



**Universidad**  
Zaragoza

**Calibración multivariable en métodos con imágenes digitales**

**Multivariable calibration in digital image based  
methods**

TFG Tipo II

DEPARTAMENTO DE QUIMICA ANALITICA

TRABAJO FIN DE GRADO

**SERGIO CALLEJA CRUZ**

DIRECTOR: ÁNGEL LOPEZ MOLINERO

FACULTAD DE CIENCIAS

Convocatoria de Septiembre de 2018



# Abstract

---

The TFG presents a theoretical study on the variance of analytical information and its modeling through different regressions.

It has been applied to chromatic systems that are captured from digital images in different lighting conditions. And of which the intensity of the primary reflected colors R-G-B is measured.

It has been shown that the chromatic variance is significantly increased using different type of lighting. Thus pan-chromatic, and also monochromatic illuminations have been used. So that up to 12 chromatic variables can be obtained.

The use of multi-variable regressions has allowed modeling the response of the chromatic system formed by sulfite ions in the Boedecker reaction.

Through quality parameters of the regressions, the differences between regressions were shown: MLR, PCR and PLS. Likewise, the influence of the number of variables in the prediction capacity is evident.

Under the best conditions, it has been possible to propose a regression to overcome the effects of variance in white wine samples and predict total sulfite contents. The results have been validated against results of reference methods.

**Key words:** Digital Image Colorimetry, RGB colours, MLR, PCR, PLS, Sulphite, Boedecker, Variance

# Resumen

---

El TFG presenta un estudio teórico sobre la varianza de la información analítica y su modelización mediante diferentes regresiones.

Se ha aplicado a sistemas cromáticos que son captados de imágenes digitales en diferentes condiciones de iluminación. Y de los cuales se mide la intensidad de los colores reflejados primarios R-G-B.

Se ha demostrado que la varianza cromática se incrementa significativamente utilizando diferente tipo de iluminación. Así se han utilizado iluminaciones pan-cromáticas, y también monocromáticas. De modo que pueden obtenerse hasta un máximo de 12 variables cromáticas.

El uso de regresiones multi-variables ha permitido modelizar la respuesta del sistema cromático formado por los iones sulfitos en la reacción de Boedecker.

Mediante parámetros de calidad de las regresiones, se ha puesto de manifiesto, las diferencias entre regresiones: MLR, PCR y PLS. Así mismo, se evidencia la influencia del número de variables en la capacidad de predicción.

En las mejores condiciones, se ha podido proponer una regresión para superar los efectos de varianza en muestras de vino blanco y predecir los contenidos de sulfito total. Los resultados han sido validados frente a resultados de métodos de referencia.

**Palabras clave:** Colorimetría imagen digital, colores RGB, Sulfitos, Boedecker, RLM, RCP, PLS, Varianza

# Índice

---

## **CAPITULO I: INTRODUCCIÓN Y OBJETIVOS**

1- Calibración simple en métodos instrumentales de análisis.....	1
2- Calibración múltiple y multivariante.....	1
2.1- Regresión Lineal Múltiple.....	1
2.2- Regresión de Componentes Principales.....	2
2.3- Regresión de Mínimos Cuadrados Parciales.....	2
3- Colorimetría de Imágenes Digitales.....	3
3.1- Principios.....	3
3.2- Empleo de la colorimetría RGB en investigaciones precedentes.....	4
4- Uso de las técnicas en la literatura analítica.....	4
5- Los sulfitos.....	5
6- Reacción de Boedecker: método cromogénico.....	6
7- Objetivos del TFG.....	6

## **CAPITULO II: EXPERIMENTAL**

1- Instrumentación.....	7
2- Reactivos, Disoluciones .....	8
2.1- Preparación de disoluciones: .....	8
3. Métodos	
3.1- Experiencia de rango dinámico.....	9
3.2- Captura de imágenes.....	9
3.3 – Lectura de colores.....	10

## **CAPITULO III: RESULTADOS Y DISCUSIÓN**

1. Caracterización de la reacción de sulfitos.....	10
1.1. Tendencias en los colores básicos .....	10
1.2. varianza cromática con solo analito.....	12
1.3. Varianza teniendo en cuenta la influencia de la matriz muestral.....	13
2. Características de las Calibraciones.....	15
2.1. – Calibración MLR.....	15
2.2. – Calibraciones PCR, PLS.....	16
3. Aplicación de una calibración multivariable en caso de muestra real.....	18

## **CAPITULO IV: CONCLUSIONES Y BIBLIOGRAFÍA**

1- Conclusiones.....	20
2- Bibliografía.....	21



# Capítulo I. Introducción y objetivos

---

## 1 - Calibración simple en métodos instrumentales de análisis

Desde la metrología, calibrar es un proceso de comparación de la medida de un equipo con valores de referencia conocidos. Se distingue entre calibración directa e indirecta. Nuestro interés se centra en esta última. Mediante una calibración indirecta típica en Química Analítica, una variable señal instrumental se calibra y modela frente a una variable de concentración. Su aplicación permite predecir la concentración de un problema una vez conocida la medida instrumental.

El tipo de calibración y regresión lineal simple, RLS, es bien conocido<sup>1</sup>:

$$y = b_0 + b_1 x;$$

con y como variable dependiente o señal instrumental,  
con x variable independiente que suele ser la concentración.

b<sub>0</sub> y b<sub>1</sub> son los coeficientes de ajuste lineal

El modelo anterior cuando está operativo predice valores de concentración a partir de valores de señales. Modelo  $\hat{y} = b_0 + b_1 [x]$  → predicción por retroceso  $[x] = (\hat{y} - b_0) / b_1$

## 2 - Calibración múltiple y multivariante<sup>2</sup>

Se habla de dos tipos de calibraciones. En la primera se predice una única respuesta,  $\hat{y}$ , a partir de varias variables predictoras, xi. Este procedimiento calcula un modelo que se dice de regresión múltiple. La segunda calibración predice varias respuestas,  $\hat{y}_i$ , a partir de las variables predictoras, xi. El cálculo deduce un modelo que se dice de regresión multivariante.

En estos modelos la variable independiente “xi” es siempre predictora: la señal instrumental. Siendo siempre la respuesta, variable dependiente, ‘y’ única o bien ‘yi’ varias. Habitualmente es concentración de analito/s. Este convenio se mantiene igual en las etapas de calibración y de predicción. Se denomina ‘modelo de avance’.

La utilidad de este tipo de calibraciones reside en su capacidad para predecir la respuesta de un sistema cuando es función de distintas variables. Lo cual es muy habitual. Es decir, en problemas complejos cuya varianza necesita diferentes factores para poder ser explicada-modelizada.

### 2.1 - Regresión Lineal Múltiple –RLM-

Nuestro primer interés en el TFG se ha centrado en el modelo de regresión múltiple basado en una relación lineal. Se describe con la expresión de Regresión Lineal Múltiple, RLM:

$$\text{Modelo } \hat{y} = b_0 + b_1 [x_1] + b_2 [x_2] + \dots + b_i [x_i]$$

siendo y la variable dependiente: concentración analito,  
x<sub>1</sub>, x<sub>2</sub>, x<sub>i</sub> variables independientes: señales instrumentales.  
b<sub>0</sub>, b<sub>1</sub>, b<sub>2</sub>...b<sub>i</sub> coeficientes de ajuste lineal

El modelo se ha aplicado para predecir la concentración de analito en un procedimiento cromático que se mide en imágenes digitales. Estas son tomadas con cámara fotográfica compacta. El procedimiento se dice 'simple' y habitualmente se desarrolla con calibración RLS. Es decir, utilizando un solo componente de color, R-G-B, de la imagen digital. Sin embargo, se tratará de demostrar que la predicción en estos procedimientos cromáticos se mejora conforme aumenta el número de variables cromáticas. La respuesta analítica de estos métodos instrumentales basados en la medida DIC- Digital Image Colorimetry- se mejora aumentando la varianza cromática que se capta en el sistema. Es decir, la predicción analítica podría verse mejorada con la aplicación de métodos multivariante.

La evaluación de la calidad de los distintos modelos de regresión aplicados se ha llevado a cabo mediante:

R<sup>2</sup>: Coeficiente de correlación; R<sup>2</sup>=varianza que explica el modelo de regresión / varianza total de los datos de calibración.

RMSE: Root Mean Square Error de los residuales en **Calibración**. Este valor se estima sobre los valores de la experiencia de calibración directa, es decir sobre variables Y-X. También inversa, con las variables a ejes cambiados X-Y. En el caso de la inversa los valores RMSE pueden ser estimados directamente en concentración.

RMSE en Validación: evaluar el RMSE del modelo en 'Cross Validation- Leave One Out': CV-LOO. En estos casos también se calculan con calibración inversa de variables.

RMSE en test: evaluar el RMSE del modelo eligiendo un grupo de resultados para verificar como test.

## **2.2 - Regresión en Componentes Principales –RCP-**

En este tipo de regresión múltiple se incide en los problemas de co-linealidad y por tanto de sobre-información que puede presentarse entre las variables elegidas en el modelo RLM. El RCP se basa en deducir nuevas variables predictivas, o componentes principales, a partir de la combinación de las variables iniciales. De modo que retengan solo la varianza de los datos correlacionada con la respuesta y eliminar la varianza no correlacionada. Es decir, se selecciona la información evitando el sobre-exceso.

