes para definir un color. Una vez la imagen es tomada en este espacio de color, si se desea, se pueden obtener las coordenadas de otros espacios mediante operaciones matemáticas. Estos espacios pueden ser HSV, CIELAB o CMYK entre otros. En el espacio RGB las coordenadas varían desde la tupla (tipo de dato secuencial que sirve para agrupar varios valores como si fueran un único valor) $[0, 0, 0]$ que define el color negro, hasta el $[2^{n-1}, 2^{n-1}, 2^{n-1}]$, que sería el color blanco y donde n es el número de bits para cada canal. Cada uno de los valores de las coordenadas se refiere a los colores rojo verde y azul, en ese orden ^[22]. La figura 5 muestra el desglose de una imagen tomada en este espacio de color en los diferentes canales que lo forman.

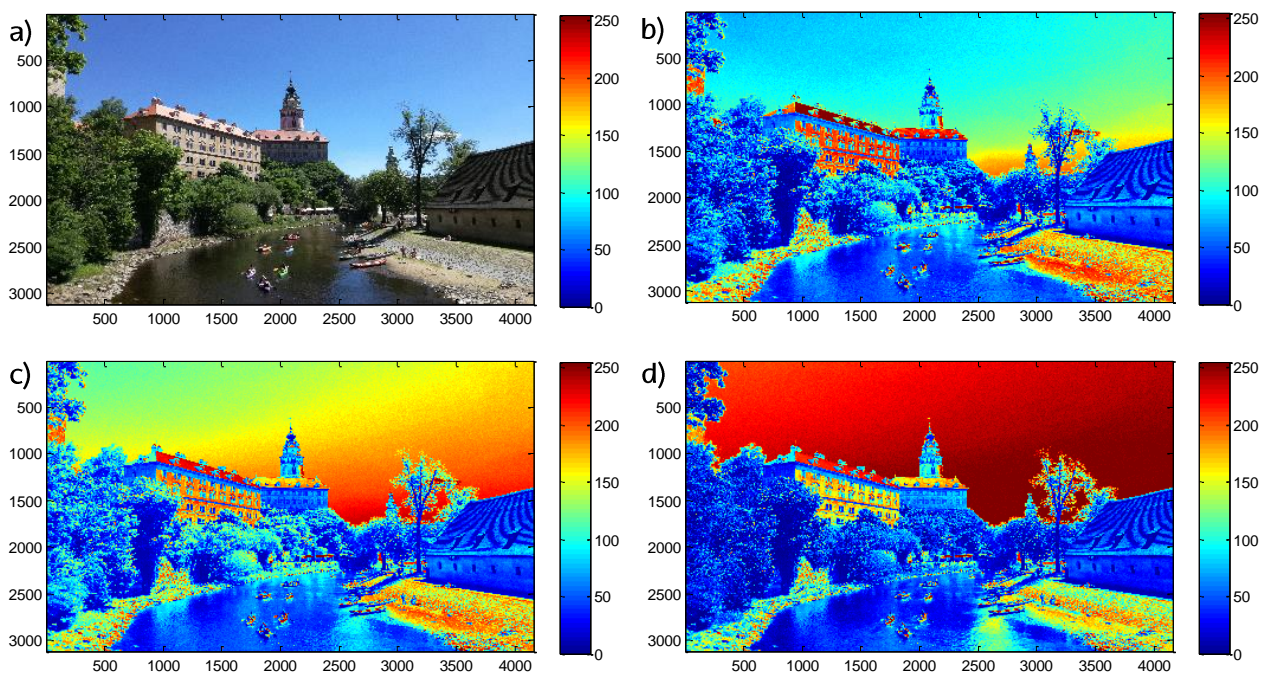


Figura 5. Desglose de una imagen en el espacio de color RGB. a) Imagen original. b) Canal R. c) Canal G. d) Canal B.

La instrumentación que puede emplearse en química analítica para la toma de imágenes digitales, ya sea un escáner o una cámara digital, tiene una fuente de luz, un detector o dispositivo de captura para llevar a cabo el reconocimiento, y la electrónica necesaria para procesar y transmitir la información. El funcionamiento de estos consiste en absorber energía luminosa y convertirla en energía eléctrica, generándose una señal que se suele convertir en un código digital que caracteriza las coordenadas de color correspondientes ^[22].

En este trabajo, el dispositivo empleado para la adquisición de la imagen fue un escáner de oficina. Como en la mayoría de escáneres normales, el elemento a analizar se coloca sobre una placa de vidrio, debajo de la cual hay un elemento móvil compuesto por una fuente de luz y

los sensores. La fuente de luz barre la muestra mientras que los sensores captan la luz reflejada convirtiendo la intensidad recibida en señales eléctricas que se convierten en información digital [22]. En la literatura hay una gran cantidad de trabajos que emplean el escáner como aparato de toma de imagen, e incluso para determinar cuantitativamente colorantes en alimentos, como son los trabajos de F. Soponar et. al. [24] y B.G. Botelho et. al. [25], aunque también se ha realizado un análisis cuantitativo de colorantes empleando teléfonos móviles [26]. El escáner, junto al análisis de imagen, también ha sido empleado por ejemplo para determinar cocaína en muestras tras realizarles el test de Scott [27], en la determinación del contenido de proteínas en arroz [28], en la determinación del brillo en electrodepositos de níquel [29][30] o en la evaluación morfológica de recubrimientos de cromo [31].

En los párrafos anteriores se ha hablado de fuentes de luz, y es que la iluminación es muy importante a la hora de la toma de la imagen. Por eso, a la hora de llevar a cabo un análisis de imagen, este debe hacerse en un ambiente cuyas condiciones de iluminación estén bien controladas, ya que si se emplean diferentes fuentes de luz, los resultados que se obtienen son muy diferentes [22]. Esta es la razón por la cual a la hora de obtener las imágenes con una cámara digital o un teléfono móvil, siempre se crea un sistema más o menos complejo mediante el cual las condiciones de iluminación puedan ser controladas, al igual que los escáneres convencionales tienen una tapa con la que se consigue que toda la iluminación que llegue a la muestra provenga del propio aparato. En este caso, se ha empleado un sistema muy sencillo, que consiste en una tapa que imita a la tapa del escáner, ya que al colocar en el escáner cosas más altas que un papel (como es la placa de microtitulación) la propia tapa del escáner no se puede cerrar y puede haber problemas con la iluminación. A modo de control se ha empleado un patrón de color (Figura 6). Viendo si los colores del patrón cambiaban de una imagen a otra se controlaba la iluminación. Si de una imagen a otra existen cambios habría que corregir las imágenes como viene descrito en los trabajos de M. Vidal et. al. [29][30].



Figura 6. Patrón de color para el control de la iluminación.

Una vez tomada la imagen y almacenada, se pasa al procesado de la imagen digital, lo que implica una serie de operaciones para mejorar la imagen o para sacar información de la misma. Uno de los procesamientos más utilizados es el de la segmentación de la imagen, que tiene como objetivo escoger la parte de la imagen que contiene la información analítica, conocida como ROI (*Region of Interest*) y eso se puede hacer basándose en los valores de píxeles. También hay otras operaciones como pueden ser la del ajuste del brillo, suavizado, etc.

Finalmente, se lleva a cabo el análisis de la imagen digital con la intención de extraer información química (como la concentración) de la región de interés (ROI) de la imagen. Para ello se puede utilizar diferentes parámetros, los cuales se pueden obtener empleando

herramientas y softwares como Photoshop o MatLab. Estos parámetros pueden ser simplemente los valores promedios de los canales R, G y B o de otros espacios de color. En muchas ocasiones no existe una relación sencilla entre la información que se ha extraído y la concentración, y para solucionarlo, una de los métodos que se pueden utilizar, es el de la calibración multivariable ^[22].

Si tuviéramos que dividir el proceso del análisis de imagen este podría dividirse en tres partes: primero estaría la toma de muestra, donde el aparato empleado juega un papel fundamental junto con el control de la iluminación. A continuación, el procesado de la imagen es una etapa importante, ya que preparamos la imagen digital para poder realizar su posterior análisis. Y por último, el análisis de la imagen, que consiste en obtener los parámetros de la imagen que intentaremos relacionarlos con el color.

La gran ventaja de esta técnica es que es una técnica barata y rápida, a pesar de que en algunos casos la obtención de los parámetros de la imagen pueda ser algo complicada, y además se trata de un análisis no destructivo.

2.3. Quimiometría

Como se ha explicado en el apartado anterior, en muchos de los casos en los que se lleva a cabo un estudio por medio del análisis de imagen, se obtienen una gran cantidad de datos, que necesitan un tratamiento complejo con el objetivo de obtener la información de interés.

En general, con el desarrollo de nuevos métodos, más modernos, se pueden obtener una gran cantidad de datos, para los cuales hay que emplear herramientas especiales como es la quimiometría, empleada también en este trabajo. La quimiometría es una disciplina que, aplicando conocimientos matemáticos, generalmente estadísticos, a procesos químicos, consigue extraer de los datos experimentales la mayor cantidad posible de información y extender el conocimiento del sistema químico. Partiendo de la base de que todos los datos no son información, por medio de la quimiometría se consigue extraer toda o la mayoría esta información disminuyendo los datos iniciales al discriminar entre la información relevante y otra de menor interés ^[32].

Esta es una herramienta compleja y que requiere de equipos y softwares especializados, permitiendo trabajar con más de una variable (datos multivariantes) a diferencia del análisis tradicional donde solamente hay una única variable (datos univariantes). Por ejemplo, el problema de conocer la concentración de un colorante en una muestra mediante espectrofotometría puede resolverse mediante un análisis univariable, escogiendo la absorbancia a la longitud de onda en la que el coeficiente de absorción del analito es máximo.

Pero cuando en la muestra hay una mezcla de colorantes cuyos espectros se solapan y las absorciones en los máximos se ven influenciadas por la presencia del otro compuesto, esto se puede resolver escogiendo una parte más amplia del espectro (más de una longitud de onda) y después empleando herramientas quimiométricas. Para ello lo que se suele hacer es crear una matriz de datos, en las que cada columna corresponde a una variable (en este caso las diferentes longitudes de onda) y cada fila a una muestra para crear un modelo robusto de calibración [32].

La quimiometría tiene diferentes campos de aplicación. Esta se puede aplicar para el diseño de experimentos, ya que con ella se pueden ver que variables no son relevantes y de esa manera evitar experimentos que no son necesarios, y también estudia dónde, cómo y cuándo deben realizarse las experiencias para que contengan la información necesaria [33].

Otro campo de aplicación es la búsqueda de tendencias en los datos con el objetivo de reducir el espacio en el que observar nuestras muestras, reduciendo el espacio original de múltiples dimensiones a uno de dimensiones reducidas. Para ello se deben combinar adecuadamente las variables medidas formando nuevas variables que se conocen como variables latentes o componentes principales que contengan la información inicial, y la técnica análisis de componentes principales (PCA, *Principal Component Analysis*) es la más empleada para ello. De esta manera será más fácil comprobar si el conjunto de muestras es homogéneo, encontrar muestras discrepantes o que se diferencien del resto (*outlier*) y ver si todas las variables son independientes o si tienen relación entre ellas y por lo tanto hay algunas redundantes [33].

Una tercera aplicación es la de relacionar conjuntos de datos. Si por ejemplo se quiere relacionar los parámetros que han sido medidos con las propiedades que miden la calidad de un producto se emplea la calibración multivariable. Al no haber modelos teóricos y formarse estos ajustándose a los datos que se introducen (modelos "*ad hoc*") es muy importante llevar a cabo pruebas de validación [33].

Una última aplicación que cabe destacar es la posibilidad de analizar sistemas en evolución. Si por ejemplo queremos conocer la concentración de un sustrato en un sistema en evolución se deben coger muestras y analizarlas continuamente, algo que resulta laborioso. En cambio, existen técnicas más sencillas de determinación de la concentración. Un ejemplo sería el empleo de medidas en el infrarrojo cercano (NIR, *Near InfraRed*). Con una sonda se pueden tomar continuamente espectros del sistema, espectros que varían con la concentración de sustrato. Estos espectros se podrían resolver por medio de la quimiometría y conocer así la concentración de sustrato en cada instante [33].

En este trabajo, se ha aplicado la quimiometría, y más concretamente la calibración multivariable, por un lado disminuyendo el número de variables mediante un PCA y por otro relacionando las variables reducidas con la concentración de los colorantes en las muestras

utilizando el algoritmo de la regresión parcial por mínimos cuadrados (PLS, *Partial Least Squares*).

2.3.1. Análisis de componentes principales

Como se ha dicho anteriormente, el objetivo del PCA es la reducción de datos. Para ello, a partir de un conjunto inicial de variables (observables) se obtiene otro conjunto de variables (no observables) por combinaciones lineales de las originales y que son capaces de describir la muestra. Se obtiene la misma cantidad de variables que al principio, y estas se conocen como variables latentes o componentes principales (PC, *Principal Components*), pero en este caso se elige como primer componente principal (PC1) a aquel que recoge la mayor variabilidad de los datos, y el siguiente (PC2) recogerá la siguiente mayor parte de esta y así sucesivamente y finalmente se cogerán los PCs necesarios siempre que la varianza que expliquen sea significativamente mayor al nivel de ruido de los datos^[34].

La mejor forma de ilustrarlo es por medio de un ejemplo en la que solo tenemos 3 variables: partiendo de una matriz de datos que contiene 20 muestras y 3 variables que corresponden a tres longitudes de onda, se hace una representación de los datos en un espacio tridimensional, ya que tenemos 3 variables (Figura 7a). A continuación se buscan los componentes principales, como ya se ha explicado antes. El PC1 tendrá la dirección de máxima variabilidad de los datos, y el PC2 explicará la segunda máxima variabilidad pero siempre siendo ortogonal a PC1 (Figura 7b). El PC3, también existirá, pero si la variabilidad que logra explicar en los datos es mínima, se puede prescindir de él y así el sistema experimentará una reducción de variables. Finalmente se cambia el eje de coordenadas para una visualización más sencilla de los datos (Figura 7c)^[33].

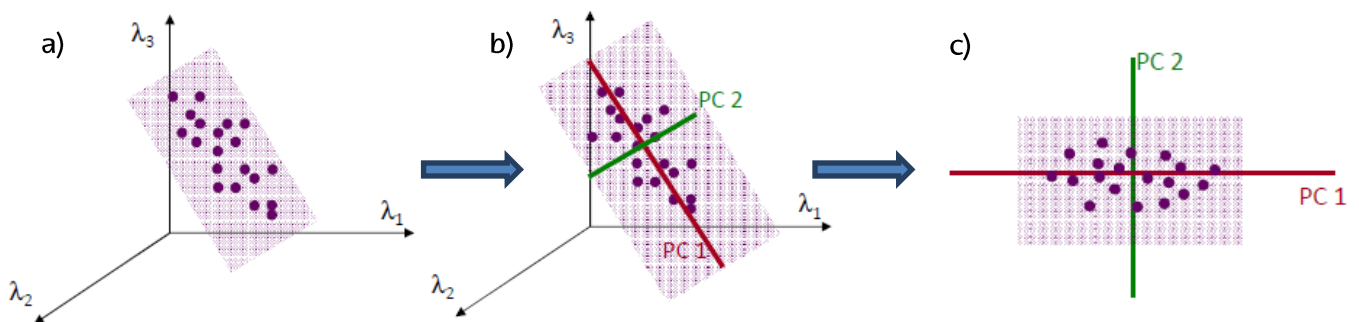


Figura 7. Reorganización de la matriz de datos por medio del algoritmo del PCA.

Las muestras se relacionan con los nuevos ejes, es decir con las nuevas variables o PCs por medio de los *scores*. Los *scores* son la distancia de las muestras a estos nuevos ejes, por lo que cada muestra tendrá un *score* diferente dependiendo de a que PC se refiera. Lo mismo ocurre con las variables iniciales, que por medio de los *loadings* se relacionan los ejes antiguos

con los nuevos ejes. De esta manera se puede ver de manera muy sencilla qué variables están relacionadas y dan la misma información.

