

Métodos multivariantes aplicados a la efectividad de los disolventes en la industria farmacéutica

Application of multivariate methods in the effectiveness of solvents in the pharmaceutical industry

<https://doi.org/10.17981/ingecuc.14.1.2018.12>

Artículo de investigación científica. Fecha de recepción: 12/02/2018. Fecha de aceptación: 25/05/2018

Enrique De La Hoz Domínguez 

Universidad Oberta de Catalunya. Barcelona (España).
edelahoz@uoc.edu

Tomás Fontalvo Herrera 

Universidad de Cartagena. Cartagena (Colombia).
tfontalvoh@unicartagena.edu.co

Adel Mendoza 

Universidad del Atlántico. Barranquilla (Colombia).
adelmendoza@uniatlantico.edu.co

Para citar este artículo:

E. De La Hoz Domínguez, T. Fontalvo Herrera y A. Mendoza, "Métodos Multivariantes aplicados a la efectividad de los disolventes en la industria farmacéutica," INGE CUC, vol. 14, no. 1, pp. 133-140, 2018. DOI: <https://doi.org/10.17981/ingecuc.14.1.2018.12>

Resumen

Introducción– Los disolventes son agentes químicos usados en la industria farmacéutica y su importancia radica en que su presencia puede acelerar o retardar la reacción de un fármaco hasta un millón de veces.

Objetivo– La presente investigación analiza los diferentes tipos de disolventes para evaluar la existencia de grupos en los cuales se puedan identificar patrones relacionados con la efectividad de esos disolventes en la producción de fármacos.

Metodología– El estudio consta de 4 fases: 1) análisis de componentes principales, 2) análisis de clúster, 3) análisis discriminante, y 4) interpretación de resultados y conclusiones.

Resultados– Se identificaron tres conglomerados categorizados como supercríticos, microesféricos y biodegradables. La prueba de T de Hotelling arroja un p-valor de 0, poniendo en evidencia la diferencia entre grupos. El discriminante cuadrático arroja una precisión del 96% para la clasificación de disolventes.

Conclusiones– El análisis multivariante permite modelar la efectividad de los disolventes en la industria farmacéutica, generando una metodología objetiva de decisión para la clasificación de disolventes según un enfoque de efectividad.

Palabras clave– Análisis multivariante, disolventes, componentes principales, clustering, clasificación

Abstract

Introduction– Solvents are chemical agents used in the pharmaceutical industry. Their importance is that their presence can accelerate or delay the reaction of a drug up to a million times.

Objective– The present investigation analyzes the different types of solvents in order to evaluate the existence of groups in which patterns related to the effectiveness of the mentioned solvents in medicine production can be identified.

Methodology– The study is comprised of 4 phases: 1) Principal component analysis; 2) Cluster analysis; 3) Discriminant analysis; 4) Interpretation of results and conclusions.

Results– Three clusters, categorized as supercritical, microspherical and biodegradable, were identified. The Hotelling T test yields a p-value of 0, evidencing the difference between groups. The quadratic discriminant yields a precision of a 96% for the classification of solvents.

Conclusions– The multivariate analysis allows modeling the effectiveness of solvents in the pharmaceutical industry. Hence, generating an objective decision methodology for the classification of solvents according to an effectiveness approach.

Keywords– Multivariant analysis, solvents, principal components, clustering, classification

I. INTRODUCCIÓN

Los disolventes son importantes en el manejo de residuos peligrosos, la industria farmacéutica y la protección del medio ambiente. En esta investigación se presta atención particular a la producción farmacéutica, conceptualizando los disolventes como componentes claves en este proceso de producción. Su importancia radica en que su presencia puede acelerar o retardar la reacción de un fármaco hasta un millón de veces [1]. Dependiendo del tipo de fármaco que se intente fabricar, primero, se clasifican los disolventes en polares y apolares.