El método de cálculo de los Componentes Principales utiliza procedimientos de Algebra Matricial, que se ven en la asignatura de Matemáticas y que ahora aplicamos en el contexto de la calibración analítica. En el TFG se han utilizado programas de cálculo que simplifican el procedimiento.

Al igual que en el caso de la RLM, se evaluará la calidad del ajuste del modelo mediante los parámetros: R<sup>2</sup>, RMSE : Calibración- validación en CV-LOO y test.

## **2.3 - Regresión de Mínimos Cuadrados Parciales –RMCP-**

Comparte fundamento con la RCP. Es decir, se basa en deducir unas nuevas variables<sup>3,4</sup>, como combinación de las variables iniciales y que se dicen nuevas variables latentes. Este procedimiento se caracteriza por retener la máxima varianza de los datos, además incluir la varianza de la respuesta y seleccionar las variables latentes que presentan la máxima covarianza con la respuesta.

El procedimiento de cálculo es similar al expuesto en la RCP. Es decir, con álgebra matricial que se lleva a cabo con programas de ordenador.

La calidad de la regresión de los modelos se evalúa del mismo modo que en MLR y RCP. Es decir, con los parámetros:  $R^2$ , RMSE: Calibración, Validación CV-LOO y test.

### 3 – Colorimetría de Imágenes Digitales-DIC-

#### 3.1 - Principios

La técnica DIC<sup>5</sup> se ha desarrollado en los últimos años gracias al continuo avance de las cámaras digitales<sup>6</sup>, también presentes actualmente en todo tipo de dispositivos inteligentes como tablets, Smartphone y web-cams.

Otorga lectura de datos cromáticos de manera inmediata, por lo que se puede encuadrar dentro de las técnicas rápidas de análisis.

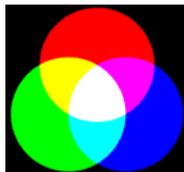
Los detectores que utilizan este tipo de cámaras digitales son en gran parte tipo CMOS (Complementary Metal Oxide Semiconductor), o CCD (Charge Coupled Device), los cuales otorgan medidas de los colores Rojo, Verde y Azul (RGB) de la imagen.

El principio en el que se basa es:

- Obtención de la imagen digital
- Lectura de la intensidad de la luz reflejada en los colores básicos RGB
- Correlación de los datos con la concentración del analito de interés.

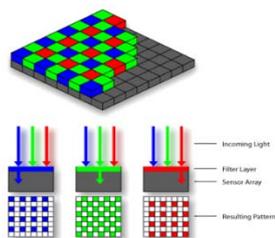
En la colorimetría tri-cromática RGB, los colores se definen y codifican en base a la suma aditiva de los tres filtros RGB. Como utilizamos sensores de 8 bits, las respuestas se encuentran escaladas en el rango 0-255 unidades de color, o código de color ( $2^8$ ).

Por tanto, tendremos la escala de colores desde negro (0, 0, 0) hasta blanco (255, 255, 255).



*Figura 1: Imagen de una composición aditiva de colores básicos*

La obtención de los colores básicos en la lectura digital se logra superponiendo un filtro de corte a los detectores. El filtro típico más utilizado es el mosaico de Bayer que se ve en la Figura 2. Se distribuye de forma sistemática filtros verdes, rojos y azules en cantidades 50-25-25 % respectivamente.



*Figura 2: Mosaico de Bayer utilizado en las cámaras digitales y la distribución de filtros R-G-B.*

### 3.2 - Empleo de la colorimetría RGB en investigaciones precedentes

Nuestro grupo de investigación ha sido pionero en la implantación de esta metodología. Algunos ejemplos de diferentes metodologías analíticas realizadas para determinaciones han sido: Ti (IV) como aditivo en plásticos<sup>7</sup>, demanda química de oxígeno en aguas<sup>8</sup>, peróxido de hidrógeno (H<sub>2</sub>O<sub>2</sub>) en desinfectantes y productos químicos<sup>9</sup>, dureza en agua<sup>10</sup>, Ca (II) en productos farmacéuticos<sup>11</sup> y aminos biógenas<sup>12</sup>, entre otros.

### 4 - Uso de las técnicas en la literatura analítica

Se ha realizado un estudio bibliográfico del uso de las técnicas de calibración aplicadas en el TFG para conocer su impacto actual y su evolución a lo largo de los últimos años.

Se ha llevado a cabo mediante búsquedas en la web de 'Elsevier- Science Direct', proporcionado por la Biblioteca Universitaria, con las palabras clave que aparecen en las Figuras y seleccionando todo tipo de publicaciones.

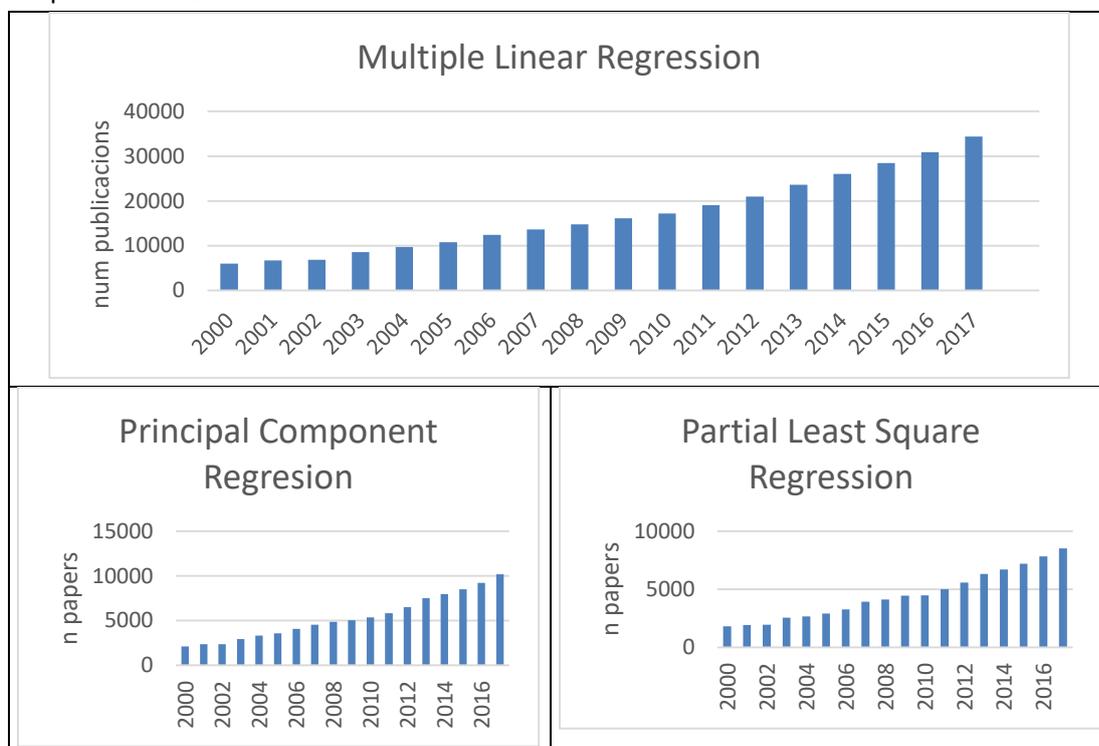


Figura 3: Número de resultados pertenecientes a cada año por búsqueda

Los resultados se han reducido al periodo 2000-2017. Las figuras muestran cómo se ha producido un aumento exponencial en el uso de este tipo de calibraciones. La utilización de la RLM es aproximadamente 4 veces mayor respecto a RCP y RMCP, cuya utilización es similar.

Incluyendo en los parámetros de búsqueda una 'intersección' con los 'métodos de imágenes digitales', se ha encontrado que aproximadamente se reducen al 10% los resultados para cada año. Que, no obstante, sigue siendo un número alto de publicaciones y da idea del creciente interés por su aplicación.