2.3.2. Calibración multivariable

La calibración es una operación con la que se busca relacionar un valor indicado por un instrumento de medida (X) con valores conocidos, que generalmente serán concentraciones (y). Si X fuera una sola variable, la operación que se realiza se conoce como calibración univariable, pero al ser una matriz, para cada valor de concentración (y) hay un vector de información (más de una variable), y es entonces cuando se realiza la calibración multivariable.

Para llevar a cabo esta última, volvemos a valernos de la quimiometría, siendo la regresión de componentes principales (PCR, *Principal Component Regression*) y la ya mencionada regresión parcial por mínimos cuadrados (PLS, *Partial Least Squares Regression*), algoritmos muy utilizados. Ambos se basan en una reducción de variables por medio de un PCA, con la diferencia de que en el PCR se establece una relación entre los scores, obtenidos en un PCA previo de los valores de la matriz X , con los valores de y , mientras que con el PLS lo que se realiza es un PCA en el que se tiene en cuenta tanto los valores de la matriz X como los de y buscando que el modelo explique la máxima variabilidad de ambas matrices.

Antes de realizar cualquiera de estas operaciones, ya sea un PCA, un PCR o un PLS, se suele realizar un pre-procesado de los datos, ya que si una de las variables tiene otras unidades o una magnitud diferente, el modelo puede resultar distorsionado. Esto ocurre porque si una de las variables tiene un rango de variabilidad mayor (una varianza mayor) al de las demás puede tener mayor influencia en el modelo. El pre-procesado consiste en realizar operaciones matemáticas. Uno de estos pretratamiento es el centrado, en la que se le resta a cada dato la media de los valores de la variable a la que corresponde, de forma que el centro de todas las variables se sitúa en cero. Junto al centrado, el otro pre-procesado muy empleado es el autoescalado, donde una vez los datos han sido centrados, estos se dividen por la desviación estándar de los valores de la variable. La operación de centrado se suele emplear casi siempre, y la de autoescalado cuando las variables tienen diferentes varianzas^[35].

2.4. Espectrofotometría ultravioleta-visible

En este trabajo, se ha empleado la espectrofotometría UV-Vis como referencia, ya que los colorantes alimentarios son capaces de absorber luz de la región del visible y por ello esta técnica es una de la más utilizada en el análisis cuantitativo de estos compuestos^{[17][36]}.

La espectrofotometría es el conjunto de procedimientos que utilizan la luz para medir concentraciones químicas, en este caso lo que ocurre es que al radiar una molécula a diferentes longitudes de onda, cuando la energía de los fotones coincide con la energía de algún salto de energía de la propia molécula, esta absorbe el fotón pasando a un estado excitado. En la espectrofotometría de absorción UV-Vis lo que se mide es la transmitancia (T), que indica la fracción de luz incidente que pasa a través de la muestra, por lo que se calcula dividiendo la radiación que no ha absorbido la muestra (P) entre la radiación inicial (P₀):

$$T = \frac{P}{P_0} \quad (1)$$

Para relacionar la transmitancia con la concentración primero se debe definir la absorbancia (A), ya que es esta la que es directamente proporcional a la concentración (c).

$$A = \log_{10} \left(\frac{P_0}{P} \right) = -\log_{10} T \quad (2)$$

La relación entre absorbancia y concentración viene dada por la Ley de Beer:

$$A = \epsilon b c \quad (3)$$

En la ecuación 3, ϵ es la absorptividad molar, que se define como la cantidad de luz que absorbe una especie determinada a una longitud de onda concreta por unidad de concentración en molaridad. b es el paso óptico, y la c es la concentración de la especie absorbente.

Gracias a los monocromadores de los que disponen los espectrofotómetros, se obtiene la información a cada longitud de onda, haciendo un barrido que se resuelve en un espectro^[37]. Cuando en la muestra únicamente hay una única sustancia que absorbe se puede obtener la concentración de este a partir de un análisis univariable, escogiendo la absorbancia a la longitud de onda que más absorba el analito. Pero cuando se tratan de mezclas, como es el caso de este trabajo, en el que hay presencia de dos colorantes, los espectros se pueden solapar y para resolverlos se debe utilizar el análisis multivariable^[38].

3. Objetivo

El objetivo de este trabajo es desarrollar un método sencillo, barato y rápido basado en el análisis de imagen digital que permita determinar con una exactitud aceptable la concentración del colorante alimenticio E 129 en mezclas, concretamente en presencia de otro colorante como es el E 102, y utilizando la espectrofotometría UV-Vis como técnica de referencia. Mediante la espectrofotometría UV-Vis se determinarán cuantitativamente ambos colorantes.

Para ello se utilizará la calibración multivariable, a través de la cual se obtendrán modelos para ambos métodos, modelos que se validarán. Se calcularán parámetros estadísticos como son la exactitud y la precisión, se harán estudios de recuperación y se aplicarán los modelos a muestras reales que contengan estos colorantes. Finalmente, se compararán los resultados obtenidos por análisis de imagen con los obtenidos por espectrofotometría UV-Vis.

4. Experimental

4.1. Preparación de muestras y optimización del pH

Primero se pesaron los gramos de colorante alimentario necesarios para preparar las disoluciones stock tanto de tartrazina (E 102) como de rojo allura (E 129) de 100 mg/L en agua bidestilada. A partir de ellas se prepararon todas las muestras necesarias para la calibración y la validación de los dos métodos, además de emplearlas en el cálculo de las recuperaciones. Ambos colorantes fueron suministrados por la empresa Roha Epsa S.L. (Valencia, España), y la balanza empleada para pesar todos los reactivos y muestras reales empleados durante el trabajo fue una METTLER TOLEDO (Greifensee, Suiza).

Todas las disoluciones utilizadas en la obtención del calibrado se prepararon a partir de las stocks haciendo diluciones de éstas con agua bidestilada para obtener las mezclas de concentraciones deseadas. A continuación se prepararon las muestras reales, las cuales tuvieron una preparación algo más larga. Las muestras reales seleccionadas para analizar fueron un colorante alimentario comercial, muy utilizado sobre todo como colorante para la paella, y un helado alargado refrescante de naranja (flash).

Antes de comenzar a preparar cualquier muestra, se llevó a cabo un estudio de la influencia del pH en la absorbancia de ambos colorantes por separado. Para ello, de las disoluciones stock, se cogieron las cantidades necesarias para preparar nuevas disoluciones de 30 mg/L a y se les varió el pH. Primero se midió el pH de la disolución original de cada colorante para después añadir ácido acético glacial o diluido, si lo que se deseaba era obtener pHs ácidos, o una disolución de amoníaco al 10% v/v. Añadiendo volúmenes diferentes de ambos, se obtuvieron los pH que se muestran en la tabla 1, de las cuales se obtuvieron los espectros correspondientes.

Tabla 1. Valores de pHs obtenidos para las diferentes disoluciones de ambos colorantes. La muestra 1 en ambos casos se refiere a la disolución de colorante en agua

| E 102 | | E 129 | |
|---------|-------|---------|-------|
| Muestra | pH | Muestra | pH |
| 1 | 8,33 | 1 | 7,75 |
| 2 | 4,92 | 2 | 3,51 |
| 3 | 10,67 | 3 | 2,89 |
| 4 | 3,59 | 4 | 10,51 |
| 5 | 2,84 | | |

Observando como variaban los espectros con el pH se estableció el rango de pH óptimo, considerando este rango como el rango de pH donde los valores de absorbancia a la longitud de onda máxima eran más altos. Para ajustar posteriormente el pH de las muestras reales en el

rango de pH óptimo, se emplearon las mismas disoluciones de ácido acético y NH_3 10% v/v. Este procedimiento de optimización se llevó a cabo con un pH-metro de sobremesa Crison Basic 20 (Barcelona, España).

Antes de medir los espectros y de obtener las imágenes, las muestras reales se filtraron con filtros de acetato de celulosa de $0,45 \mu\text{m}$ de poro distribuidos por FILTER-LAB® (Barcelona, España) con el objetivo de eliminar todas las impurezas que pudiesen quedar en las disoluciones.

Por un lado, el colorante alimentario, que contiene tartrazina, es un sólido, del cual $0,0480 \text{ g}$ se disolvieron en $100,0 \text{ mL}$ de agua destilada (D1). Posteriormente, se hizo una dilución 1/10. Para obtener los porcentajes de recuperación, tanto para rojo allura (por análisis de imagen y espectrofotometría) como para tartrazina (por espectrofotometría), lo que se hizo fue coger $1,00 \text{ mL}$ de D1 y se añadieron diferentes cantidades de los colorantes a partir de las stock, para finalmente añadir agua bidestilada hasta alcanzar los $10,00 \text{ mL}$ de disolución. Este procedimiento se repitió completamente volviendo a preparar una nueva disolución (D2), en la que se pesaron $0,0576 \text{ g}$ de colorante alimentario que se disolvieron en $100,0 \text{ mL}$ de agua bidestilada.

Por otro lado, el flash de naranja (que contiene ambos colorantes) se dejó descongelar, y una vez líquido se cogieron $2,00 \text{ mL}$ y se añadieron $50,0 \mu\text{L}$ de NH_3 10% v/v, ya que a diferencia del colorante alimentario tiene un pH muy ácido y se debe ajustar al rango de pH óptimo que se especificará después. Finalmente se añadieron $8,00 \text{ mL}$ de agua bidestilada dando lugar a la disolución F1. Siguiendo el mismo procedimiento, pero añadiendo a la muestra de flash inicial diferentes concentraciones de tartrazina y rojo allura, se calcularon los porcentajes de recuperación. Con un segundo lote se siguió el mismo procedimiento para preparar F2, pero en este caso, a la hora de preparar las muestras para obtener las recuperaciones solamente se cogió $1,00 \text{ mL}$ de flash, al cual se le añadieron $25,0 \mu\text{L}$ de amoníaco y luego distintas cantidades de stock y nuevamente agua bidestilada hasta alcanzar, en este caso, los $10,025 \text{ mL}$.

4.2. Análisis de imagen

4.2.1. Aparatos y softwares

Todas las imágenes fueron tomadas por medio de un escáner de oficina como es el EPSON STYLUS DX7400 (Japón), gracias al software EPSON Scan que se encontraba instalado en el ordenador con el que se trabajó. Este ordenador era un DELL Optiplex 7010 Intel® CORE™ i7 de 465 GB , que también disponía de la herramienta de *MATLAB R2013b* (The Mathworks, Inc.

Natick, MA) con la *PLS_toolbox* (Eigenvector Research, Inc, Wenatchee, WA) que permitió el tratamiento de los datos y la posterior obtención de modelos. Para el cambio de las imágenes entre los diferentes espacios de color se usó la *Image Processing Toolbox™* (The Mathworks, Inc. Natick, MA).

4.2.2. Muestras empleadas

Para la calibración y la validación se emplearon un total de 73 muestras con mezclas de los dos colorantes preparadas a partir de los stocks de 100 mg/L, 73 puntos que formaban una matriz de 2D, en la que en el eje x aparecen las concentraciones de E 102, mientras que en el y las concentraciones de E 129. Ambos colorantes se encontraban en concentraciones entre los 0,20 mg/L y los 50,0 mg/L.

Para calcular la reproducibilidad de la técnica se prepararon muestras a dos niveles de concentración, uno bajo, de 10,0 mg/L y otro alto, de 26,0 mg/L para los dos colorantes. Se midieron 20 veces durante 5 días, y variando la posición de medida de las disoluciones en el escáner, entre una esquina del escáner y el centro. Las medidas de repetibilidad se llevaron a cabo con los mismos niveles de concentración que en el caso de la reproducibilidad, y para ello se tomaron 10 medidas, todas ellas durante un mismo día y en la misma posición.

En cuanto a las muestras reales, se tomó una muestra de flash de dos lotes diferentes, al igual que del colorante alimentario. Para el cálculo de la recuperación, en el lote 1 del flash se emplearon dos niveles de concentración, uno alto y otro bajo, con réplicas en ambos casos. En el lote 2 de flash las recuperaciones se dividieron en cuatro niveles, todos medidos en dos ocasiones, donde en dos de los niveles únicamente se añadía uno de los dos colorantes mientras que en los otros dos se añadían ambos. El procedimiento seguido para el colorante alimenticio fue muy parecido. En este caso, partiendo de ambos lotes se prepararon seis disoluciones con cantidades diferentes de las disoluciones stock añadidas para estudiar las recuperaciones. En dos de ellas se añadió E 102 a dos niveles de concentración; en otras dos E 129 a dos niveles de concentración y en las últimas dos se añadió tanto E 102 como E 129 en distintos niveles.

4.2.3. Matriz de calibración y de validación

Al azar se separaron las 73 muestras en dos grupos, donde 47 fueron al grupo de calibración y las 24 restantes al grupo de validación, manteniendo siempre las réplicas en un mismo grupo, lo que se refleja en la figura 8.

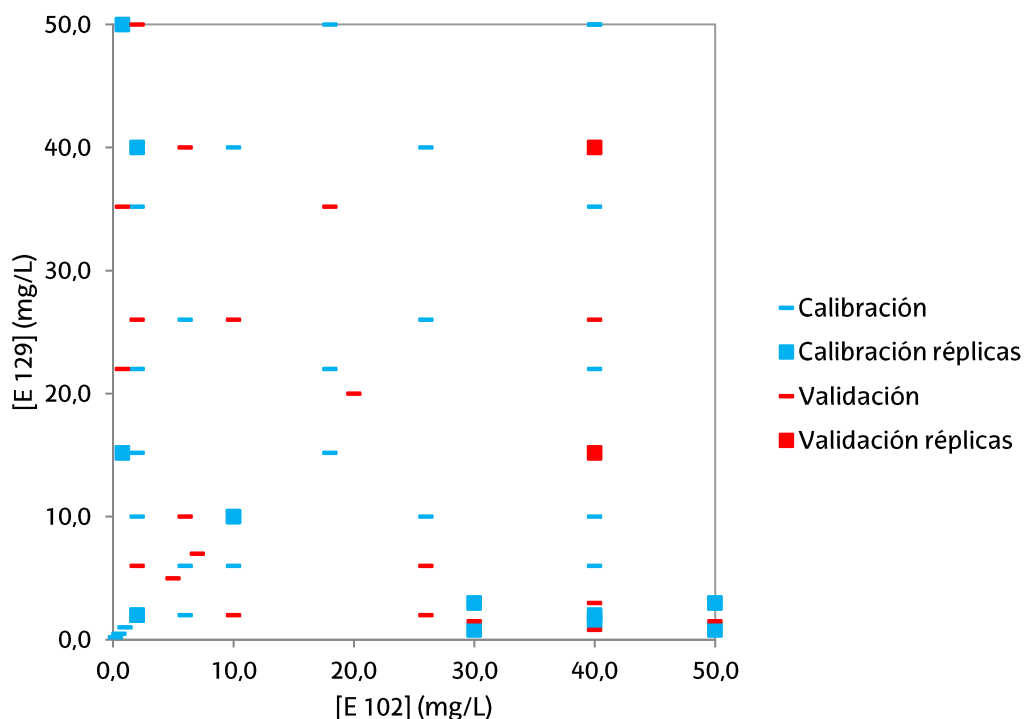


Figura 8. Matriz de concentraciones para el análisis de imagen diferenciando muestras de calibración de validación. En azul las muestras de calibración y en rojo las de validación. Las líneas corresponden a muestras únicas mientras que los cuadrados representan réplicas.