Sin embargo, nuevas investigaciones presentan las características deseables para los disolventes. Así, en [2] resaltan las ventajas de utilizar disolventes que sean amigables con el ambiente y presentan los beneficios de los disolventes orgánicos y biodegradables. Por su parte, autores [3] muestran la importancia de contar con disolventes capaces de proveer fiabilidad en la transportación del medicamento. En otras investigaciones [4], se resalta la capacidad de resiliencia del disolvente ante eventos climáticos adversos. Para otros autores [5], es importante que los disolventes permitan la rápida asimilación del fármaco y cuya cristalización para la absorción sea lo más controlable posible.

El creciente índice de epidemias y endemias nuevas demanda a la industria farmacéutica fármacos efectivos y de rápida acción. Esto, sumado a la complejidad de las interacciones entre soluto y solvente y la gran cantidad de solventes disponibles en el mercado, hacen que la elección del solvente ideal para un proceso de fabricación farmacéutica sea realmente difícil. Usualmente, se usan las propiedades físico-químicas, como el punto de ebullición y de fusión, presión de vapor, índice de refracción, densidad, viscosidad, tensión superficial, conductividad, entre otras. Sin embargo, muchas de las clasificaciones toman en cuenta solo unas pocas propiedades, resultando en clasificaciones pobres, considerando el gran número de propiedades que pueden ser usadas para caracterizar un solvente.

Existen diferentes referencias acerca del uso de métodos estadísticos multivariantes para la clasificación y selección de solventes. Parker [6] implementa un análisis factorial para la clasificación de disolventes, encontrando una clasificación similar a la de [7], las cuales se basan fundamentalmente en principios químicos. En otras investigaciones [8], utilizan un doble análisis de componentes principales, generando un modelo sistemático para la selección de disolventes para reacciones químicas. Por su parte, en otros estudios [9], se propone uno de los sistemas de clasificación de disolventes más populares y usados por medio de celdas por cromatografías. Recientemente, algunos estudios [10] proponen una clasificación de las propiedades de los disolventes usando métodos de equilibrio termodinámico.

En [11] implementan una regresión multivariada para obtener ecuaciones predictivas en función de las principales variables físicas.

Consecuentemente con lo anterior, esta investigación analiza, los diferentes tipos de disolventes de acuerdo con su comportamiento en la cantidad de cada una de las características físico-químicas que posean con la finalidad de evaluar la posible existencia de grupos en los que se puedan identificar patrones relacionados con la efectividad de dichos disolventes, teniendo en cuenta el fármaco que se quiere producir. Para lograr esto, se utiliza como metodología una serie de criterios articulados y relacionados con el análisis multivariante (Fig. 1) desarrollando, en primer lugar, un análisis de componentes principales para reducir las dimensiones del conjunto de datos de estudio. Luego, se desarrolla un análisis de clúster, intentando identificar la existencia de grupos en los disolventes estudiados, para, finalmente, implementar un sistema de clasificación basado en análisis discriminante. El uso de componentes principales y el análisis de clúster permiten evaluar las diferencias entre los grupos de disolventes, analizando las características comunes entre los miembros de cada clúster y definiendo su nivel de relacionamiento con los dos primeros componentes principales.

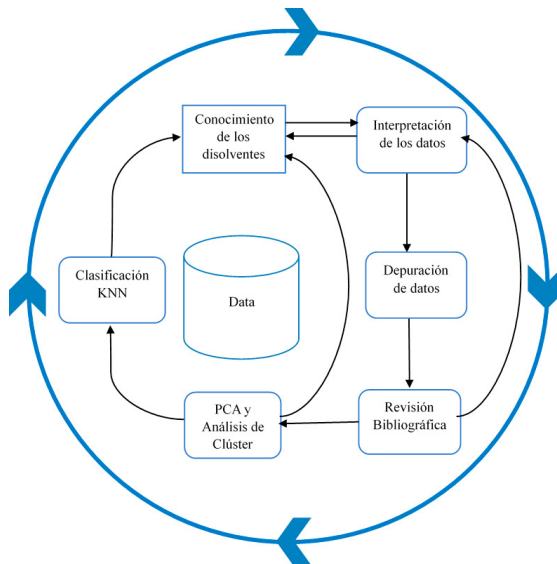


Fig. 1. Metodología de análisis multivariante.
Fuente: Autores.