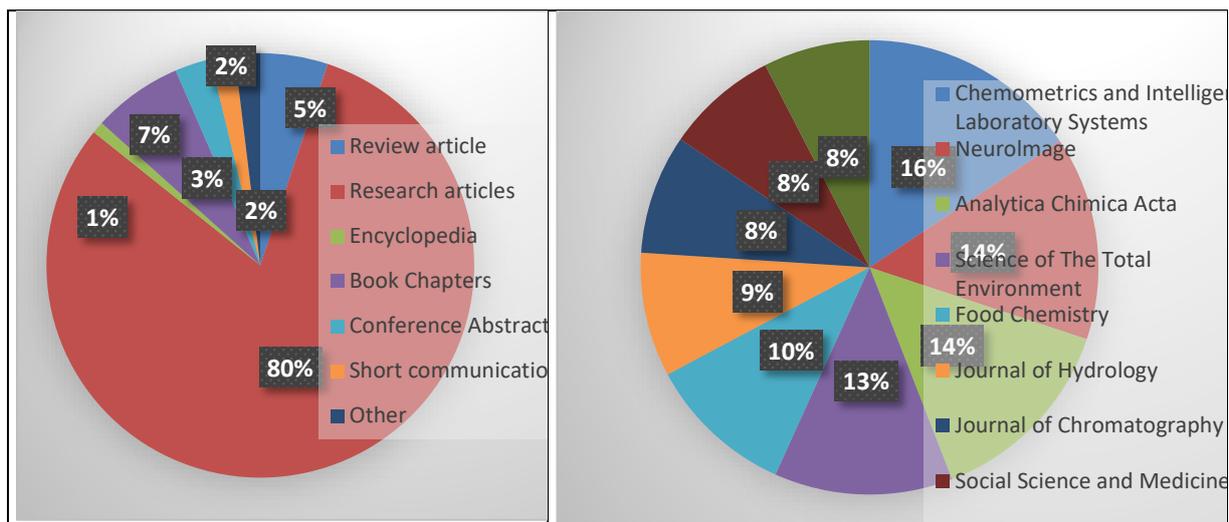


Figura 4: Distribución de la búsqueda resultados

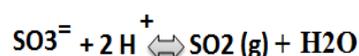
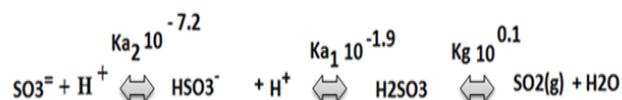
Se encuentra, como es de esperar, que la mayor parte de los resultados se relacionan con artículos, pero que éstos se encuentran divididos equitativamente en campos de estudio y aplicaciones muy diversas entre sí.

## 5 - Los sulfitos

El TFG trata de demostrar que la captación de una mayor varianza cromática en un método DIC permite modelizar un sistema cromogénico y predecir su resultado analítico con menor error. El estudio cromático se ha desarrollado sobre la reacción de Boedecker que se produce por los iones sulfitos con nitroprusiato y los iones Zn(II). El sistema cromogénico ha sido ya aplicado en un TFG anterior<sup>13</sup> (Métodos analíticos con imágenes digitales, por Lorena Gómez) pero ahora en el presente TFG se ha modelizado su comportamiento mediante calibraciones multivariantes.

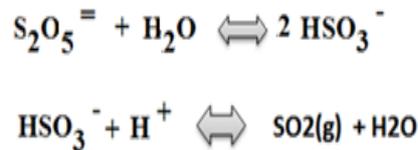
Las características químicas y analíticas de los sulfitos están bien descritas en la bibliografía y los trabajos del grupo. Resumimos los aspectos que nos son de mayor interés para este TFG. Se destaca el carácter reductor del semi-sistema  $\text{SO}_4^{2-}/\text{SO}_3^{2-}$ , con  $E^\circ = -0.936 \text{ V}$ . De modo que la oxidación de sulfito a sulfato se da con oxígeno atmosférico, aunque la cinética es lenta. Los iones metálicos como Co(II), Ce(IV) o Fe(III) actúan como catalizadores favoreciendo la oxidación. Sin embargo, alcoholes y glicerina son inhibidores; es por ello que adicionamos una cantidad de EDTA suficiente que actúe como estabilizador en las disoluciones de sulfito al acomplejar especies metálicas.

Es importante tener en cuenta en este ión, las diferentes especies de S(IV) y sus equilibrios en función del pH. Los equilibrios entre especies se describen en el esquema siguiente:



Los sulfitos son aditivos muy empleados en la conservación de alimentos: antioxidante y antibacteriano. En especial en la conservación de vinos. Siendo muy

significativo en el caso de los vinos blancos. Es muy habitual su adición en forma de meta bisulfito. En el esquema siguiente se describen los equilibrios de la especie, y en función del pH:



Los sulfitos empleados en alimentación en Europa se clasifican en la denominación E 220-228. La UE, en el Reglamento (CE) Nº 1493/1999<sup>14</sup> y el Reglamento (CE) Nº 1622/2000<sup>15</sup> estableció una concentración límite de sulfuro total en 160 mg/l para vinos tintos y 210 mg/l para vinos blancos. La Organización Internacional de la Viña y el Vino (OIV), recomienda un máximo de 150 mg/l para vinos tintos y 200 mg/l para vinos blancos y rosados que contengan 4 g/l de azúcares reductores, pudiendo aumentar hasta 300 y 400 mg/l respectivamente si estos azúcares también aumentan.

## 6 - Reacción de Boedecker

En el esquema se expresa el mecanismo de la reacción iónica cromogénica: los sulfitos reaccionan con nitroprusiato pentaciano ferrato (III), en presencia de  $\text{Zn}^{2+}$  para formar un precipitado de sulfito-nitro prusiato de zinc, de color rojo<sup>16</sup>. La adición de Piridina -u otras aminas- aumenta la sensibilidad como consecuencia de la producción de pares iónicos menos solubles y más intensamente coloreados.

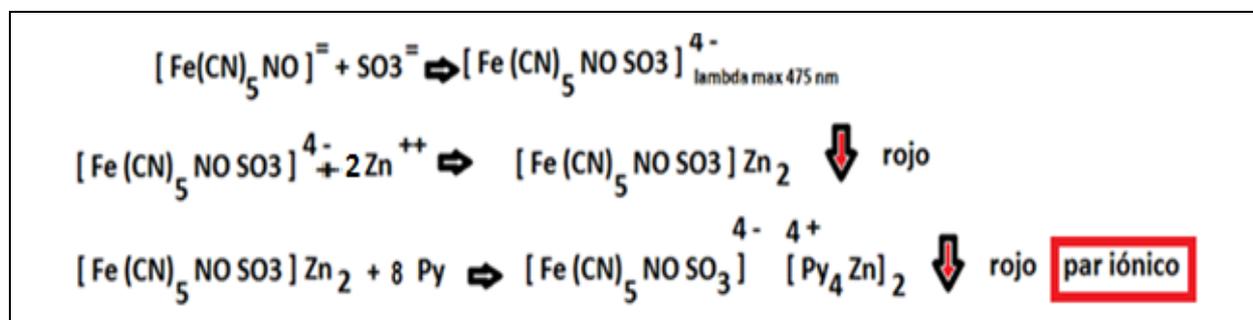


Figura 5: Reacción de Boedecker

En el caso del TFG<sup>11</sup> de Lorena Gomez, 'Métodos analíticos con imágenes digitales', la reacción se modificó ligeramente, ya que se propuso Hexametilentetramina como amina. Los resultados producen una similar intensidad colorimétrica y se evita la Piridina que tiene una toxicidad aguda y con riesgos de irritación cutánea y ocular. Los estudios posteriores se realizarán con la inclusión de esta modificación,

## 7 – Objetivos del Trabajo de Fin de Grado

El objetivo general trata de demostrar que la captación de una mayor varianza cromática en métodos DIC podría conducir a menores errores en la predicción analítica.

En el contexto en el que se aplica del TFG se destacan los siguientes puntos:

- Es tipo 2: Por tanto 'Es un trabajo de carácter teórico para la interpretación de resultados o comprensión de fenomenologías relacionadas con la Química'.
- Se basa en un método cromogénico desarrollado en un TFG previo: método de Boedecker para la determinación de sulfitos.
- El uso de distintas iluminaciones: w, r, g, b, permite captar una mayor varianza en el sistema cromogénico.
- Las medidas de varianza de las reacciones cromogénicas deben de realizarse en condiciones controladas entre sí. Por tanto, el estudio requiere de planificación, diseño de preparaciones y medidas experimentales. Y sobre las mismas aplicar procedimientos de cálculo e interpretación de fenómenos.
- Se aplican métodos diferentes de regresión: RLS, RLM, PCR y PLS para la determinación de sulfitos mediante datos R, G, B, medidos en imágenes digitales.
- Los procedimientos de cálculo de resultados y obtención de medidas se han llevado a cabo con programas de ordenador desarrollados en el grupo de investigación.

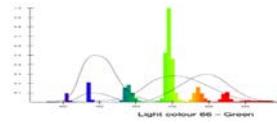
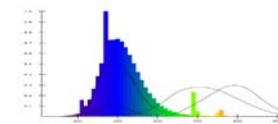
## Capítulo 2. Experimental

### 1 - Instrumentación

#### Materiales de laboratorio

1. Micropipetas Pipet-Lite 'RAININ' 5-50  $\mu\text{L}$ , 20-200  $\mu\text{L}$ , 100  $\mu\text{L}$  – 1 mL y 0,5-5 mL.
2. Matraces aforados de 50, 100 y 250mL.