En la figura 8 se pueden diferenciar las muestras que fueron seleccionadas para formar parte de la calibración (en azul) de las de validación (en rojo) para la técnica del análisis de imagen. También se pueden ver en ambos grupos, indicados con cuadrados de su color correspondiente, las concentraciones en las que se han medido más de una muestra, es decir, en las que hay réplicas.

4.2.4. Toma de las imágenes

Para la obtención de la imagen, primero se emplearon unas placas Petri, pero finalmente todo el estudio se llevó a cabo utilizando unas placas de microtitulación de Bioster (República Checa) (empleadas también en trabajos como el de R. Ahirwar et. al.^[39] o el de D. Damasceno et. al.^[40]), de esta manera se ahorra mucha cantidad de muestra y además en una misma imagen se podían escanear varias muestras.

Para la toma de imágenes se empleó una tapa blanca, hecha a medida, para simular las condiciones normales de un escáner con la tapa bajada, ya que al introducir las placas la tapa no podía bajarse y por lo tanto no se conseguía aislar las muestras de la iluminación exterior. Además de esto, en toda imagen que se tomaba se colocaba un patrón de color, (Figura 9a), para poder controlar la estabilidad de las medidas. El patrón de color se fabricó a partir de piezas de LEGO® para emplearlo en trabajos anteriores en el Grupo de Quimiometría Aplicada de la UPV/EHU^[41].

En los pocillos de las placas de microtitulación únicamente había que emplear 400 μ L de muestra. En cada placa hay un total de 94 pocillos que se podían rellenar, pero solamente unos pocos se utilizaban en cada medida para que no hubiera problemas por reflexión de la luz de pocillos llenos con muestra, sobre pocillos contiguos (Figura 9b).

Una vez preparada la placa con las muestras en los pocillos, ésta se llevaba al escáner y se intentaba colocar siempre en la misma posición (como muestra la figura 9c), y junto con ella, siempre el patrón de color. Una vez colocados, se tapaban con la tapa blanca para aislar las muestras de la luz del exterior y se procedía al escaneado. El escaneado se realizó siempre bajo las mismas condiciones (Figura 9d). De esta manera se obtenían imágenes de 5100 x 7019 x 3 píxeles, de los cuales sólo interesaban aquellas regiones en las que se encontraban las muestras (ROI). Por ello, tras tomar la imagen, éstas eran recortadas reduciéndose cada muestra a una imagen de dimensiones 21 x 21 x 3, y se calculaban los parámetros de las imágenes que se deseaban.

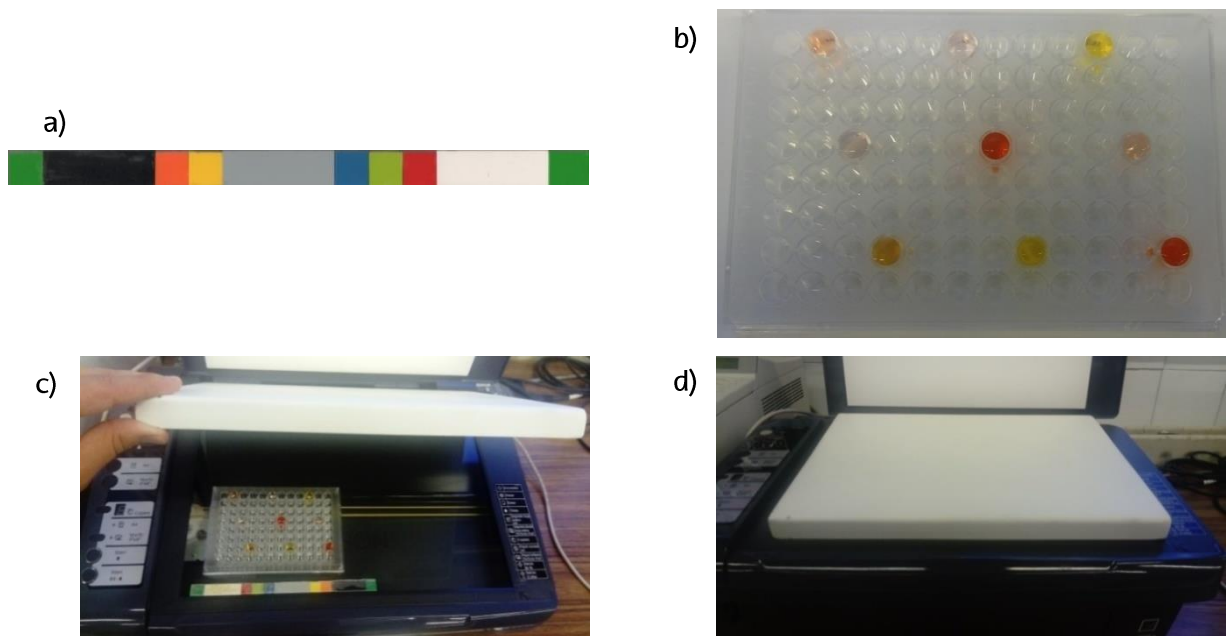


Figura 9. Toma de imagen. a) Patrón de color para controlar la estabilidad del escáner. b) Placa de microtitulación con 9 pocillos llenos con mezclas de diferentes concentraciones de los colorantes E 102 y E 129. c) Posición de la placa y del patrón de color a la hora de la toma de imagen. d) Situación al colocar la tapa para evitar problemas con la iluminación.

Además de las muestras, también se recortaba la parte de la imagen específica en cada pieza coloreada del patrón de color para el posterior estudio de estabilidad de la iluminación.

4.2.5. Procesamiento de la imagen

Una vez realizados los recortes, se intentó encontrar una relación lineal entre la concentración de cada colorante y valores extraídos de las imágenes.

El problema se abordó de tres formas. En la primera se intentó aplicar la calibración univariable buscando una relación directa entre cada canal R, G y B con la concentración. En la segunda, se aplicó un método obtenido de la bibliografía (*método bibliográfico*) con el que anteriormente se ha relacionado el color extraído de diferentes imágenes con fines cuantitativos ^{[42][43]}, y en la tercera se propuso un procedimiento propio (*método propio*), y sencillo, de extracción de la información de color que puede ser relevante para relacionarla con la concentración. Los últimos dos métodos se van a explicar con detalle a continuación.

4.2.5.1. Método bibliográfico

Se buscó obtener una huella dactilar (*colorgrama*) parecida a la que se emplea en dos artículos de la bibliografía ^{[42][43]}. En el artículo de A. Antonelli et. al. ^[42], del año 2004, con esta huella dactilar se clasificaron distintas salsas de pesto de cinco productores diferentes, y se obtuvieron resultados muy alentadores. En cambio, en el segundo de los artículos, de A. Ulrici ^[43], este más reciente, del 2012, se buscó identificar muestras de jamón que estuvieran dañadas por falta de color rojo en su piel. Para poder construir esta huella dactilar, primero se deben obtener los valores de R, G y B para cada pixel de cada una de las imágenes, con lo que se construye una matriz tridimensional de datos para cada muestra de $X \times Y \times 3$, donde X e Y son las dimensiones de imagen, y la tercera dimensión hace referencia a los tres canales R G y B. Una vez obtenida esta matriz, se transforma a una matriz de dos dimensiones de forma que: $X \times Y \times 3 = XY \times 3$, donde las dimensiones XY son el producto de X e Y . Al cambio de las dimensiones de una matriz se llama *reshape*. Después se realizan tres PCAs, donde las variables son los tres canales R, G y B: Un PCA sin ningún pretratamiento (PCA_RAW), un PCA con un centrado de los datos (PCA_MNCN) y finalmente un PCA con autoestacalado (PCA_AUTO). De cada uno de estos tres PCAs se sacan los valores del primer score.

A continuación se genera una nueva matriz para cada muestra, donde cada pixel estará definido por 13 variables. Las tres primeras serán los valores de R, G y B, por ese orden, y a continuación el valor de la suma de los tres canales (L). Las variables 5, 6 y 7 son los valores de los colores relativos (rR, rG y rB), que para obtenerlos hay que dividir cada canal por la suma de los tres. Las siguientes tres variables (8 a 10) son las que pertenecen a los valores de los tres canales de un nuevo espacio de color como es el HSV, que se divide en tres nuevos canales como son el H, el S y el V. Las últimas variables de esta matriz corresponden a los valores de los scores del primer PC obtenido en cada uno de los PCA anteriores, siendo la variable 11 el score del PC1_RAW, el 12 el de PC1_MNCN y el 13 el de PC1_AUTO.

Una vez obtenida la matriz se realiza un histograma para cada columna, obteniéndose así una nueva matriz de dimensiones 256 x 13 para cada muestra, que se transforman a su vez en un vector de 1 x 3328 (256 x 13) (*reshape*). A este nuevo vector se le añaden los valores de los tres *loading* y el primer *eigenvalue* de cada PCA anterior. De manera que por cada PCA se añaden 3 *loadings* y un *eigenvalues*, dando un total de 12 nuevas variables y formándose así el vector definitivo de 3340 variables. Al juntar todos los vectores de cada una de las 73 muestras, se obtiene una matriz de 73 x 3340. Este procedimiento, se explica de manera esquemática en la figura 10.

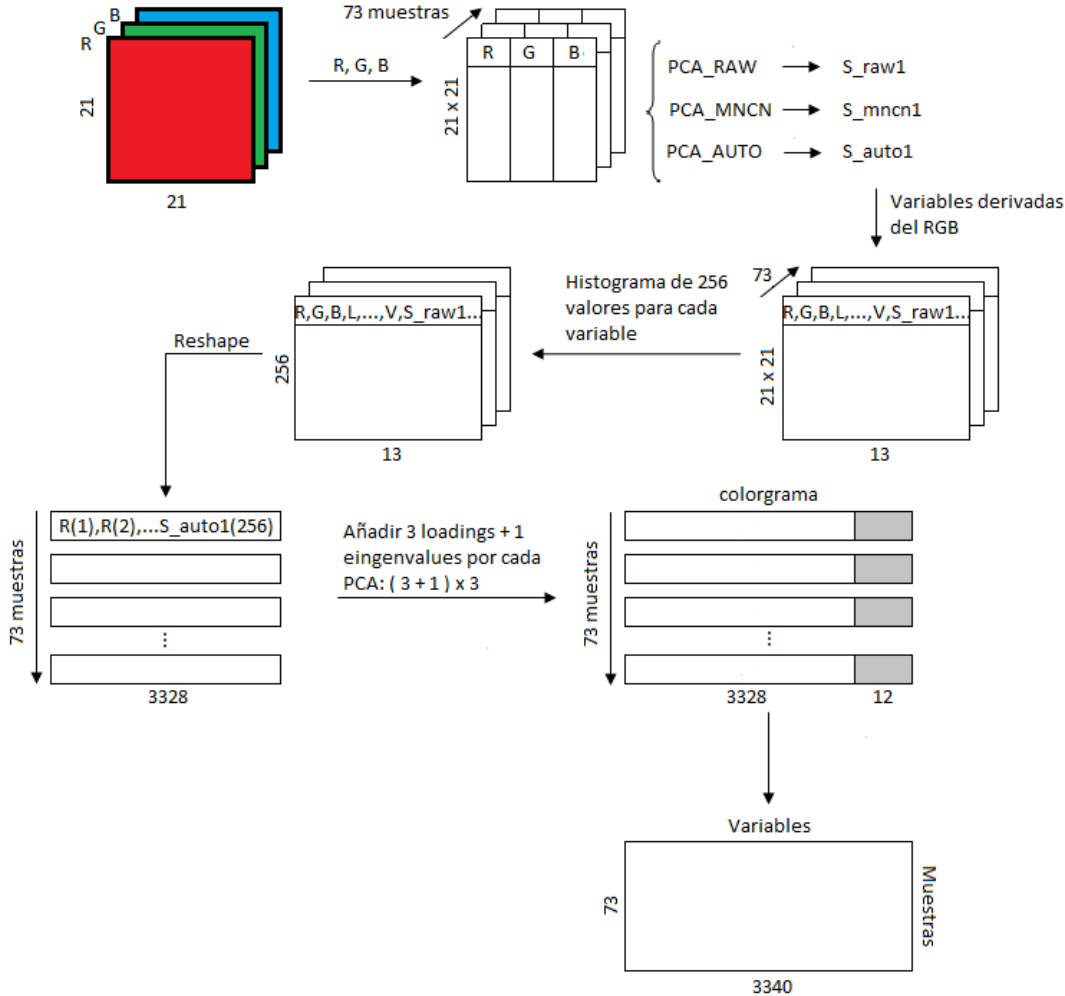


Figura 10. Esquema del procedimiento para la obtención de los colorgramas para las 73 muestras del calibrado tras hacer el recorte de 21 x 21 x 3.

4.2.5.2. Método propio

Aquí se buscó una nueva huella dactilar mucho más simple que la anterior compuesta únicamente por las medias de canales de diferentes espacios de colores. Los espacios de colores seleccionados fueron:

- RGB
- HSV
- CIELAB, donde el espacio de color se divide en los canales L, A y b.
- CMYK, donde el espacio de color se divide en los canales C, M, Y y K.
- YCBCR, donde el espacio de color se divide en los canales y, CB y CR.

De esta manera, se obtuvo un nuevo colorgrama donde cada muestra está definida por un total de 16 variables muy fáciles de obtener (Figura 11).

| Espacio | RGB | | | HSV | | | CIELAB | | | CMYK | | | | YCBCR | | |
|---------|-----|---|---|-----|---|---|--------|---|---|------|---|---|---|-------|----|----|
| Canal | R | G | B | H | S | V | L | A | b | C | M | Y | K | y | CB | CR |
| M 1 | | | | | | | | | | | | | | | | |
| M 2 | | | | | | | | | | | | | | | | |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| M 73 | | | | | | | | | | | | | | | | |

Figura 11. Esquema del colorgrama obtenido con la huella dactilar del método propio.

4.3. Espectrofotometría UV-Vis

4.3.1. Aparatos y softwares

En la obtención de los espectros por medio de la espectrofotometría UV-Vis se emplearon dos espectrofotómetros, por un lado T92+UV spectrophotometer PG y por otro un Agilent 8453 UV-Visible spectrophotometer. Para su posterior análisis se utilizó el mismo ordenador que en el de análisis de imagen, pero en este caso se usó un software diferente, *The Unscrambler X 10.3 Client* (Como, Trondheim, Noruega) que permitió obtener los calibrados por medio de un análisis multivariable.

4.3.2. Muestras empleadas

Para esta técnica, se añadieron 21 estándares más a los 73 analizados en el análisis de imagen, dando un total de 94, mientras que las muestras reales analizadas fueron exactamente las mismas.

En este caso, el estudio de la precisión se realizó de la siguiente manera: para conocer la reproducibilidad, lo que se hizo fue medir un único nivel, donde ambos colorantes tenían una concentración cercana a los 8 mg/L, tres veces al día durante tres días. La repetibilidad en cambio se obtuvo de medir este mismo nivel de concentración en seis ocasiones durante un mismo día.

4.3.3. Matriz de calibración y de validación

La división de las muestras para la calibración y validación de esta segunda técnica se llevó a cabo también al azar, dando lugar a dos grupos de 47 muestras cada uno como muestra la figura 12, donde se puede observar que las muestras han sido representadas de la misma forma que se ha hecho en la figura 7.

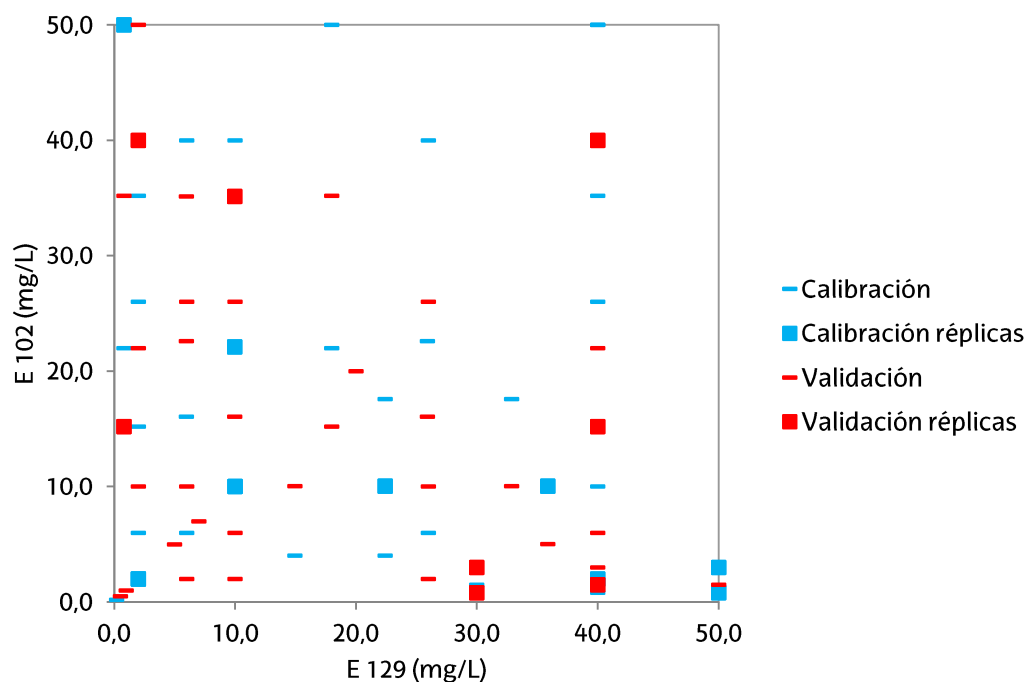


Figura 12. Matriz de concentraciones para el análisis de imagen diferenciando muestras de calibración de validación. En azul las muestras de calibración y en rojo las de validación. Las líneas corresponden a muestras únicas mientras que los cuadrados representan réplicas.

Una vez realizados los espectros, para la medida de los cuales se usó siempre la misma cubeta, éstos se recortaron para obtener el intervalo de longitudes de onda de interés. En este caso ese intervalo iba desde los 350 hasta los 610 nm, intervalo en el cual se encuentran los máximos de absorción de los dos colorantes, y por lo tanto el que más información daba sobre ambos.

4.4. Cálculo de los errores de los modelos de calibración multivariable

Para valorar los modelos obtenidos se calcularon los errores relativos (E.R., %) siguiendo la ecuación 4, donde n es el número de muestras, $y_{hallado}$ es la concentración predicha por el modelo y donde y_{puesto} es el valor de concentración de la disolución preparada. Por medio de esta ecuación se calculan los errores relativos tanto de la calibración (E.R._{cal}), como de la *cross-validation* (E.R._{cv}) y la de la validación externa (E.R._{val}).

$$E. R. (\%) = 100 \sqrt{\frac{\sum_{i=1}^n (y_{hallado} - y_{puesto})^2}{\sum_{i=1}^n (y_{puesto})^2}} \quad (4)$$

4.5. Cálculo de los límites de detección

Finalmente cabe destacar la forma en la que se calculó el límite de detección (LOD, *Limit Of Detection*) para los dos métodos empleados. El LOD es uno de los parámetros más importantes en los métodos analíticos, y representa la mínima cantidad de especie a identificar que genera una señal analítica significativamente diferente de la señal del blanco o ruido de fondo^[32]. A pesar de que en la calibración univariable el cálculo del LOD está bien definido, en la calibración multivariable no existe un método definitivo. En este caso se calculó siguiendo el método propuesto por M. Ostra et. al^[44], y que fue empleado en la Tesis de la Dra. M Vidal^[41] siguiendo la ecuación 5. En esta ecuación se toma como desviación estándar del blanco (S_{blanco}) la desviación estándar calculada a partir de los residuales obtenidos del modelo de calibración de PLS y b es el vector de regresión, donde se indica cuanto contribuye cada variable en el modelo final. El límite de cuantificación (LOQ, *Limit Of Quantification*) se calculó multiplicando el LOD por 3,3.

$$LOD = 3 \cdot S_{blanco} \cdot b \quad (5)$$

5. Resultados y discusión

5.1. Optimización del pH

Como se ha explicado en el apartado 4.1., lo primero que se llevó a cabo fue la optimización del pH midiendo los espectros de las muestras a los pHs indicados en la tabla 1. De esta manera, se obtuvieron los espectros de la figura 13:

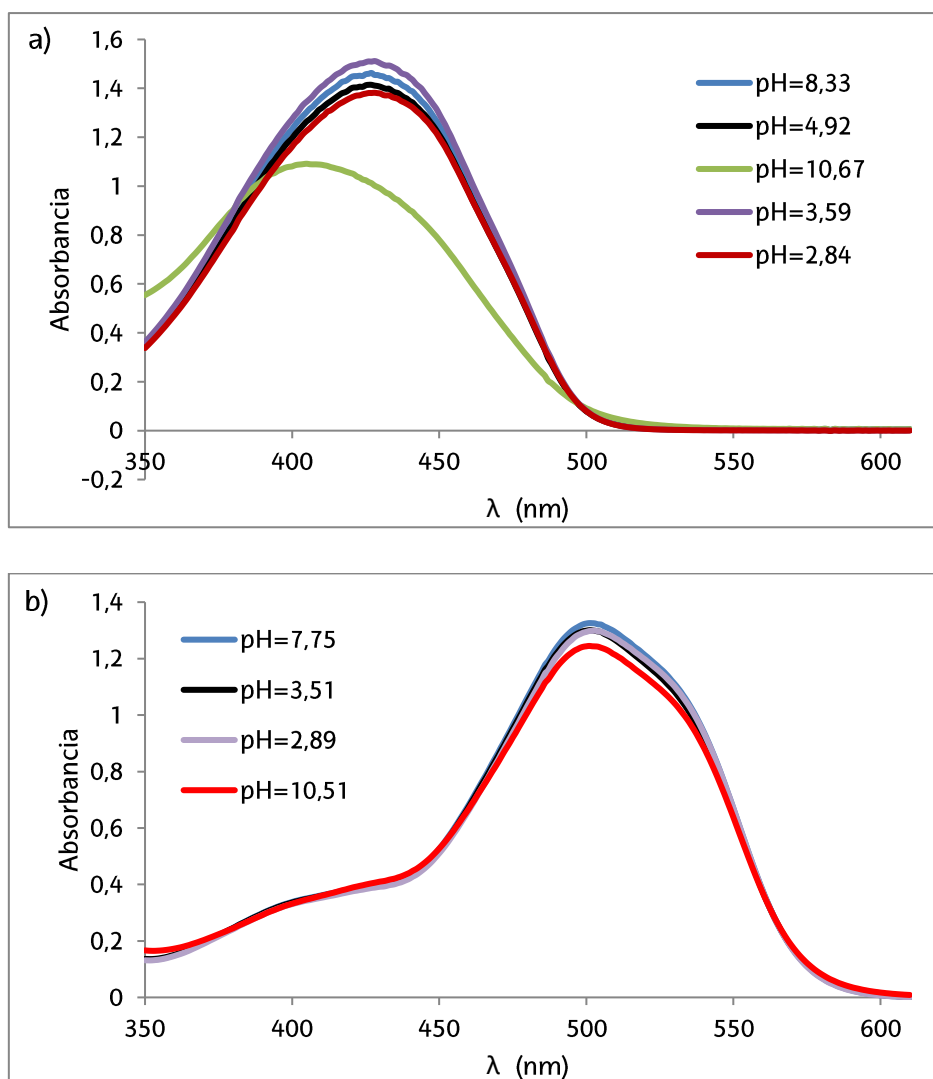


Figura 13. Optimización del pH. a) Espectros de disoluciones de E 102 a distintos pH. b) Espectros de disoluciones de E 129 a distintos pH.

De la figura 13 se puede interpretar que los pH básicos hacen que la absorbancia disminuya, ya que los espectros de las muestras a pH 10,67 para E 102 y a pH 10,51 para E 129 son las que dan un menor valor de absorbancia en el máximo. En el caso del E 129 los espectros son muy parecidos en todo el rango de pH medido, pero en el caso del E 102 un pH básico como es el de 10,67 hace que el espectro sea muy diferente al de las demás muestras. Además, para la

tartrazina, al aumentar el pH, no solo disminuye la absorbancia, si no que el máximo se ve desplazado a longitudes de onda menores, lo que se conoce como desplazamiento hipsocrómico, o desplazamiento al azul.

Por todo ello, se llegó a la conclusión de que el rango de pH óptimo para la medida de los colorantes se encontraba alrededor de un pH de entre 3-8.

5.2. Análisis de imagen

El principal objetivo de este trabajo es el de establecer un método nuevo de medida de colorantes mediante análisis de imagen. A continuación se muestran los resultados obtenidos para la determinación de E 129 en muestras en presencia de E 102.

5.2.1. Patrón de color, medida de la estabilidad de la imagen

A modo de control de la estabilidad de la toma de imágenes se empleó el patrón de color ya mencionado en el apartado de la parte experimental. De esta manera, se puede comprobar que todas las imágenes se han tomado bajo las mismas condiciones de iluminación y por lo tanto los resultados obtenidos no están sometidos a un error sistemático. Para ello, en todas las imágenes tomadas, junto con la placa de microtitulación se colocaba este patrón de la forma que se indica en la figura 9c.

Inicialmente se tomaron 5 imágenes del patrón de color medidas a lo largo de diferentes meses, y se calculó la media del espacio de color RGB para cada uno de los colores del patrón (Figura 14). Con esto se generaron diez diagramas de control, uno para cada uno de los colores (Figura 15) donde se representó la media de las 5 medidas (representada por una línea negra discontinua en los diagramas de control de la figura 15) y sumándole y restándole 2 y 3 veces la desviación estándar para obtener las líneas de aviso (líneas verdes en los diagramas de la figura 15) y las líneas de acción (líneas rojas en los diagramas de la figura 15) respectivamente. Si en las medidas de las imágenes posteriores las líneas de aviso son superadas debemos estar atentos a posibles inestabilidades en la toma de la imagen, mientras que se superan las líneas de acción se deben tomar medidas. En cada uno de los diagramas de la figura 15 están representadas las medias RGB de los colores del patrón de color de todas las imágenes que se han utilizado durante todo el trabajo, tanto de aquellas imágenes empleadas en la calibración y validación del modelo, como las de las muestras reales y las de los cálculos de los parámetros estadísticos.

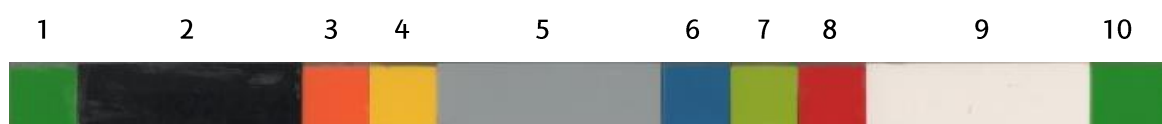


Figura 14. Patrón de color y asignación de un número a cada color. 1: Verde_1; 2: Negro; 3: Naranja; 4: Amarillo; 5: Gris; 6: Azul; 7: Verde_claro; 8: Rojo; 9: Blanco; 10: Verde_2.

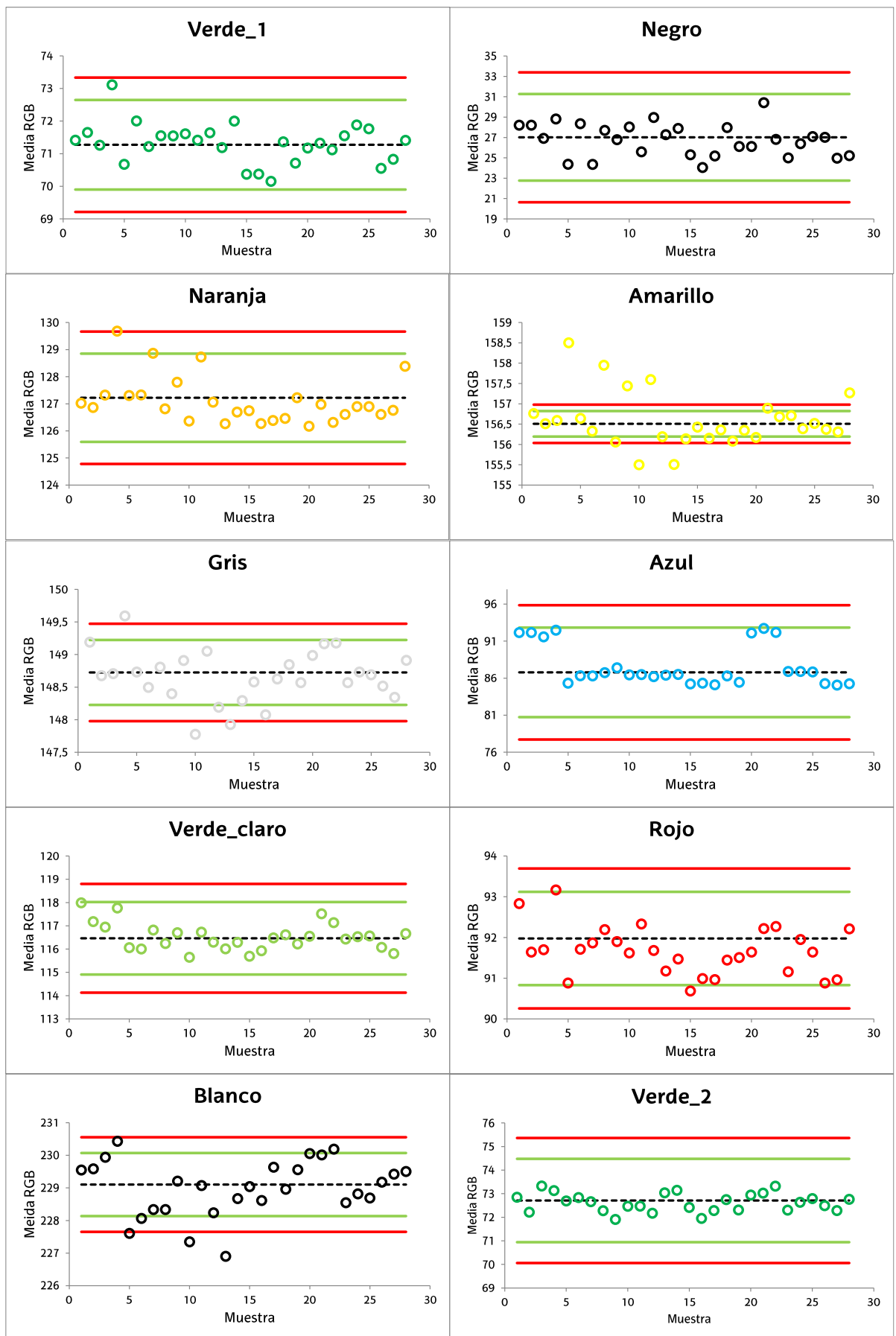


Figura 15. Diagramas de control con la representación de los valores de las imágenes tomadas a lo largo del trabajo para cada uno de los colores del patrón.

Para todos los colores, exceptuando el amarillo, las muestras están en su mayoría dentro de los límites de acción. Cabe destacar la muestra número 4, ya que tiene un comportamiento distinto a las demás muestras, y en muchos de los colores del patrón representados se trata de un punto máximo en los diagramas. Pero como en la mayoría de los casos se encuentra dentro del rango deseado, no se realizó ninguna corrección sobre ella.

Como se puede ver en la figura 15, en el diagrama de control correspondiente al color amarillo, muchas muestras se encuentran fuera de los límites de acción, pero igualmente no se consideró corrección alguna, ya que en los demás colores esto no ocurría. En el caso del color amarillo, lo que ocurre es que el rango que hay entre las dos líneas de acción es muy pequeño (de 156 a 157, una sola unidad) si lo comparamos con cualquier otro color.