Cámara digital CANON: Compacta: Power Shot SX230 HS	
Resolución	12,1 megapíxeles( 12.1x 10 <sup>6</sup> pixeles )
Sensor `CMOS´: 1 /2.3 pulgadas = 6.17x4.55 mm Área pixel= (6.17x4.55) x10 <sup>6</sup> /12.1x10 <sup>6</sup> = 2.32 $\mu\text{m}^2$ Procesador: DIGIC 4 iSAPs	
Cámara réflex NIKON: D100	
Resolución	6 megapíxeles
<ul style="list-style-type: none"> <li>◦ Conversor analógico-digital de 12 bits.</li> <li>◦ ISO 200-1600. Velocidad 30-1/4000 sec.</li> <li>◦ Lente con focal equivalente a 36-105mm.</li> <li>◦ Tarjeta CompactFlash de 256 MB.</li> <li>◦ Formatos: JPEG, TIF y RAW.</li> </ul>	
"Lighting-Box":	

		
Soportes de reacción:		
Placas de ceramica para ensayos a la gota		
Iluminantes: Lámpara LED , EGLO		
Características LED: 7.5 watt, 470 lumen		
<b>Emision tipo en LED: Red emission</b> 	<b>Green emission</b> 	<b>Blue emission</b> 

## 2 - Reactivos, Disoluciones.

### 2.1 - Preparación de disoluciones:

- Disolución de nitroprusiato sódico: SNP- Panreac: Nitro ferrocianuro sódico, dihidratado  $\text{Na}_2[\text{Fe}(\text{CN})_5\text{NO}] \cdot 2 \text{H}_2\text{O}$ , Cód. 131705, Peso molecular 297.95 g/mol. Se disuelven 0.25 g con agua desionizada y se llevan a volumen de 10ml.
- Disolución de nitrato de zinc: Probus: Se disuelven 0.6 g de  $\text{Zn}(\text{NO}_3)_2$  con agua desionizada y se diluye a 25ml.
- Disolución patrón de sulfitos: Panreac. Cód. 131717. Peso molecular 126,04 g/mol. 1000 mg/l: se pesan 0.08 g de  $\text{Na}_2\text{SO}_3$  y se disuelve y diluye a 50 ml en matraz aforado con agua desionizada. Se estabiliza con adición de 0.18g de  $\text{EDTA} \cdot \text{Na}_2 \cdot 2\text{H}_2\text{O}$ .
- Disolución combinada de Hexametilentetramina(Hx)+NaOH: Se pesan 5.3g de Hexamina y 0.03g de NaOH se disuelven con agua desionizada en un volumen de 20ml. Hex: Panreac, 140,19g/mol. NaOH: Merck, 40,00 g/mol.
- Suspensión de micro-partículas de nitroprusiato y zinc: snp-zn:'nps': Se pesa 0.6g de  $\text{Zn}(\text{NO}_3)_2$  y se disuelven en 25ml de agua desionizada. Se enfrían a temperatura entre 5 y 10°C con baño de agua y se mantiene en agitación vigorosa.

Se añade 1 ml de TX-100

Por otro lado, se pesan 0.25g de  $\text{Na}_2 [\text{Fe}(\text{CN})_5\text{NO}] \cdot 2 \text{H}_2\text{O}$  y se disuelven en 10ml de agua desionizada. Esta disolución se vierte gota a gota sobre la disolución anterior de nitrato de zinc. Una

vez terminada la adición se acidifica con HCl 0.01M hasta un pH de 3. Y se mantiene en agitación constante.

- Tensoactivo neutro Tritón X 100: Polyethylene glycoltert-octylphenyl ether: Sigma: 60 mg/l: se vierten 4 ml del tensoactivo en 10ml de agua desionizada.

### 3 Métodos:

- **3.1 Experiencia cromogénica tipo: rango dinámico**

Se lleva a cabo un diseño de experiencias para cubrir el rango de 0-150 mg/l de sulfitos, con una concentración constante de nitroprusiato pre-reaccionado con iones Zn, a pH alcalino 8-10, y con Hex amina a nivel de 0.63 mol/l. Así mismo se determinó que la adición de tensoactivo Triton X-100 mejoraba la reproducibilidad de las señales. El soporte de reacción son pocillos de una placa cerámica de ensayos a la gota. Y se mantiene el volumen final a 1,5 mL. Otras condiciones se han aplicado pero los valores anteriores fijan condiciones de referencia.

En la tabla siguiente se ofrece la relación de concentraciones en un ensayo tipo. Los experimentos se preparan siempre tomando volúmenes con micro-pipetas de las disoluciones patrones de referencia.

<b>disol=&gt;conc</b>	<b>SO<sub>3</sub>=, mg/L</b>	SNP,mol/l	Zn, mol/L	%, TX100	NaOH, mol/L	Hx, mol/l	EDTA, real
bl	0,00	0,0071	0,0168	0,3333	0,0120	0,6322	0,0000
2	6,52	0,0071	0,0168	0,3333	0,0120	0,6322	0,0001
3	16,30	0,0071	0,0168	0,3333	0,0120	0,6322	0,0002
4	22,81	0,0071	0,0168	0,3333	0,0120	0,6322	0,0002
5	29,33	0,0071	0,0168	0,3333	0,0120	0,6322	0,0003
6	39,11	0,0071	0,0168	0,3333	0,0120	0,6322	0,0004
7	45,63	0,0071	0,0168	0,3333	0,0120	0,6322	0,0005
8	52,15	0,0071	0,0168	0,3333	0,0120	0,6322	0,0005
9	58,67	0,0071	0,0168	0,3333	0,0120	0,6322	0,0006
10	68,44	0,0071	0,0168	0,3333	0,0120	0,6322	0,0007
11	74,96	0,0071	0,0168	0,3333	0,0120	0,6322	0,0008
12	81,48	0,0071	0,0168	0,3333	0,0120	0,6322	0,0008
13	88,00	0,0071	0,0168	0,3333	0,0120	0,6322	0,0009
14	97,78	0,0071	0,0168	0,3333	0,0120	0,6322	0,0010
15	104,30	0,0071	0,0168	0,3333	0,0120	0,6322	0,0011
16	110,81	0,0071	0,0168	0,3333	0,0120	0,6322	0,0011
17	120,59	0,0071	0,0168	0,3333	0,0120	0,6322	0,0013
18	127,11	0,0071	0,0168	0,3333	0,0120	0,6322	0,0013
19	136,89	0,0071	0,0168	0,3333	0,0120	0,6322	0,0014
20	143,41	0,0071	0,0168	0,3333	0,0120	0,6322	0,0015
21	149,93	0,0071	0,0168	0,3333	0,0120	0,6322	0,0016

Tabla 1: Diseño Experimentos para calibración

### 3.2 - Captura de imágenes

Se sitúa la placa de ensayo en el centro de la 'lightning-box', sobre fondo blanco, de forma que quede encuadrada para la correcta toma de la fotografía. Se utiliza cámara compacta digital que está fijada mediante un 'soporte-regulable', para conseguir capturas reproducibles. La iluminación utilizada es indirecta y artificial, mediante dos LED situadas en contraposición diagonal. La iluminación puede ser

cambiada a diversos colores; así principalmente: w blanca, b azul, g verde y r roja. Las fotos se realizan en modo automático y en ausencia de flash. Se realiza un pre-enfoque antes de tomar cada imagen para sistematizar los balances de colores automáticos.

### 3.3- Lectura de colores

Se obtienen imágenes en formato .jpg en cámaras Nikon y Canon y además en formato .raw en la cámara Nikon. Se utilizan las imágenes .jpg para el tratamiento de los datos. A través de un programa desarrollado por nosotros para la plataforma Matlab, se mide la intensidad de color RGB a lo largo de las distintas regiones de interés (ROI). Se obtiene el valor promedio y de la desviación estándar de 150 píxeles de las mismas. Las zonas ROI son seleccionadas en cada pocillo.

El tratamiento matemático de los datos de intensidad se realiza mediante programas desarrollados por nosotros para cada una de las calibraciones múltiples: MLR, PCR, PLS. En el caso de la MLR también se realizó con Microsoft Excel.

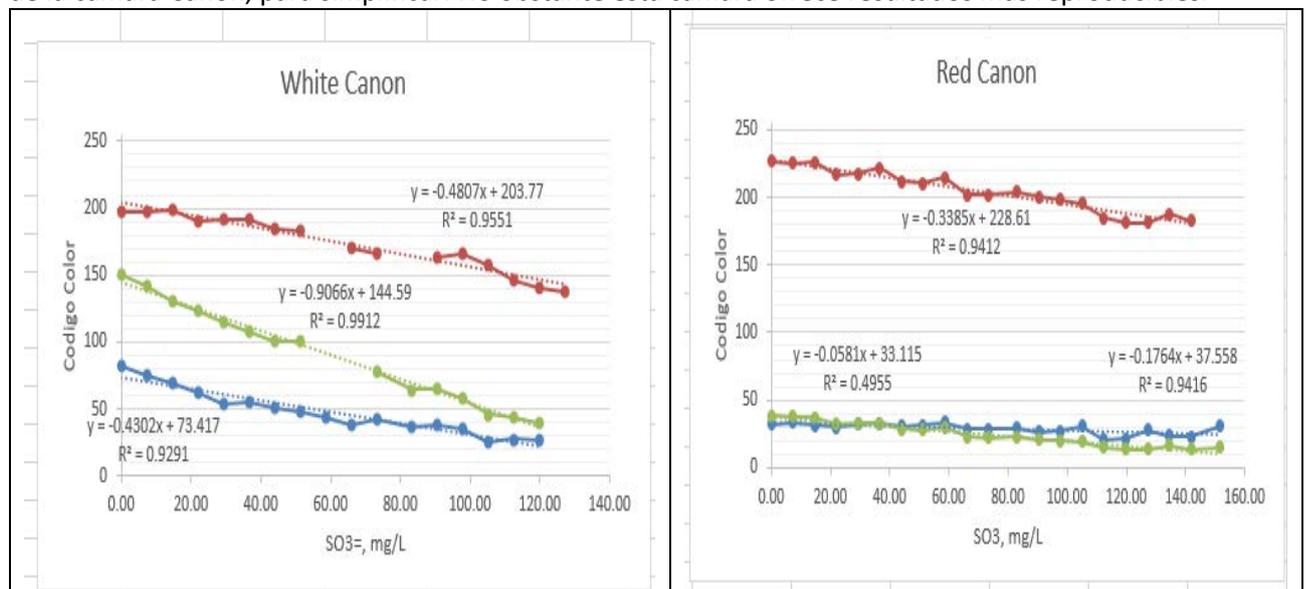
## Capítulo 3. Resultados y Discusión

### 1 – Caracterización de la reacción de sulfitos

#### 1.1 – Tendencias en los colores básicos con solo analito:

Se prepara el experimento básico de rango dinámico, según el diseño de la Tabla1. Se captura imagen en condiciones de iluminación: blanca-w, roja-r-, verde-g, y azul-b. Y se utilizan las cámaras: Canon y Nikon en formatos jpg y raw.