Con este análisis, podemos concluir que la toma de imágenes es reproducible, siendo una técnica estable en el tiempo, y que por lo tanto podemos hacer uso de las imágenes tomadas. Sin tener que hacer ningún preprocesado para corregir efectos de la iluminación.

5.2.2. Influencia de los pocillos con muestra sobre pocillos vecinos

Mediante un análisis corto y sencillo se midió la influencia que puede tener el colocar dos muestras en dos pocillos contiguos a la hora de escanear las placas, ya que a simple vista se pudo ver que estos pocillos podían verse afectados. Para ello se cogió una de las imágenes escaneadas (Figura 16) en la que una serie de pocillos habían sido ocupados con diferentes disoluciones preparadas de E 129, y se estudiaron 18 de los pocillos que estaban vacíos. Siempre se estudiaron pocillos que se encontraban contiguos a otros llenos en dirección horizontal según se visualiza la imagen en el ordenador, que es en el eje en el que la influencia parece más evidente.

Los 18 pocillos se dividieron en tres grupos, cada uno con 6 muestras (Tabla 2). El primer grupo está formado por los pocillos que en la figura 16a se enumeran del 1 al 6. Estos se encuentran contiguos a pocillos que contienen una disolución de concentración alta de E 129, por encima de 15 mg/L. Los pocillos de este primer grupo se considera que están muy influenciados. Los del segundo grupo, que se consideraron con influencia baja, estaban junto a pocillos que contenían muestras más diluidas (entre 12,5 y 4 mg/L). Finalmente, el tercer grupo se formó con pocillos que en principio no debían estar influenciadas por no estar contiguos a muestras y se tomaron como referencia.

Como se puede ver en la figura 16b, la media de R, G y B fue elegida para comparar los 18 pocillos. En esa misma figura, se puede comprobar que las medias obtenidas para el grupo 1 tienen una mayor dispersión que las de los otros dos grupos, lo que refuerza la idea de que existe una influencia en los pocillos contiguos. A los valores de los tres grupos se les realizó un análisis

ANOVA de un factor, según el cual existen diferencias significativas entre los tres grupos, al ser la F experimental mayor a la F tabulada en un 95% de confianza (Tabla 3) [34].

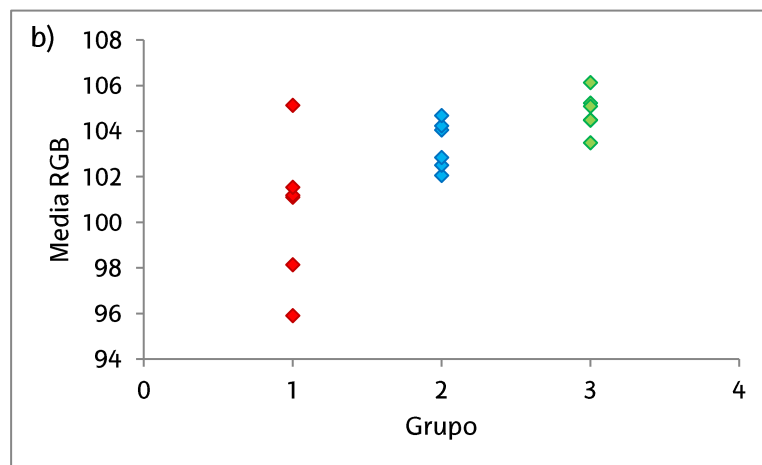
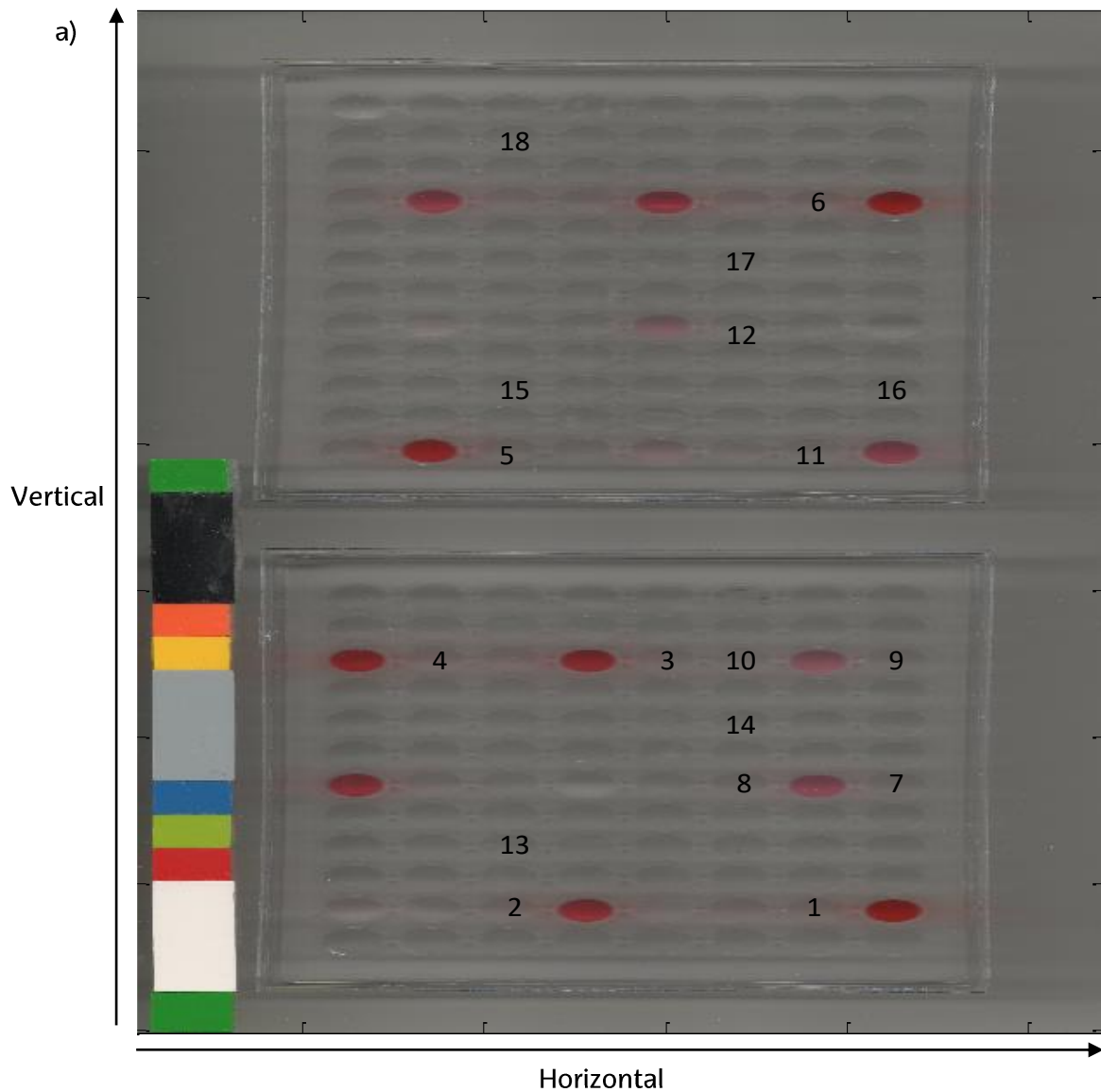


Figura 16. a) Imagen empleada para el estudio de la influencia de los pocillos, indicando los pocillos seleccionados para el mismo. b) Representación del promedio de los canales R, G y B frente a los tres grupos, en rojo el primer grupo, en azul el segundo y en verde el tercero.

Tabla 2. Resumen de los datos empleados para el análisis ANOVA.

| Grupos | Cuenta | Suma | Promedio | Varianza |
|--------|--------|-------|----------|----------|
| 1 | 6 | 602,9 | 100,5 | 10,0 |
| 2 | 6 | 620,3 | 103,4 | 1,14 |
| 3 | 6 | 628,8 | 104,8 | 0,789 |

Tabla 3. Resultados del análisis ANOVA.

| Origen de la varianza | Suma de cuadrados | Grados de libertad | Promedios de los cuadrados | F _{experimental} | Valor de P | F _{tabulada} |
|-----------------------|-------------------|--------------------|----------------------------|---------------------------|------------|-----------------------|
| Entre grupos | 58,1 | 2 | 29,0 | 7,30 | 0,00611 | 3,68 |
| Dentro de grupos | 59,7 | 15 | 3,98 | | | |
| | | | | | | |
| Total | 117,7 | 17 | | | | |

Tras este estudio, se llegó a la conclusión de que los pocillos vacíos contiguos a uno con muestra se ven afectados por la presencia de ésta, sobre todo los que están a la derecha o a la izquierda de la muestra en la imagen tomada de la placa.

Este efecto de que los pocillos contiguos en la dirección horizontal a otros que contengan muestra se vean afectados puede deberse a que al moverse la luz del escáner a lo largo de la dirección vertical, la luz reflejada por las muestras se expanda algo en la dirección contraria al movimiento, influyendo a los pocillos que ahí se encuentran.

Con ello, se tomó la decisión de no colocar ninguna muestra contigua a otra en la dirección horizontal.

5.2.3. Modelos de calibración para los colorantes a partir de las imágenes

Sabiendo la forma más adecuada de llenar los pocillos se procede a la obtención de las imágenes, dando lugar a imágenes como la figura 17. Como puede observarse, con una única imagen pueden medirse simultáneamente varias muestras. Se lleva a cabo el recorte de las mismas, para conseguir imágenes de 21 x 21 x 3 píxeles para cada muestra seleccionando de la imagen completa las regiones de interés (ROI) (figura 17):

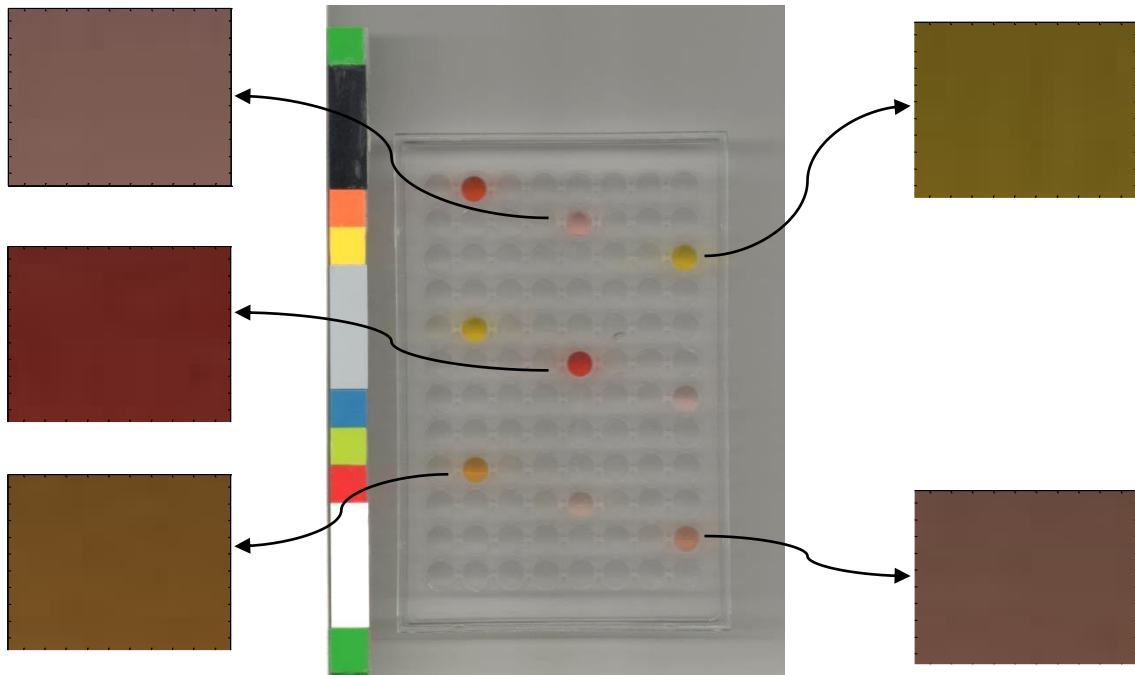


Figura 17. Ejemplo de una imagen escaneada empleada durante el trabajo. Imágenes ampliadas de algunos recortes de los pocillos de 21 x 21 x 3 píxeles.

Una vez realizados los recortes se extrajeron tanto variables individuales como colorgramas de las imágenes que pudieran definir el color de cada muestra y por consiguiente la concentración de los colorantes.

5.2.3.1. Calibración univariable

Lo primero que se intentó aplicar fue la calibración univariable. En este caso, no sólo se intentó buscar un modelo de calibración para rojo allura, sino también para tartrazina con la media de alguno de los canales R, G o B. Para ello se representaron los valores citados frente a las concentraciones de cada colorante, obteniéndose las gráficas que se muestran en la figura 18.

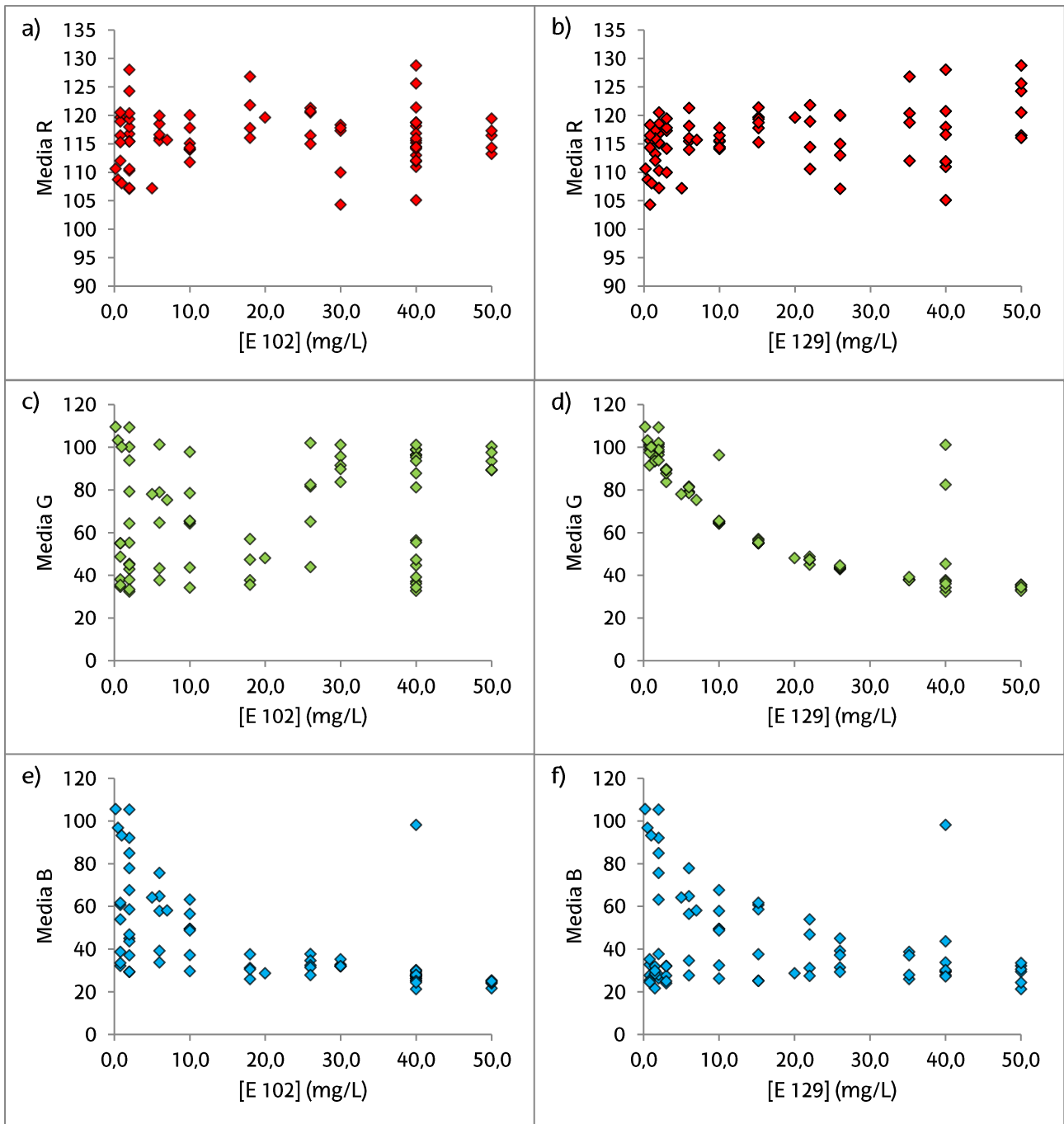


Figura 18. Calibración univariable para ambos colorantes respecto de las medias R (en rojo), G (en verde) y B (en azul). a) Representación de la media R frente a la concentración de E 102. b) Representación de la media R frente a la concentración de E 129. c) Representación de la media G frente a la concentración de E 102. d) Representación de la media G frente a la concentración de E 129. e) Representación de la media B frente a la concentración de E 102. f) Representación de la media B frente a la concentración de E 129.

Como era de esperar, el análisis individual de los canales R, G ó B no se relaciona de manera sencilla con las concentraciones de los colorantes en la mezcla, probablemente porque el color final en cada canal no depende únicamente de la concentración de uno de los colorantes, si no que existe dependencia del otro colorante. En ninguno de los casos se aprecia un comportamiento lineal en el que se relacione la concentración con la media de alguno de los canales. En todo caso, al representar la media del canal B frente a ambas concentraciones se observa un comportamiento que podría ser logarítmico. Además a medida que aumenta la concentración del colorante estudiado la dispersión de los datos disminuye, algo que es lógico, ya que cuanto mayor sea la presencia de uno de los colorantes, el color se verá menos afectado por la presencia del otro. Este comportamiento logarítmico, puede verse también en la relación entre la media G y [E 129] (mg/L) (Figura 18d). Por ello, y a pesar de las tendencias logarítmicas observadas, se descartó este tipo de calibración y se optó por una calibración multivariable con la esperanza de obtener modelos lineales y más robustos, como son los que a continuación se describen.

5.2.3.2. Calibración multivariable siguiendo un método de la bibliografía

Seguido de este sencillo intento de calibración, se optó por un procedimiento basado en la obtención de un colorgrama encontrado en el trabajo de A. Ulrici et. al. ^[43], el cual se explica en el apartado 4.2.5.1. y en la figura 10. Se decidió empezar por encontrar un modelo capaz de predecir la concentración de rojo allura en presencia de interferencia de tartrazina. En función de los resultados obtenidos, se podría pantear como trabajo futuro la obtención de un modelo predictivo para tartrazina en presencia de rojo allura.

Para ello se empleó la matriz de concentraciones indicada en la figura 9, y de cada una de las muestras se calcularon los parámetros necesarios para formar el colorgrama y crear así una matriz de datos de 73 x 3340. Una vez hecho esto, con las 49 muestras de calibración, se realizó un PLS con la *PLS_toolbox*.

Algunas muestras fueron eliminadas del modelo porque se comprobó que en muestras en las que la concentración de uno de los dos colorantes, o la suma de las concentraciones era muy grande, el color intenso hacía que se saturase el color de la imagen y no fuera posible realizar una buena predicción a partir de ellas. Para visualizar las muestras que fueron eliminadas, en la figura 19a se puede ver cómo en la matriz de concentraciones empleada en el análisis de imagen se ha introducido un línea recta. Las muestras con concentraciones por encima de la recta son las que fueron eliminadas, figura 19b.

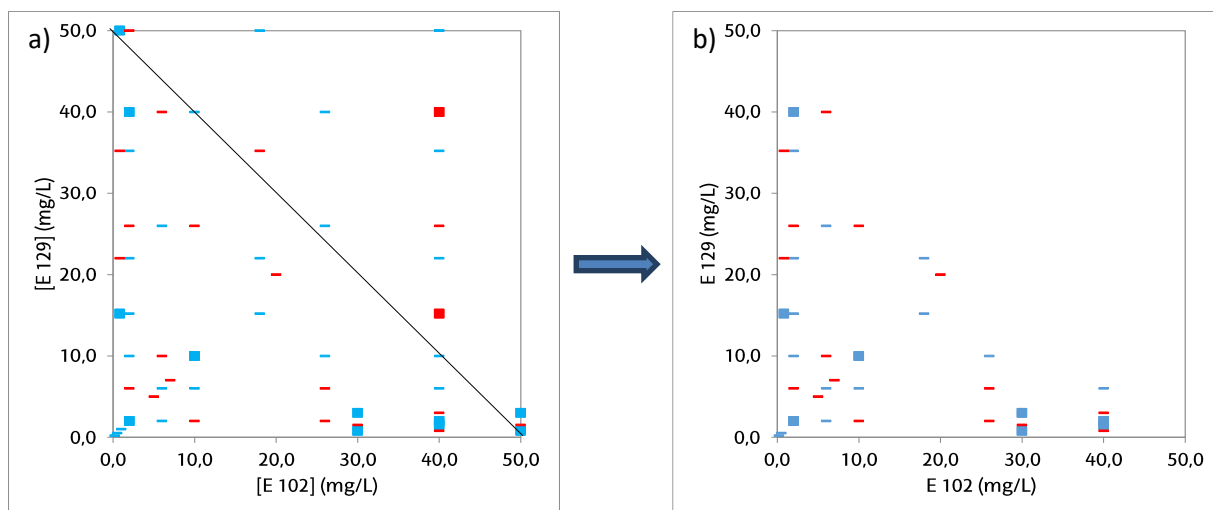


Figura 19. a) Matriz de concentraciones inicial y sobre ella la línea recta. b) Matriz de concentraciones final. En azul las muestras de calibración y en rojo las de validación. Las líneas corresponden a muestras únicas mientras que los cuadrados representan réplicas.

Además de estas muestras también se eliminó alguna que otra muestra a la hora de definir los modelos siguiendo las recomendaciones del programa, por tener un residual y/o un leverage muy alto. Un residual alto indica que la muestra no está bien explicada por el modelo, mientras que el leverage indica cuanto se aleja el comportamiento de esa muestra de la de las demás que conforman el modelo y por lo tanto cuanto lo puede distorsionar. Además de varias muestras también se eliminaron algunas variables atendiendo a la poca variabilidad que explicaban en relación a la concentración. Así pues, finalmente la matriz de datos con la que se realizó el modelo PLS fue una matriz de 29 muestras y 119 variables (matriz X), mientras que, los valores de concentración conocidos de E 129 formaban el vector y .

La figura 20 muestra los valores predichos por el modelo para la concentración de E 129 en mg/L para los datos de la calibración (Cal) junto a los obtenidos de validación interna o *cross-validation* (CV), una validación que se realizó mediante el método de *leave one out*, que consiste en realizar tantos calibrados como muestras haya, dejando en cada caso una fuera y prediciendo su valor.

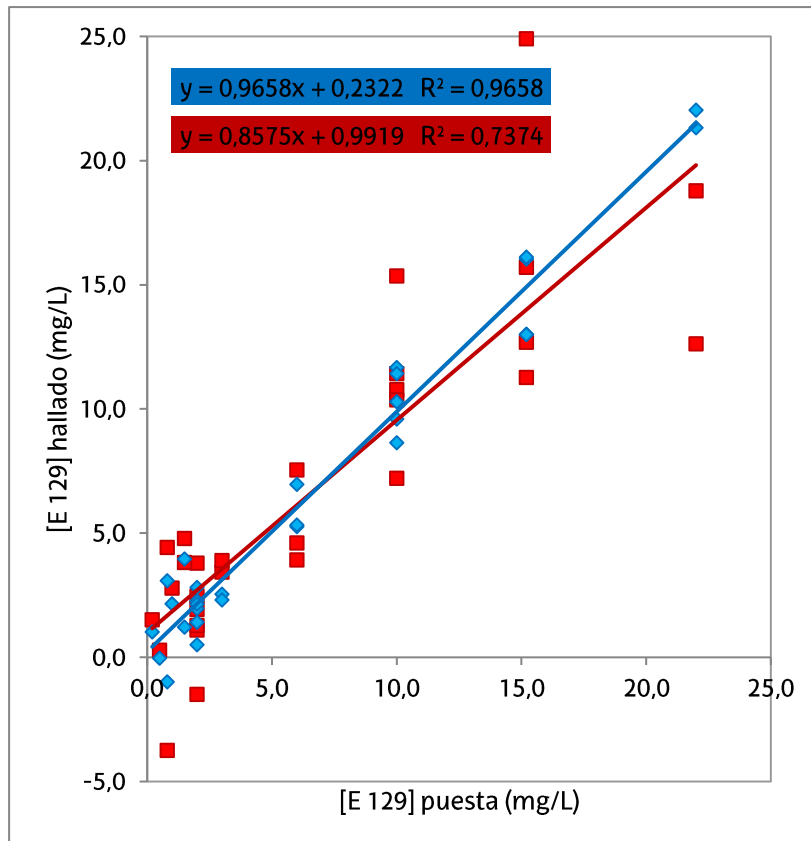


Figura 20. Representación del hallado vs puesto de las concentraciones de E 129 utilizando el método de la bibliografía. Los rombos en azul se refieren a las muestras calibración, mientras que los cuadrados rojos a las de *cross-validation*. Se dan, además, las rectas de regresión obtenidas.

Este modelo PLS que se obtuvo con 4 PCs (con los cuales solo se lograba explicar un 64,19% de la variabilidad de la matriz X y un 96,58% de las concentraciones en y) no resultó lo suficientemente satisfactorio, por los altos errores y por su diferencia entre el calibrado y la validación interna (Tabla 4). Como se puede ver en la tabla 4, donde se recogen los valores de los errores y de las regresiones de este modelo, este está totalmente sobreajustado, ya que a pesar de que el error en la calibración es aceptable (cercano al 10%), en la *cross-validation* este es muy alto en comparación con el primero. Al emplear un modelo con menos PCs se obtenían errores mayores, y no se eliminaba el problema del sobreajustado.

Por eso se procedió a la realización de un nuevo método, al cual se le llamó *método propio*. Viendo este mal resultado no se calculó el límite de detección ni la validación externa y directamente se pasó a buscar otro modelo. También cabe destacar que la obtención de la matriz de datos con la cual se lleva a cabo el *método bibliográfico* es mucho más laboriosa que el nuevo que se propone.

Tabla 4. Datos de la calibración y *cross-validation* del método bibliográfico.

| | Ecuación | R ² | E.R. % |
|-----|-----------------|----------------|--------|
| Cal | 0,9658 + 0,2322 | 0,9658 | 12,73 |
| CV | 0,8575 + 0,9919 | 0,7374 | 36,59 |

5.2.3.3. Calibración multivariable por el método propio

Al ver que el modelo anterior no fue satisfactorio, por los motivos explicados en el apartado 5.2.3.2., se optó por estudiar el que se llamó *método propio*. Como ya se ha explicado en el apartado 4.2.5.2., este es un método en el cual las variables que se emplean son muy sencillas de obtener, lo que facilita mucho el trabajo.

Con el mismo objetivo que antes, el de la determinación del colorante E 129 en presencia de E 102, se realizó la misma eliminación de muestras que se indica en la figura 17, y una vez reducido el número de muestras se intentó obtener un modelo satisfactorio. Para ello, además de esas muestras que se encontraban por encima de la línea recta se eliminó alguna que otra muestra siguiendo el mismo razonamiento que en el apartado anterior, atendiendo a su valor del residual y del *leverage*. Además de muestras, también se eliminaron variables, ya que, de forma empírica, se vio que al quedarse únicamente con las variables del espacio de color RGB y las del LAB el modelo mejoraba considerablemente. El modelo se obtuvo con una matriz de 28 muestras y 6 variables y con 4 PCs. El LOD y el LOQ se calcularon como se explica en el apartado 4.5. y se obtuvieron valores de 1,8 mg/L y 5,9 mg/L respectivamente.

Así pues, los puntos que se encontraron por debajo del LOD también se eliminaron, quedándose únicamente 19 muestras para calibración y validación interna, en un rango de concentraciones entre 3,0 y 35,2 mg/L para E 129 con interferencia del otro colorante en concentraciones que se encontraban en el rango entre 0,8 y 30,0 mg/L. Se obtuvieron las rectas hallado vs puesto que se muestran en la figura 21 con 4 PCs, con los cuales se explicaba el 99,98% de la variabilidad de la matriz X y un 98,67% de la concentración de E 129 (y). En este caso, la calibración, además de con *cross-validation leave one out*, también se validó por medio de una validación externa, con 15 muestras que no habían sido incluidas en el modelo (validación externa). Las muestras de este grupo se redujeron de 24 a 15, ya que se eliminaron las muestras con concentraciones muy altas como en la obtención del modelo. En la figura 21 se pueden ver las rectas de regresión obtenidas para la calibración, la validación interna y validación externa.

Los errores relativos obtenidos con este modelo (Tabla 5), tanto para la calibración ($ER_{Cal}\%$), como para la *cross-validation* ($ER_{CV}\%$) y la validación externa ($ER_{Val}\%$), todos ellos entre 5,7% - 12,4%, mejoran los valores obtenidos con el anterior método reflejados en la tabla 4.

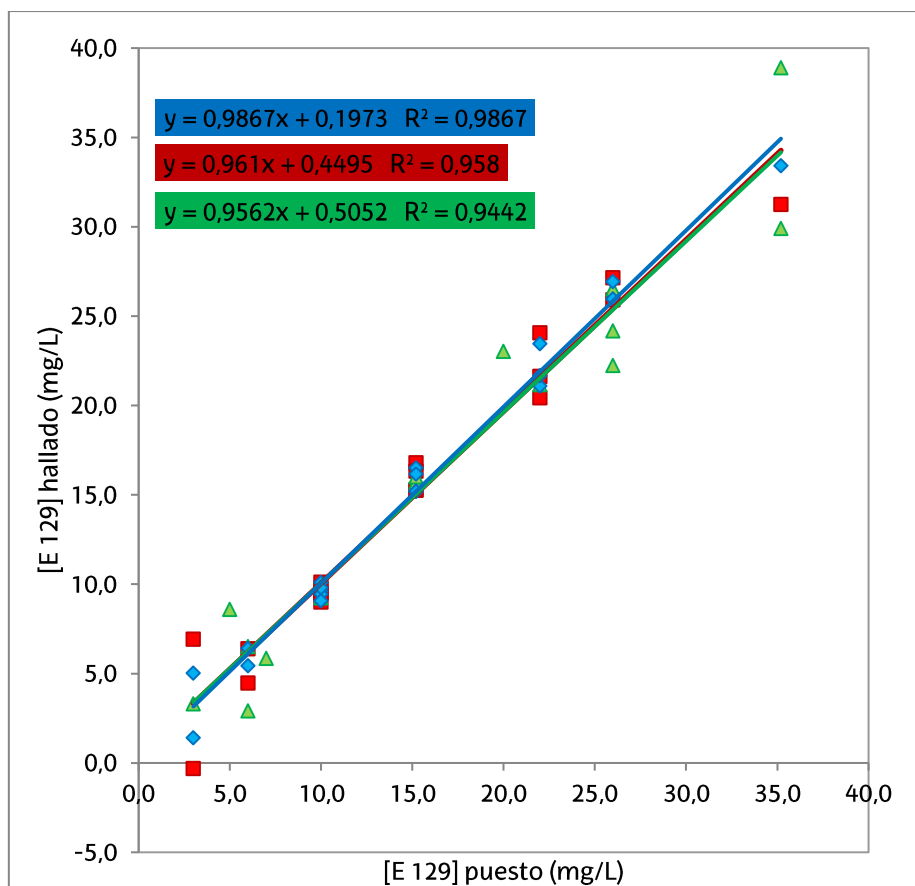


Figura 21. Representación del hallado vs puesto de las concentraciones de E 129 utilizando el método propio. Los rombos en azul se refieren a las muestras de calibración, mientras que los cuadrados rojos a las muestras de *cross-validation*. Por último, los triángulos verdes corresponden a las muestras de la validación externa.