En la tabla se ofrecen las tendencias en la intensidad de color con la concentración de analito, solo los de la cámara Canon, para simplificar. No obstante esta cámara ofrece resultados más reproducibles.



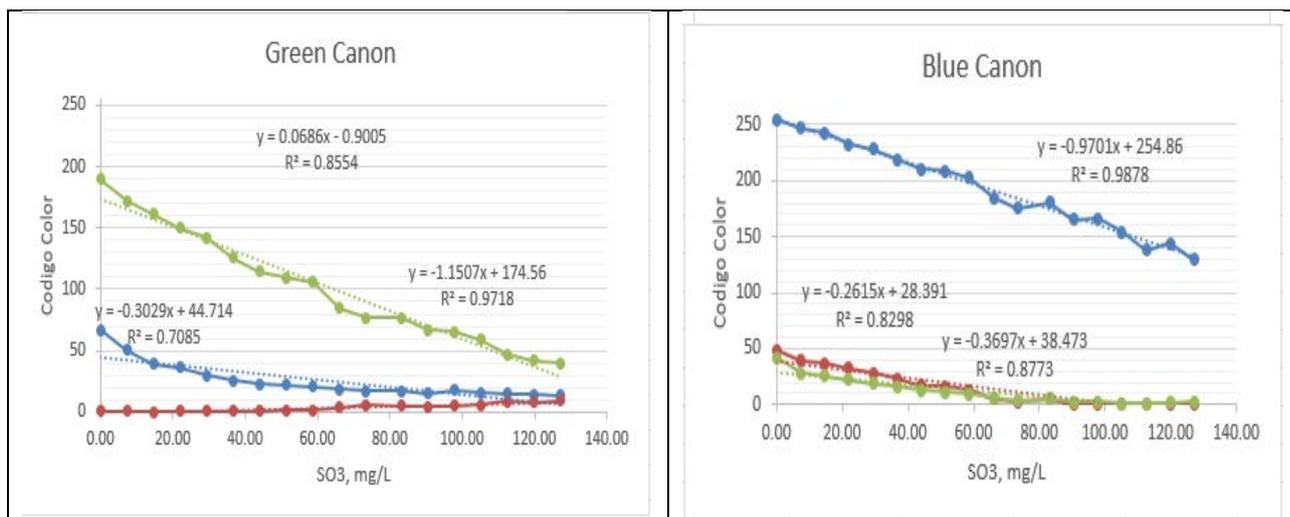


Figura 6: Regresión del rango dinámico con diferentes colores de iluminación medida sobre colores RGB.

Se muestran tendencias muy importantes entre las diferentes iluminaciones. Así en iluminación 'Pancromática-blanca' se aprecian cambios en el color verde y azul. Mientras que en iluminación monocromática se ven favorecidos esos mismos colores pero reflejados. Es importante ver como se inhiben las regresiones en los colores que no coinciden con la iluminación.

En la siguiente tabla se ofrecen los resultados de la calibración lineal RLS sobre colores individuales. Solo se ofrecen las calibraciones con valores  $r > 0.9$ .

Como parámetros de referencia se señalan los coeficientes de regresión, valores  $R^2$  y 'RMSE' de los residuales.

### Regresión RLS

	b1	b0	
desvest	-2.1625	170.3863	
r <sup>2</sup>	0.1546	7.8192	error típico
r	0.9288	10.5017	
Minimo	7.3333	Maximo	127.111

	b1	b0	
desvest	-2.1719	442.3781	
r <sup>2</sup>	0.1210	20.6389	error típico
r	0.9498	11.3227	
Minimo	0.0000	Maximo	151.555

	b1	b0	
desvest	-1.0376	161.9625	
r <sup>2</sup>	0.0372	3.8531	error típico
r	0.9824	5.8698	
Minimo	0.0000	Maximo	127.111

	b1	b0	
desvest	-2.7805	639.8132	
r <sup>2</sup>	0.1638	33.6134	error típico
r	0.9412	11.0771	
minimo	0.0000	maximo	141.7778

	b1	b0	
desvest	-5.1387	198.4656	
r <sup>2</sup>	0.2597	6.7994	error típico
r	0.9561	9.5772	
minimo	0.0000	maximo	141.7778

### Iluminación roja

iluminación blanca

Iluminación verde

	b1	b0	
	-0.84452994	149.211432	
desvest	0.03598678	4.01230386	
r <sup>2</sup>	0.9717682	6.96095699	error tipico
r	0.98578304		
minimo	0.00	maximo	127.11

Iluminación azul

	b1	b0	
	-1.01824277	260.286018	
desvest	0.02831202	5.58293447	
r <sup>2</sup>	0.98778144	4.57941119	error tipico
r	0.99387195		
minimo	0.00	maximo	127.11

Figura 7: Regresión RLS

En la siguiente tabla se recogen los valores de calidad de las regresiones con iluminación 'w' y en colores R-G-B, pero utilizando calibración inversa (intercambio de variables). Así mismo los valores RMSE de validación según método CV-LOO.

Iluminación w:

		Regresion Inversa	LOO
RMSE, mg/L	Blue	10,50	12,46
	Red	11,32	12,46
	Green	5,87	6,75
R2	Blue	0,93	
	Red	0,95	
	Green	0,98	

Tabla 2: Resultados RMSE por CV-LOO

## 1.2 Varianza cromática.

La **varianza cromática** representa el espacio o amplitud de color de un experimento cromogénico. Es una medida del potencial analítico del sistema.

Se obtiene de la matriz de varianza-covarianza de la intensidad de los colores del experimento

Los valores de varianza cromática en cada una de las iluminaciones y para el experimento básico se aportan en el diagrama siguiente.

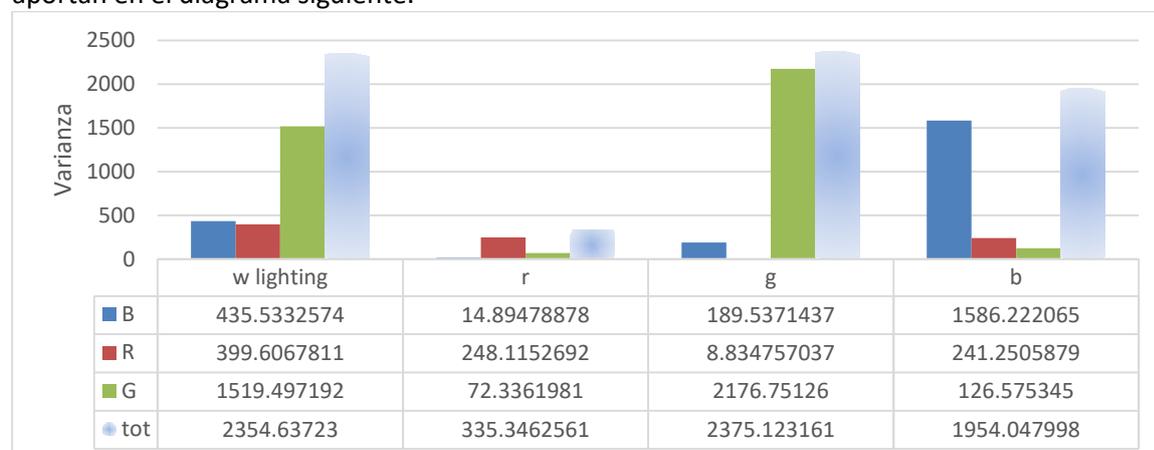


Figura 8: Varianza estudio calibración

Se pone de manifiesto cómo la iluminación w y la g producen la máxima varianza cromática. Es decir, son estas condiciones las pueden dar lugar a un mayor cambio en las variables y por tanto a una mayor sensibilidad analítica.

En el caso de la iluminación g, la varianza se proyecta sobre el color reflejado G, mientras que prácticamente se inhibe el color reflejado R.

Es importante destacar como la iluminación r reduce drásticamente la varianza cromática del sistema en un factor de 1/7 sobre los valores de varianza máxima

Por otra parte, los resultados evidencian que una manera de incrementar la varianza cromática, es decir ampliar la información de un sistema, se consigue incluyendo las intensidades cromáticas en las diferentes condiciones de iluminación.