La precisión del método se evaluó calculando la desviación estándar relativa (RSD%, *Relative Standar Derivation*) de la reproducibilidad y la repetibilidad a dos niveles de concentraciones diferentes (Tabla 5) como se explica en el apartado 4.2.2. Para la reproducibilidad se obtuvieron valores de RSD% de 4,8% y 12,2%, valores parecidos a los obtenidos para la repetibilidad de 5,4% y 14,2%. En ambos casos la mayor desviación corresponde al nivel más bajo de concentración. Por debajo de un 10% los valores de la desviación estándar se consideran buenos. En este caso los valores para el nivel alto están por debajo y por lo tanto son satisfactorios, mientras que los del nivel bajo, superan por poco ese valor y por lo tanto también se consideran aceptables.

Por otro lado, la exactitud se midió por medio de la recuperación en muestras reales. Para los dos lotes de flash (F1 y F2) se obtuvieron recuperaciones de E 129 en los rangos de 94,7% - 104% y 102% - 106% respectivamente. En el caso del colorante alimentario también se emplearon dos disoluciones iniciales (D1 y D2), y en este caso los rangos de recuperación teniendo en cuenta ambas fueron 90,7% - 99,7% y 95,4% - 97,9%, respectivamente. En todos los casos las recuperaciones obtenidas son cercanas al 100%, y por ello muy satisfactorios.

Tabla 5. Parámetros analíticos de calidad del modelo para E 129 con análisis de imagen y el método propio.

| | Rango lineal (mg/L) | R ² | LOD (mg/L) | LOQ (mg/L) | Errores relativos (E.R.) (%) | | | Reproducibilidad (RSD%) | | Repetibilidad (RSD%) | | Recuperaciones (%) | | | |
|-------|---------------------|----------------|------------|------------|------------------------------|-------------------|---------------------|-------------------------|---------|----------------------|---------|--------------------|-----------|-------------|-------------|
| | | | | | ER _{Cal} % | ER _c % | ER _{Val} % | 10 mg/L | 26 mg/L | 10 mg/L | 26 mg/L | F1 | F2 | D1 | D2 |
| E 129 | 3-35,2 | 0,987 | 1,8 | 5,9 | 5,7 | 10,2 | 12,4 | 12,2 | 4,8 | 14,2 | 5,4 | 94,7 -104 | 102 - 106 | 90,7 - 99,7 | 95,4 - 98,0 |

5.3. Espectroscopía ultravioleta-visible

Como se ya se ha citado antes, la técnica de la espectroscopia UV-Vis se empleó para validar la técnica del análisis de imagen. En este caso se crearon modelos de calibración tanto para E 102 como para E 129.

Para ello, y con la matriz de calibración de la figura 12, se utilizó el programa *The Unscrambler*. En el caso del colorante E 102, las muestras que contenían una concentración de este superior a los 40 mg/L no fueron empleadas ya que en el máximo de absorbancia se superaban las 2 unidades de absorbancia, y por lo tanto dejaba de cumplirse la Ley de Beer.

Para esta técnica de referencia se calcularon los mismos parámetros estadísticos que con la nueva técnica para poder validarla. Para empezar se realizaron las calibraciones para ambos colorantes. En este caso, y dado la diferencia espectral entre los espectros de ambos colorantes (Figura 13), se esperaba obtener un buen modelo de calibración para cada uno de ellos. Se validaron ambos, por un lado mediante *cross-validation* pero en este caso con Random, que en vez de dejar una única muestra fuera y calcular un modelo con las demás (como en *el leave-one-out*), divide todas las muestras en grupos aleatorias que son las que va sacando fuera para luego predecirlas, y por otro mediante una validación externa con muestras que no participaban en el modelo. Para los modelos de ambos colorantes se escogieron 2 PCs, con los que se lograba explicar el 100% de la varianza tanto en las variables (X) como en las concentraciones (y).

El colorante E 102 se modeló con 31 muestras para la calibración (ya que se eliminaron las muestras por encima de 40 mg/L y las que posteriormente se encontraron por debajo del LOD) y 37 muestras se dejaron para la validación externa. El número original de muestras de validación se redujo por el mismo motivo. La figura 22a muestra estas representaciones.

Por otro lado, para el colorante E 129, el número de muestras empleadas para la calibración fue 45, dos menos de las 47 iniciales que se eliminaron por encontrarse por debajo del LOD, mientras que en la validación externa se emplearon las 47. En la figura 22b, se pueden ver los resultados de predicción obtenidos para las muestras de calibración, *cross-validation* y validación externa.

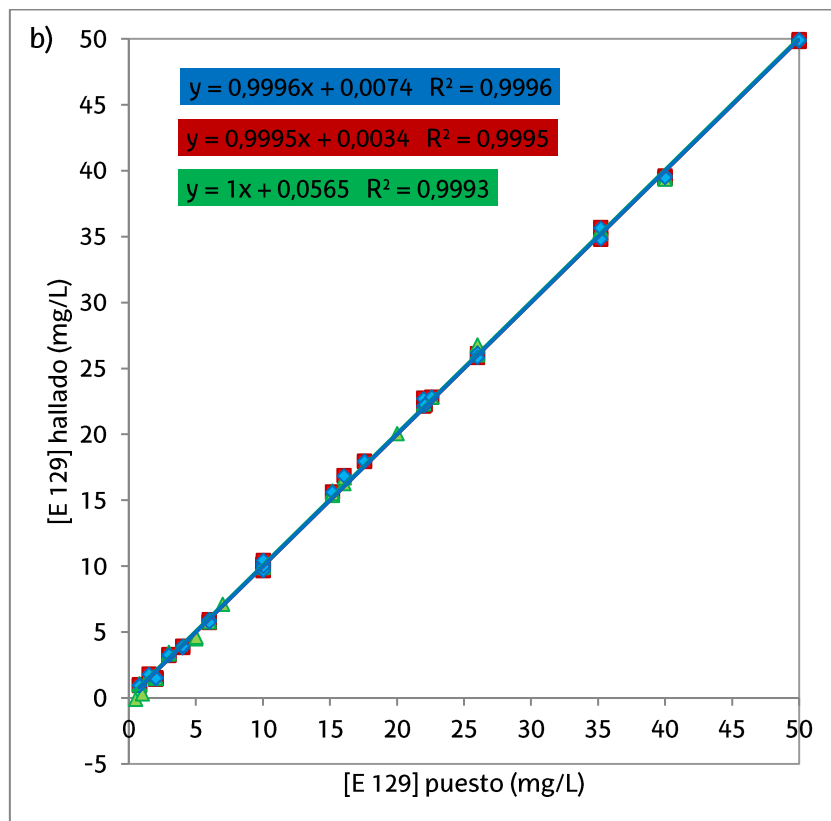
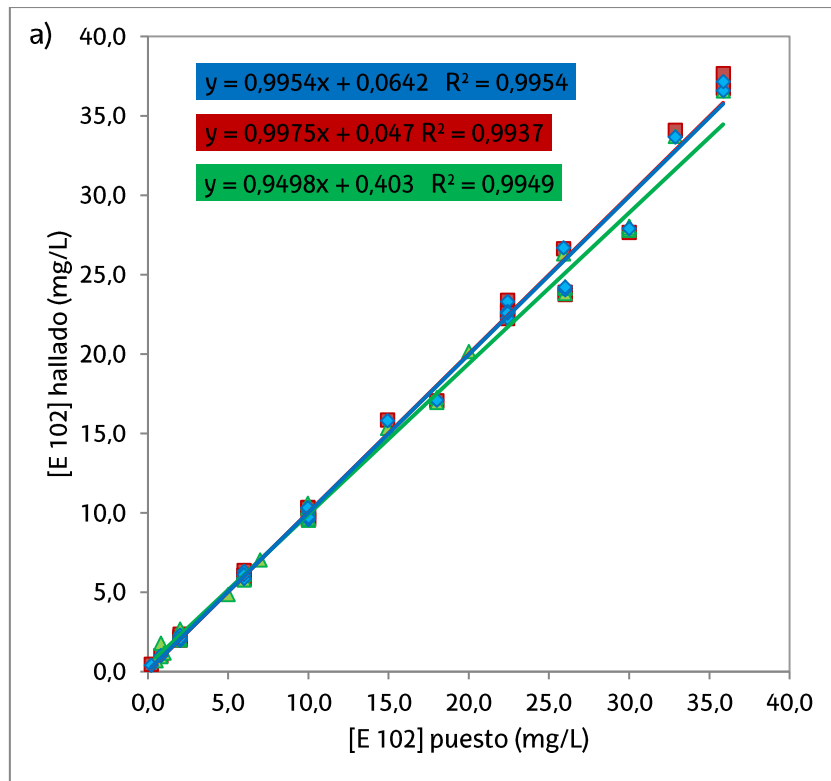


Figura 22. Representación del puesto vs hallado de las concentraciones para la espectrofotometría. a) Concentraciones de E 102. b) Concentraciones de E 129. Los rombos en azul se refieren a las muestras de calibración, mientras que los cuadrados rojos a las muestras de *cross-validation*. Por último, los triángulos verdes corresponden a las muestra de la validación externa.

Una vez más, la precisión se estudió calculando la RSD% de la reproducibilidad y repetibilidad, en este caso a una única concentración y para los dos colorantes. Para la tartrazina, los valores obtenidos para la reproducibilidad y repetibilidad fueron de 0,54% y 0,57% respectivamente, mientras que para el rojo allura fueron de 1,2% y 0,80%. (Tabla 6)

Finalmente, la exactitud, volvió a medirse por medio de la recuperación. Para el E 102, con el lote 1 de flash (F1) se obtuvieron recuperaciones entre 99,7% y 104% mientras que para el lote 2 (F2) fue entre 100% y 109%, mientras que en el colorante alimentario comercial, se obtuvieron valores entre 97,6% y 98,2% para D1 y entre 99,4% y 103% para D2. En cambio, para el E 129, con el F1 las recuperaciones obtenidas se encontraban en un rango entre el 94,6% y 104% y para el F2, 96,8% y 106%, y en el caso del colorante alimenticio se obtuvieron rangos de 100% - 104% y 102% - 114% para el D1 y D2 respectivamente.

Como cabía esperar los resultados obtenidos del cálculo de la exactitud y de la precisión para esta técnica son muy buenos, y cabe destacar también, que para ambos colorantes, todos los errores relativos calculados se encuentran por debajo del 6%.

Todos estos resultados se pueden ver mejor en la tabla 6, que está a continuación.

Tabla 6. Parámetros analíticos de calidad de los modelos para ambos colorantes con la técnica de espectroscopía UV-Vis.

| | Rango lineal (mg/L) | R ² | LOD (mg/L) | LOQ (mg/L) | Errores relativos (E.R.) (%) | | | Reproducibilidad (RSD%) | Repetibilidad (RSD%) | Recuperaciones (%) | | | |
|--------------|---------------------|----------------|------------|------------|------------------------------|--------------------|---------------------|-------------------------|----------------------|--------------------|------------|-------------|------------|
| | | | | | ER _{cal} % | ER _{cv} % | ER _{val} % | | | 0,08 mg/L | 0,08 mg/L | F1 | F2 |
| E 102 | 0,2-35,9 | 0.995 | 0.14 | 0.46 | 4,3 | 5,0 | 5,7 | 0.54 | 0.56 | 99,7 - 104 | 100 - 109 | 97,6 - 98,2 | 99,4 - 103 |
| E 129 | 0,8-50,0 | 0.999 | 0.65 | 2.2 | 1,4 | 1,5 | 1,7 | 1.20 | 0.80 | 94,6 - 104 | 96,8 - 105 | 100 - 104 | 102 - 114 |

5.4. Comparación del método del análisis de imagen con el de referencia

Como se ha venido diciendo durante todo el trabajo, el objetivo ha sido desarrollar un método para la determinación de colorantes mediante análisis de imagen que fuera validado por medio de una técnica muy contrastada como es la espectrofotometría UV-Vis.

Para ello, se representaron las predicciones de las concentraciones de E 129 de las muestras reales (siempre y cuando estas estuvieran por encima de los límites de detección) de las del análisis de imagen frente a la técnica de referencia, como se puede ver en la figura 24.

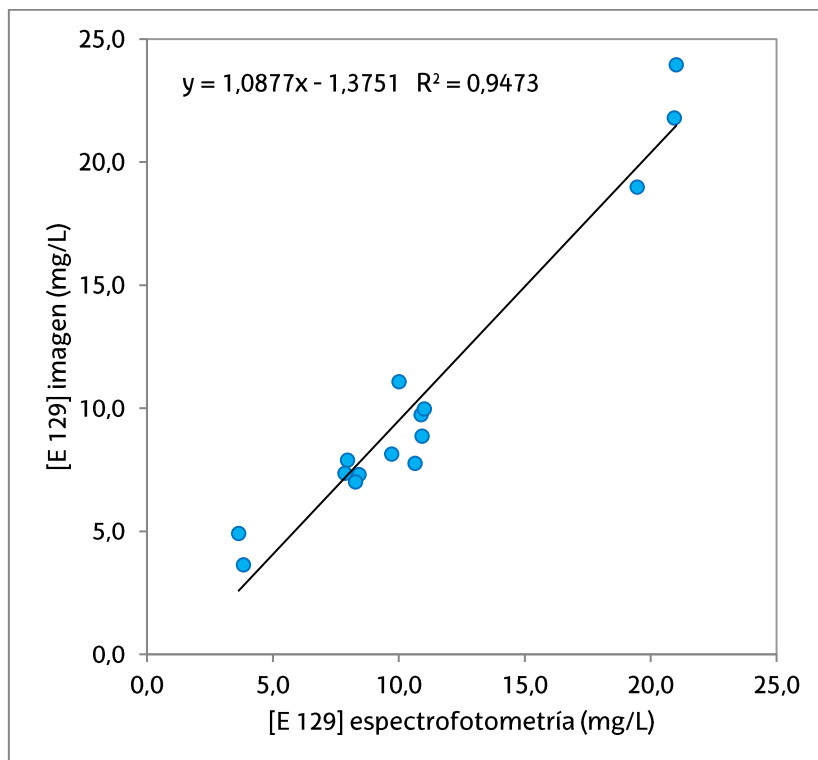


Figura 23. Comparación de las predicciones de la espectrofotometría UV-Vis y el análisis de imagen en muestras reales.

De esta representación en la figura 23, se espera obtener una regresión cuya pendiente sea 1 y su ordenada en el origen (0,0), ya que lo ideal sería que con ambas técnicas se predijeran los mismo valores de concentración de E 129 en las muestras reales.

Con el objetivo de comprobar que no existen diferencias significativas entre ambos métodos se aplicó a los datos de la figura 23 el test de la elipse a un 95% de confianza (Figura 24), donde EJCR se refiere a región de confianza elíptica (*Elliptical Joint Confidence Region*), en cuyo centro se sitúan los valores de pendiente y ordenada en el origen de la recta de la figura 23. Este test también ha sido empleado con el mismo objetivo de confirmar la fiabilidad de un nuevo método en el trabajo de L. Zhi et. al.^[45].