En el diagrama siguiente se presentan los resultados de varianza según la fuente de iluminación y los colores reflejados.

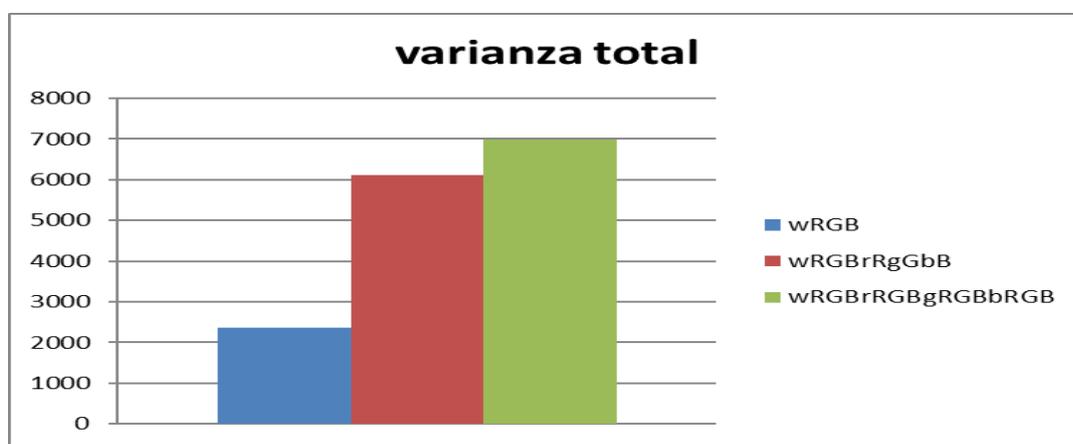


Figura 9: Estudio de la varianza total

Codificación:

Iluminación w colores reflejados RGB: 2354

Ilumina w -RGB ilumina r -R ilumina g -G ilum b - B: 6110

Ilumina w-RGB ilumina r-RGB ilumina g- RGB ilumina b- RGB: 7000

Esta manera de obtener información demuestra que: tomando los datos RGB de iluminación blanca y los datos de cada color RGB según la iluminación de ese mismo color, conseguimos casi triplicar la varianza aumentando las variables de 3 a 6.

Sin embargo, tomar todas las variables (12) aumenta la varianza, pero no de manera significativa. Veremos posteriormente que los resultados de calibración son algo peores que en el caso de 6 variables, debido a que por un exceso de información estamos “confundiendo” al modelo de regresión.

### **1.3 - Varianza teniendo en cuenta la influencia de la matriz muestral**

Se realiza un experimento para determinar la influencia de la matriz de una muestra sobre el desarrollo de color y su varianza. En este estudio se ha elegido un vino blanco como influencia de matriz.

En el experimento realizado se adiciona un rango de analito patrón a diferentes cantidades de muestra.

Es un experimento tipo de adición estándar generalizado.

El diseño de experimentos se expone en el Anexo.

Los resultados más significativos se presentan en representaciones 3D:  
 Con X = vol de vino,  $\text{SO}_3^{2-}$  en la muestra,  
 Y= conc  $\text{SO}_3^{2-}$  añadido analito patrón y  
 Z = intensidad color reflejado: R-G-B.

Solo se ofrecen resultados con iluminación w.

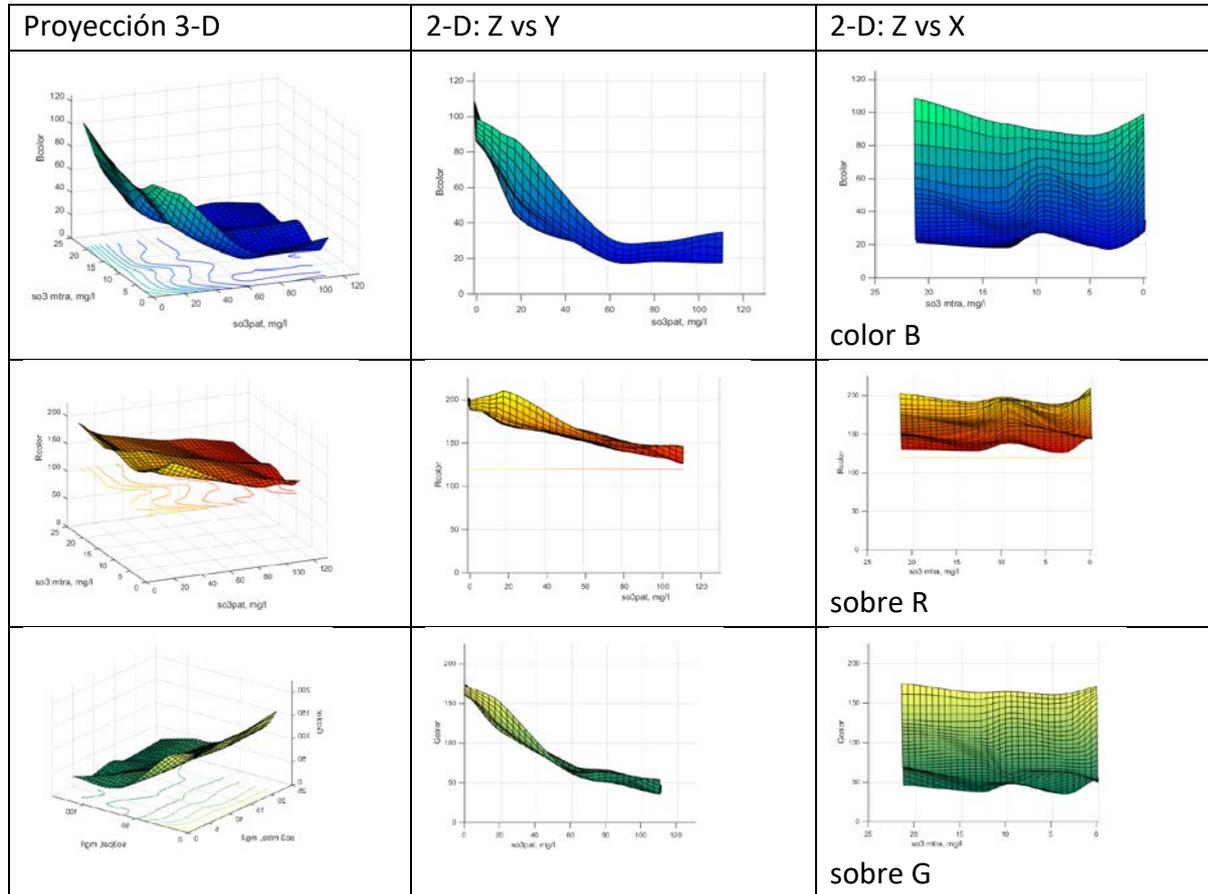


Figura 10: Representación 3D de la varianza muestral

Los resultados evidencian una influencia de la matriz sobre la varianza cromática. En las figuras, como tendencia resumida, se demuestra que la matriz disminuye la varianza del sistema, pero prácticamente no es influyente el nivel de muestra que se toma. La grafica siguiente refleja como la disminución de la varianza es prácticamente constante en el rango de muestra 30  $\mu\text{l}$ -120  $\mu\text{l}$ . En el diseño se pasan los valores de volumen a concentración.

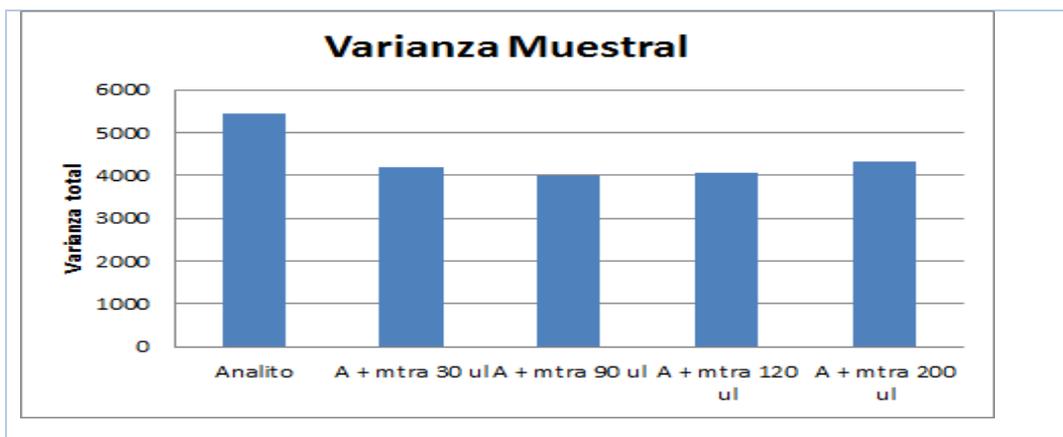


Figura 11: Influencia de la muestra en la varianza del analito

Los estudios realizados han presentado cómo la varianza cromática es fuente de información. Cómo se logra aumentarla, y también que la matriz es fuente de información, para el caso tratado.

## 2 - Características de las Calibraciones

Se realiza un estudio comparativo entre la calidad predictiva de diferentes calibraciones- solo se muestran las más relevantes- y teniendo en cuenta las diferentes informaciones cromáticas.

- 2.1 Calibración MLR:

	Error ,RMS estimacion: mg/l			
	Calibracion		Validacion	Validacion
	RMSE-C	R <sup>2</sup>	CV- LOO	14-4 test:20%
wBRG	2,626	0,996	2,833	2,543
wBRGrRbBgG	2,583	0,997	3,103	1,993
wBRGrBRGgBRGgBRG	2,789	0,999	8,628	1,455

Figura 12: Resultados para MLR

	PCR				
	Calibration		Validation		Test
wRGB	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
1PC	6,9335	0,9685	4,2145	0,9886	10,0359
2PC	6,2609	0,9743	3,5202	0,9918	9,5457
wRGBrRgGbB	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
1PC	4,3206	0,9878	4,278	0,988	7,5489
2PC	3,7493	0,9908	3,731	0,991	5,9622
wRGBrRGBgRGBbRGB	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
2PC	4,7629	0,9851	4,88396	0,98364	8,1321
3PC	4,22	0,9883	4,18426111	0,98848889	6,7072

Figura 13: Resultados para PCR

	PLS				
	Calibración		Validación		Test
<b>wRGB</b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>RMSE</b>
1 PLS	6,8986	0,9688	<b>4,1364</b>	<b>0,9883</b>	10,0208
2 PLS	6,1679	0,9751	<b>3,0892</b>	<b>0,9937</b>	9,4166
3 PLS	5,5839	0,9796			8,9224
<b>wRGBrRgGbB</b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>RMSE</b>
1 PLS	4,3085	0,9878	4,33544	0,98712	7,5188
2 PLS	3,4848	0,992	3,51828	0,99152	5,4795
3 PLS	2,1968	0,9968	2,15829444	0,99682778	2,8849
<b>wRGBrRGBgRGBbRGB</b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>RMSE</b>
2 PLS	4,5775	0,9863	4,67596	0,9842	7,7
3 PLS	3,6491	0,9913	3,6202	0,99102	5,761
4 PLS	2,9595	0,9943	2,90925	0,99428333	4,7536

Figura 14: Resultados para PLS

Estas tablas resumen los resultados obtenidos en los ejercicios de Calibración, Validación Cruzada y Validación Test para los diferentes conjuntos de datos pertenecientes a 3, 6 y 12 variables.

Conviene señalar que los mejores resultados son los que producen:

- mayor valor de  $R^2$
- menores valores de RMSE.

Con los datos obtenidos, se puede apreciar que esto ocurre en MLR con solo 6 variables: wRGB, rR, gG, bB. Al introducir más variables, es decir, 12 variables con todo el conjunto de valores de intensidad, se podría pensar que se obtendrán mejores resultados, pero esto no es así debido a que se genera un exceso de información que hace que los modelos matemáticos den peores resultados. Es decir, los métodos de calibración multivariable producen mayores valores de  $R^2$  y por tanto de justificación de la varianza cromática. Y además se reduce el error de predicción del modelo.

## • 2.2 Calibración PCR y PLS

Los principales resultados con calibraciones PCR y PLS se muestran en las gráficas siguientes: A fin de simplificar la exposición se muestran solo los resultados obtenidos en Calibración. Los resultados obtenidos en Validación y Test, son coincidentes y redundantes con los anteriores.

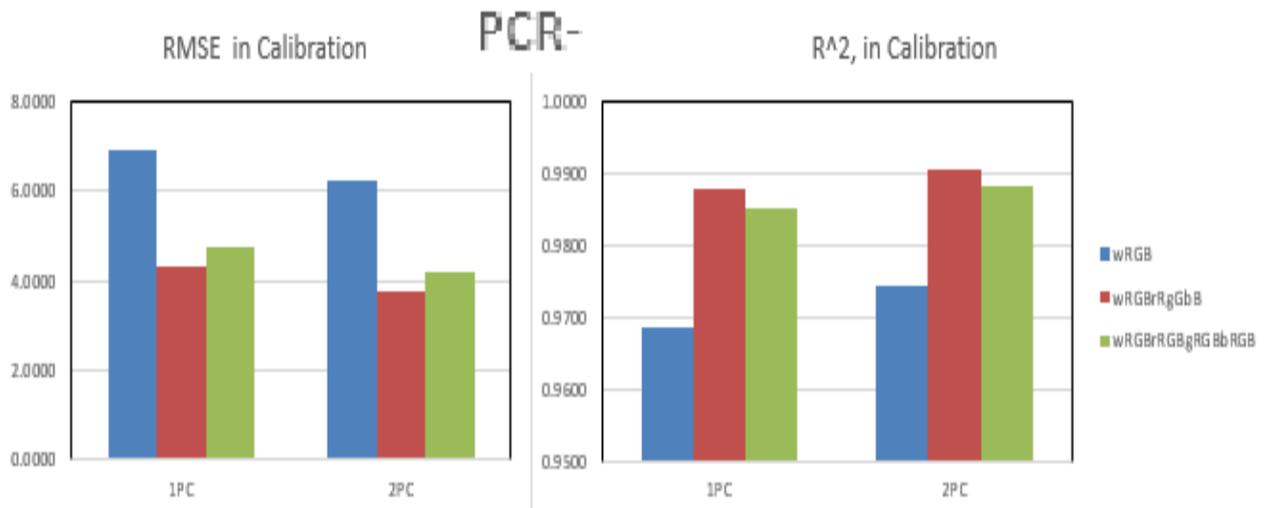


Figura 15a: Predicción para PLS

Se observa como los valores RMSE disminuyen utilizando 6 variables: wRGB rR gG bB. La inclusión de las 12 variables no disminuye el error, más bien lo aumenta ligeramente.

El uso de PC1 da un error que apenas disminuye con la introducción de PC2. En este caso, se propondría el uso de solo PC1 para una regresión más simplificada.

Por otra parte el estudio de  $R^2$ , como la varianza justificada por el modelo, evidencia resultados coherentes con lo anterior. Así la inclusión de 6 variables, wRGB rR gG bB, aumenta de nuevo y de modo importante la varianza justificada por el modelo. Sin embargo, la inclusión de las 12 variables vuelve a disminuir el valor de  $R^2$ .

La argumentación sobre la inclusión de PCs en la regresión pone de manifiesto que con PC1 y 6 variables se obtienen valores superiores al 0.987, y con la inclusión del PC2 prácticamente se justifica el 0.99 de la varianza total de los datos. Es decir, teniendo en cuenta el valor de  $R^2$ , la calibración con PCR si que experimenta mejora con la inclusión de PC1 + PC2.

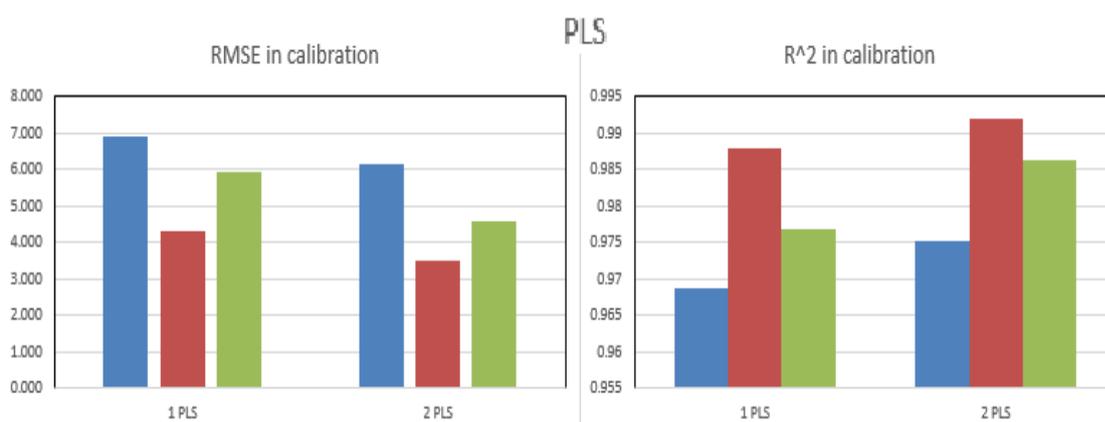


Figura 15b: Predicción para PLS

En la calibración con PLS se evidencia muy claramente la suficiencia de las calibraciones con 6 variables. Ya que la inclusión de las 12 variables aumenta RMS y disminuye  $R^2$ .

Por otra parte, el incremento de variables latentes, PLS1 a PLS1+2 disminuye el RMSE y aumenta la varianza y el valor  $R^2$ .

Como resumen se compara en la tabla los valores de RMSE, u unidades similares, en las diferentes condiciones de calibración:

RMSE, u	Calibración			Validación			Test		
	MLR	PCR	PLS	MLR	PCR	PLS	MLR	PCR	PLS
wBRG	2,6	6,3-6,9	6,2-6,9	2,8	3,5-4,2	3,1-4,1	2,5	9,5-10,0	9,4-10,0
wBRGrRbBgG	2,6	3,5-4,3	2,2-4,3	3,1	3,4-4,3	2,2-4,3	2,0	5,2-7,5	2,9-7,5
wBRGrBRGgBRGgBRG	2,8	4,0-4,7	3,0-4,6	8,6	4,0-4,9	2,9-4,7	1,5	6,0-8,1	4,7-7,7

Figura 16: Resultados RMSE en conjunto

Los resultados evidencian que la regresión MLR, con 6 variables cromáticas, produce los menores valores de RMSE, **en conjunto**. Y resaltan de modo concreto sus bajos valores en la aplicación de ‘test’ . Bien es verdad que son algo mejorados ( RMSE pasa de 2,0 a 1,5) utilizando 12 variables. Pero éstas 12 variables se ha visto que tienden a proporcionar ‘sobre-información’ y mayor variabilidad en los resultados.