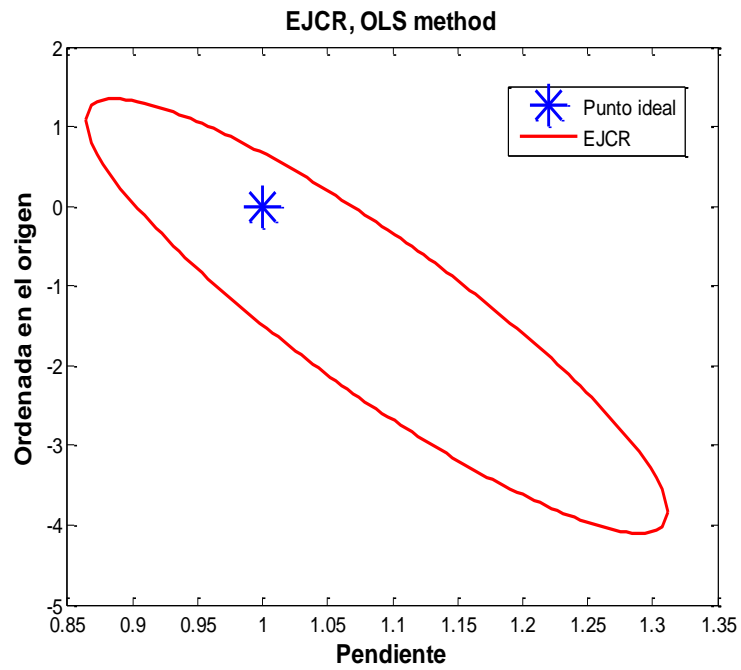


Figura 24. Resultado del test de la elipse en un 95% de confianza.

Al encontrarse el punto ideal, que indica pendiente 1 y ordenada en el origen 0, dentro de la elipse, se puede afirmar que no existen diferencias significativas al aplicar uno u otro método.

6. Conclusiones

Se ha logrado desarrollar un método para la predicción de rojo allura en presencia de tartrazina por medio del análisis de imagen. Para ello se ha construido un modelo PLS, relacionando el color de la muestra con parámetros de la imagen sencillos de obtener. Este modelo se ha validado y comparado con la espectrofotometría.

Como era de esperar los resultados obtenidos con el método de referencia, son mejores a los obtenidos con el nuevo método aquí desarrollado (menor LOD, mayor rango lineal, menores errores, etc). Sin embargo, hay que tener en cuenta, que el análisis de imagen es una técnica muy sencilla y que en este caso emplea un escáner de oficina muy sencillo y barato a diferencia del espectrofotómetro necesario para el desarrollo de la espectrofotometría. Además, las placas de microtitulación empleadas en este trabajo permiten la medida simultánea de más de una muestra con una sola imagen, lo que se refleja en una reducción del tiempo a la hora de medir las muestras. La tercera ventaja que cabe destacar de este método, es la disminución de volumen de muestra a emplear, de unos 3 mL a 400 μ L.

Por otra parte, se ha conseguido emplear el modelo obtenido por medio del análisis de imagen para la predicción de rojo allura en muestras reales, predicciones que han sido validadas con el método de referencia, viendo que no existen diferencias significativas en el empleo de cualquiera de los dos métodos.

También se ha llevado a cabo un control de la estabilidad de las medidas empleando el patrón de color, y comprobando que las imágenes obtenidas son comparables sin ser necesaria la corrección de ninguna de estas.

Como un trabajo futuro, queda pendiente el desarrollo de un modelo con el que se logre predecir la concentración de la tartrazina en presencia de rojo allura, al igual que el empleo de otro tipo de dispositivo de toma de la imagen, como puede ser una cámara digital de un smartphone.

Aun así, los resultados obtenidos en este trabajo resultan satisfactorios, haciendo que sea posible emplear este método en la determinación de rojo allura en muestras, a pesar de que haya interferencias y cambios en el color por acción de la tartrazina.

6. Conclusions

A method to the prediction of allura red in presence of tartrazine has been developed. A PLS model has been built, which relates the color of the sample with easy parameters obtained from digital image. The model has been validated by comparing the results with the obtained by spectrophotometry.

As it was expected, the results obtained with the reference method are better than the ones with the new method (lower LOD, higher linearity range, lower errors, etc). Nevertheless, the image analysis is a very simple technique and it is only necessary a simple and cheap desktop scanner unlike the sophisticated apparatus that is used for the spectrophotometry. In addition, microtitration plates allows the simultaneous measurement of several samples in one image, so the time to take the measurements shortens. The third advantage of this method is that it needs less volume of sample compared to spectrophotometry, 400 μ L vs 3 mL.

On the other hand, the model obtained by image analysis has been used to predict allura red in real samples. The predictions have been validated with the reference method, and no significant differences between both methods have been arisen.

A control of the stability of the measurements was also carried out using the color pattern and checking that the obtained images are comparable without being necessary the correction of any of these.

As future work, the development of one model to predict the concentration of tartrazine in presence of allura red is pending. Likewise, the use of another device for image capture, for example a smartphone's digital camera.

Even so, the results obtained in this work are satisfactory, making possible the use of this method for the determination of allura red in samples, in spite of interferences and changes in color by the presence of tartrazine.

7. Referencias

- [1] http://www.aecosan.msssi.gob.es/AECOSAN/web/seguridad_alimentaria/subdetalle/aditivos_alimentarios.htm consultada el 08/06/2017.
- [2] J. J. Berzas Nevado, C. Guiberteau Cabanillas, A. M. Contento Salcedo "Simultaneous spectrophotometric determination of three food dyes by using the first derivative of ratio spectra" (1995) *Talanta* 42: 2043-2051.
- [3] P. Amchova, H. Kotolova, J. Ruda-Kucerova "Health safety issues of synthetic food colorants" (2015) *Regulatory Toxicology and Pharmacology* 73: 914-922.
- [4] https://es.m.wikipedia.org/wiki/Archivo:Skittles_in_HDR.jpg
- [5] J. Jiao, J. Wang, M. Li, J. Li, Q. Li, Q. Quan, J. Chen "Simultaneous determination of three azo dyes in food product by ion mobility spectrometry" (2016) *Journal of Chromatography B* 1025: 105-109.
- [6] M. Bastaki, T. Farrel, S. Bhusari, K. Pant, R. Kulkarni, "Lack of genotoxicity in vivo for food color additive Tartrazine" (2017) *Food and Chemical Toxicology* 105: 278-284.
- [7] D.W. Schab, N.H. Trinh "Do artificial food colors promote hyperactivity in children with hyperactive syndromes? A meta-analysis of double-blind placebo-controlled trials" (2004) *Journal of Developmental Behavioral Pediatrics* 25: 423-434.
- [8] D. McCann, A. Barrett, A. Coope, D. Crumpler, L. Dalen, K. Grimshaw, E. Kitchin, K. Lok, L. Porteous, E. Prince, E. Sonuga-Barke, J.O. Warner, J. Stevenson "Food additives and hyperactive behavior in 3-year-old and 8/9-year-old children in the community: a randomized, double-blinded, placebo-controlled trial" (2007) *Lancet* 370: 1560-1567.
- [9] J.T. Nigg, K. Lewis, T. Edinger, M. Falk "Meta-analysis of attention-deficit/hyperactivity disorder as attention-deficit/hyperactivity disorder symptoms, restriction diet, and synthetic food color additives" (2012) *Journal of the American Academy of Child and Adolescent Psychiatry* 51:86-97.
- [10] Blanca Del Rio Navarro "Urticaria and angioedema" (2007) *Revista Alergia México* 54(2): 54-65.
- [11] REGLAMENTO (CE) N° 178/2002 DEL PARLAMENTO EUROPEO Y DEL CONSEJO (28.01.2002). Reglamento (CE) N° 178/2002 del parlamento europeo y del consejo de 28 de enero de 2002 por el que se establecen los principios y los requisitos generales de la legislación alimentaria, se crea la Autoridad Europea de Seguridad Alimentaria y se fijan procedimientos relativos a la seguridad alimentaria. *Diario Oficial de las Comunidades Europeas*, (L 31/ 1-24).
- [12] <http://www.efsa.europa.eu/en/topics/topic/food-colours> consultada el 14/06/2017.

- [13] REGLAMENTO (CE) N° 1333/2008 DEL PARLAMENTO EUROPEO Y DEL CONSEJO (16.12.2008). Reglamento (CE) N° 1333/2008 del parlamento europeo y del consejo de 16 de diciembre de 2008 sobre aditivos alimentarios. *Diario Oficial de la Union Europea*, (L 354/16-33).
- [14] REGLAMENTO (UE) N° 1129/2011 DE LA COMISIÓN (12.11.2011). Reglamento (UE) N° 1129/2011 de la comisión de 11 de noviembre de 2011 por el que se modifica el anexo II del Reglamento (CE) no 1333/2008 del Parlamento Europeo y del Consejo para establecer una lista de aditivos alimentarios de la Unión. *Diario Oficial de la Union Europea*, (L 295/1-177).
- [15] Carlos Barros Santos, "Los aditivos en la alimentación de los españoles. La legislación que regula su autorización uso". 2 Tomos. (2008) página 55.
- [16] M. Bastaki, T. Farrel, S. Bhusari, K. Pant, R. Kulkarni, "Lack of genotoxicity in vivo for food color additive Allura Red AC" (2017) *Food and Chemical Toxicology* 105: 308-314.
- [17] K. Yamjal, M. Subramania Nainar, N. Rao Ramiseti "Methods for the analysis of azo dyes employed in food industry – A review" (2016) *Food Chemistry* 192: 813-824.
- [18] F. Turak, M. Ustun Ozgur "Validated Spectrophotometric Methods for Simultaneous Determination of Food Colorants and Sweeteners" (2013) *Journal of Chemistry*, 127847, <http://dx.doi.org/10.1155/2013/127847>.
- [19] N. Pourreza, M. Ghomi "Simultaneous cloud point extraction and spectrophotometric determination of carmoisine and brilliant blue FCF in food samples" (2011) *Talanta* 84: 240-243.
- [20] K. Rovina, S. Siddiquee, S. M. Shaarani "Extraction, Analytical and Advanced Methods for Detection of Allura Red AC (E129) in Food and Beverages Products" (2016) *Frontiers in Microbiology* 7: 1-13.
- [21] M. Šuleková, A. Hudák, M. Smrčová "The Determination of Food Dyes in Vitamins by RP-HPLC" (2016) *Molecules* 21, 1368, 10.3390/molecules21101368.
- [22] L. F. Capitán-Vallvey, N. López-Ruiz, A. Martínez-Olmos, M. M. Erenas, A. J. Palma "Recent developments in computer vision-based analytical chemistry: A tutorial review" (2015) *Analytica Chimica Acta* 899: 23-56.
- [23] E. A. Lopez-Poveda, Universidad de Salamanca, Master Neurociencias. <http://audiolab.usal.es/Documentos/Docencia/Master%20Neurociencias.pdf> Consultada el 16/06/2017.
- [24] F. Soponar, A. Cătălin, C. Sârbu "Quantitative determination of some food dyes using digital processing of images obtained by thin-layer chromatography" (2008) *Journal of Chromatography A* 1188: 295-300.

- [25] B. G. Botelho, L. P. de Assis, M. M. Sena "Development and analytical validation of a simple multivariate calibration method using digital scanner images for sunset yellow determination in soft beverages" (2014) *Food Chemistry* 159: 175-180.
- [26] B. G. Botelho, K. C.F. Dantas, M. M. Sena "Determination of allura red dye in hard candies by using digital images obtained with a mobile phone and N-PLS" (2017) *Chemometrics and Intelligent Laboratory Systems* 167: 44-49
- [27] M.C.A. Marcelo, K.C. Mariotti, R.S. Ortiz, M.F. Ferrao, M.J. Anzanello "Scott test evaluation by multivariate image analysis in cocaine samples" (2016) *Microchemical Journal* 127: 87-93.
- [28] Juan-Ping Sun, Cai-Yun Hou, Jing Feng, Xu Wang "Determination of the protein content in rice by the digital chromatic method" (2008) *Journal of food quality* 31: 250-263.
- [29] M. Vidal, J. M. Amigo, R. Bro, M. Ostra, C. Ubide, J. Zuriarrain "Flatbed scanners as a source of imaging. Brightness assessment and additives determination in a nickel electroplating bath" (2011) *Analytica Chimica Acta* 694: 38-45.
- [30] M. Vidal, J. M. Amigo, R. Bro, F. Van der Berg, M. Ostra, C. Ubide "Image analysis for maintenance of coating quality in nickel electroplating baths – Real time control" (2011) *Analytica Chimica Acta* 706: 1-7.
- [31] M. Vidal, M. Ostra, N. Imaz, E. García-Lecina, C. Ubide "Feature descriptors from scanned images of chromium electrodeposits as predictor parameters of Surface roughness and crystallographic texture" (2015) *Chemometrics and Intelligent Laboratory Systems* 149: 90-98.
- [32] C. Mongay Fernández "Quimiometría" (2005) *Universitat de València*.
- [33] Grupo de Quimiometría y Cualimetría de Tarragona "QUIMIOMETRÍA, Una disciplina útil para el análisis químico" *Departamento de Química Analítica y Química Orgánica Universitat Rovira i Virgili*. <http://www.quimica.urv.es/quimio/general/quimio.pdf> Consultado el 17/06/2017.
- [34] James N. Miller, Jane C. Miller "Estadística y Quimiometría para Química Analítica" 4º Edición (2002), *Capítulo 8*.
- [35] K.H. Esbensen "Multivariate data analysis-in practice" (2006) 5º Edición, *Camo AB Oslo*.
- [36] A. Kaur, U. Gupta "The Review on Spectrophotometric Determination of Synthetic Food Dyes and Lakes" (2012) *Gazi University Journal of Science* 25(3). 579-588.
- [37] Daniel C. Harris "Análisis químico cuantitativo" 3º Edición. Editorial Reverte. Capítulos 18 y 20.
- [38] Mohammad-Hossein Sorouraddin, Mohammad-Yaser Khani, K. Amini, A. Naseri, D. Asgari, Mohammad-Reza Rashidi "Simultaneous Determination of 6-Mercaptopurine and its Oxidative

Metabolites in Synthetic Solutions and Human Plasma using Spectrophotometric Multivariate Calibration Methods" (2011) *BiolImpacts* 1(1): 53-63.

[39] R. Ahirwar, S. Tanwar, S. Parween, A. Kumar, P. Nahar "Image-based detection of oligonucleotides – a low cost alternative to spectrophotometric or fluorometric methods" (2014) *Analyst* 139: 2186-2192.

[40] D. Damasceno, T. Gomes Toledo, A. da Silva Soares, S. Botelho de Oliveira, A. Elcana de Oliveira "CompVis: a novel method for drinking water alkalinity and total hardness analyses" (2016) *Analytical Methods* 8: 7832-7836.

[41] M. Vidal. "New methods for the analytical control of a nickel electroplating bath. Application of chemometric techniques" (2010) *PHD thesis, UPV/EHU*: 136-138.

[42] A. Antonelli, M. Cocchi, P. Fava, G. Foca, G. C. Franchini, D. Manzini, A. Ulrici "Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm" (2004) *Analytica Chimica Acta* 515: 3-13.

[43] A. Ulrici, G. Foca, M. C. Ielo, L. A. Volpelli, D. Pietro Lo Fiego. "Automated identification and visualization of food defects using RGB imaging: Application to the detection of red skin defect of raw hams" (2012) *Innovative Food Science and Emerging Technologies* 16: 417-426.

[44] M. Ostra, C. Ubide, M. Vidal, J. Zuriarrain "Detection limit estimator for multivariate calibration by an extension of the IUPAC recommendations for univariate methods" (2008) *Analyst* 133: 532-539.

[45] L. Zhi, W. Hai-Long, X. Li-Xia, H. Yong, F. Fang, S. Xiao-Dong, W. Tong, X. Rong, Y. Ru-Qin "Chemometrics-enhanced liquid chromatography-full-scan-mass spectrometry for interference-free analysis of multi-class mycotoxins in complex cereal samples" (2017) *Chemometrics and Intelligent Laboratory Systems* 160: 125-138.