### 3 – Aplicación de una calibración multivariable en caso de muestra real:

Se realizan estudios experimentales para la aplicación de la determinación de sulfitos en una muestra real. Se ha elegido: vino blanco. Este es un ejemplo de muestra que se caracteriza por tener sulfito como aditivo estabilizante.

El método diseñado pretendía:

interpolarse el resultado de una muestra directa-sin tratamiento previo- en recta de calibrado de solo analito. Es decir, ser capaces de superar los efectos de matriz mediante el uso de calibraciones robustas.

Se han realizado estudios previos, analizando diferentes cantidades de muestra real, para seleccionar condiciones en las que la calibración compensara el efecto matriz.

Los resultados muestran rangos de muestra en las que mediante calibraciones múltiples MLR, PCR y PLS y 6 variables presentan bajo sesgo entre resultados y valores de referencia.

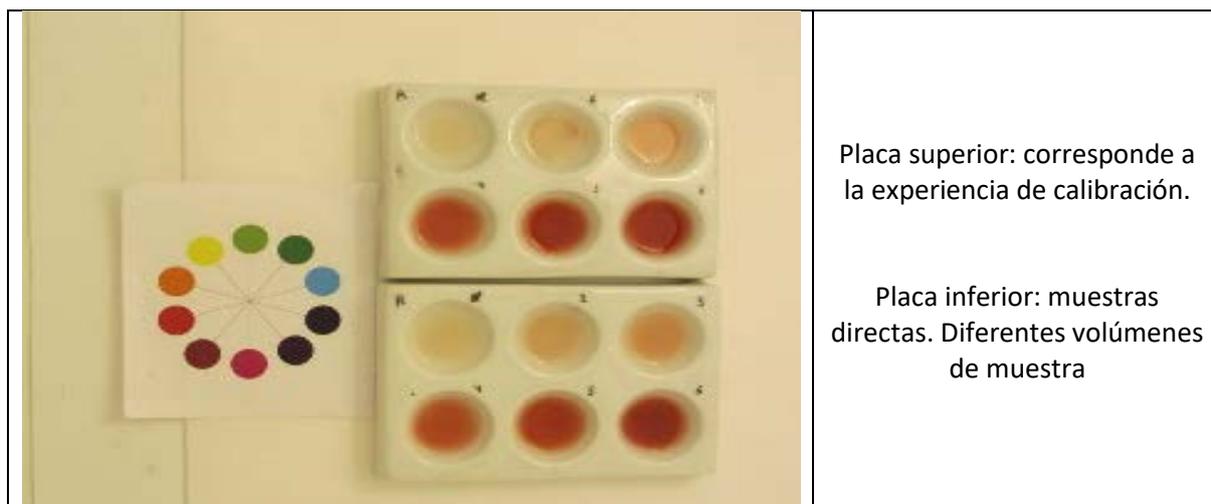
En la tabla siguiente se muestran valores de la determinación de Sulfito total: libre + combinado mediante la interpolación de la coloración de una muestra sin tratamiento en diferentes tipos de regresión.

Condiciones experimentales:

Experimentos de calibración: según la tabla 4.

Preparación de la muestra de análisis: según la tabla 4. Se toma un volumen de muestra directa comercial, sin tratamiento, entre 100-200 ul.

En la Figura se ofrece una imagen de un experimento tipo:



Placa superior: corresponde a la experiencia de calibración.

Placa inferior: muestras directas. Diferentes volúmenes de muestra

Figura 17: Ejemplo de experimento con muestra real

La muestra real ha sido analizada previamente mediante método de referencia de la OIV, método Ripper, 'iodimétrico'.

Resultados: valores medios de 2 muestras- en cada muestra 3 réplicas-, y el intervalo de confianza, IC, al nivel 95%.

	MLR	PCR		PLS		Iodometría
	wRGB	wRGB	wRGBRgGbB	wRGB	wRGBRgGbB	
mg/L SO <sub>2</sub>	260,63	242,19	240,85	241,84	237,07	251,19
IC, mg/L SO <sub>2</sub>	8,72	9,84	8,61	7,07	8,24	4,16

Tabla 4: Comparativa de resultados

El estudio estadístico de los resultados frente al valor de referencia, mediante el test t, al 95% de confianza, da los valores:

$$T_{exp} = \frac{\text{abs}(\text{valor ref} - \text{valor exp}) \times \sqrt{N}}{s}$$

	MLR	PCR		PLS		Iodometría
	wRGB	wRGB	wRGBRgGbB	wRGB	wRGBRgGbB	
	260,63	242,19	240,85	241,84	237,07	251,19
	8,72	9,84	8,61	7,07	8,24	4,16
s	0,9705449	1,09520205	0,95830179	0,78689822	0,91712041	
T exp	13,7553409	11,6215287	15,2592517	16,803821	21,7732537	
t teórico	12,7062047					

Los valores demuestran que estadísticamente al 95% de confianza la calibración PCR y seis variables muestran unas diferencias que pueden ser justificadas por error aleatorio. En el resto de calibraciones se superan las condiciones.

Otros Experimentos de pre-tratamiento de la muestra han sido realizados. En concreto se trata con NaOH 4 M, hasta alcalinidad, para desplazar sulfitos a su modo libre. Los resultados no mejoran respecto a los valores anteriores.

## Capítulo 4. Conclusiones y Bibliografía

---

### CONCLUSIONES

- Se ha estudiado la creciente implantación de técnicas de calibración multivariable en la química.
- Se han desarrollado diferentes calibraciones multivariadas (MLR, PCR, PLS) para llevar a cabo la predicción y caracterización de la varianza cromática de la reacción de boedeker en un rango dinámico de 0-150 mg/L utilizando diferentes iluminaciones y la obtención de colores RGB.
- Se ha estudiado la influencia de tensoactivos en la reacción.
- Se han caracterizado la varianza muestral y cromática, observando la influencia de la primera sobre muestras de analito patrón, y se ha visto que en nuestro caso concreto disminuía la varianza.
- Se ha estudiado, evaluado y validado la capacidad de predicción mediante MLR, PCR y PLS en sistemas cromogénicos, mejorando sobre calibración simple.
- Se obtiene que la combinación de las 6 variables wRGBrGgB proporciona los mejores resultados para determinar sulfitos en muestras reales de vino.

## BIBLIOGRAFÍA

- 1 - Ramis, G.; García, M. C.; *Quimiometría*, (2001, 9, 200).
- 2 - Ramis, G.; García, M. C.; *Quimiometría*, (2001, 9, 201-228).
- 3 - Wold, S.; Sjöström, M.; Eriksson, L.; PLS-regression: a basic tool of chemometrics, *Chemom. Intell. Lab. Syst.*, 2001, 58,109-130.
- 4 - Mehmood, T.; Hovde Liland, K.; Snipen, L.; A review of variable selection methods in Partial Least Square Regression, *Chemom. Intell. Lab. Syst.*, (2012, 118, 62-69).
- 5 - López Molinero, A.; et al, *Talanta*, (2013, 103, 236–244).
- 6 - López Molinero, A.; et al, *Microchem J* (2010, 96, 380-385)
- 7 - D. Liñán, Postgrado de Iniciación a la Investigación, Univ. Zaragoza, Fac. Ciencias (2006).
- 8 - Gómez, M.A.; Postgrado de Iniciación a la Investigación, Univ. Zaragoza, Fac. Ciencias (2009).
- 9 - Jiménez, J. Postgrado de Iniciación a la Investigación, Universidad de Zaragoza, Facultad de Ciencias (2009).
- 10 - Pérez, M. Postgrado de Iniciación a la Investigación, Univ. Zaragoza, Fac. Ciencias (2014).
- 11 - Berlín, P. Trabajo fin de Grado, Univ. Zaragoza, Fac. Ciencias (2015).
- 12 - Esteban, A. Trabajo fin de Grado, Univ. Zaragoza, Fac. Ciencias (2018).
- 13 - Gómez, L. Trabajo fin de Grado, Univ. Zaragoza, Fac. Ciencias (2016).
- 14 - Reglamento (CE) nº 1493/1999 del Consejo de 17 de mayo de 1999 por el que se establece la organización común del mercado vitivinícola. *OJ L 179, 14.7.1999, p. 1–84*
- 15 - Reglamento (CE) nº 1622/2000 de la Comisión de 24 de julio de 2000 que fija determinadas disposiciones de aplicación del Reglamento (CE) nº 1493/1999, e introduce un código comunitario de prácticas y tratamientos enológicos.
- 16 - Andrade, C.; et al., *Inorg. Chem.*, (1972, 11, 648- 650).