II. DATOS EXPERIMENTALES

Las variables del estudio son las propiedades físico-químicas de 103 solventes (Tabla 1), las cuales son tomadas de la referencia [12]. Se analizan 9 variables: 1) melting point, 2) boiling point, 3) dielectric, 4) dipole moment, 5) refractive index, 6) ET30, 7) density, 8) LogP y 9) solubility.

TABLA 1. PROPIEDADES FISIQUÍMICAS UTILIZADAS EN EL ESTUDIO.

Propiedad	Definición
Melting point	Temperatura en la cual el disolvente pasa de estado sólido a estado líquido (grados Celsius)
Boiling Point	Temperatura en la cual el disolvente pasa de estado líquido a estado gaseoso (grados Celsius)
Dielectric	Constante dieléctrica
Dipole moment	Distancia de los átomos (picómetro)
Refractive index	Índice de refracción
ET30	Parámetro de la polaridad de disolución
Density	Densidad del solvente (gramo por centímetro cúbico g/cm ³)
LogP	Coefficiente de reparto para la diferencia de solubilidad en mezclas de dos fases
Solubility	Solubilidad (gramos de soluto por 100 gramos de solvente)

Fuente: Autores.

III. MARCO CONCEPTUAL

A. Fabricación de fármacos

De acuerdo con [13], el proceso para fabricar un fármaco con la intervención de un disolvente pasa por tres etapas: la primera es la obtención del principio activo, en el que este se sintetiza, se separa y se purifica (normalmente por cristalización); durante esta síntesis el disolvente actúa como medio de reacción, colocando en contacto los distintos reactivos y solubilizándolos. La segunda etapa es la formulación, en la cual el principio activo se prepara de acuerdo con la vía de administración del fármaco (vía oral, vía sanguínea, etc.). La tercera etapa es el empaquetado, cuya función es que pueda ser removido fácilmente del medicamento, en la medida de lo posible, para evitar toxicidad, y, de ser posible, reciclarse.

Las clasificaciones de disolventes crean dos categorías: los *disolventes polares* y los *apolares*. En la primera se encuentran aquellos en cuyas moléculas la distribución de la nube electrónica es asimétrica; por lo tanto, la molécula presenta un polo positivo y otro negativo separados por una cierta distancia [14]. Por otra parte, en los disolventes apolares, por lo general, se habla de sustancias de tipo orgánico y en cuyas moléculas la distribución de la nube electrónica es simétrica; por lo tanto, estas sustancias carecen de polo positivo y negativo en sus moléculas [14].

Los científicos se enfrentan a varios retos para una elección apropiada de disolventes, ya que en los últimos años se motiva la disminución de la cantidad de disolventes usados por unos menos invasivos. Sin embargo, la sustitución de disolventes no siempre es posible por la implicación que tienen estos en el mecanismo de la reacción. Nuevos esfuerzos han hecho posibles disolventes a base de fluidos supercríticos, líquidos iónicos, medios acuosos, entre otros, los cuales serán insertados en la industria.

B. Análisis de componentes principales

El análisis de componentes principales (PCA) es un método estadístico clásico para transformar las variables de un conjunto de datos en un nuevo conjunto de variables no correlacionadas llamadas componentes principales. El PCA puede ser usado para reducir la dimensionalidad de un conjunto de datos mientras conserva la mayor cantidad posible de variables de los datos. En resumen, el PCA se constituye de: 1) los vectores W_j que están dados por los autovectores dominantes (aquellos con el autovalor asociado más grande); 2) la matriz de covarianza muestral,

$$\text{Covarianza Muestral} = \sum \frac{(t_n - \bar{t})(t_n - \bar{t})}{N} \quad (1)$$

donde es la media muestral, de tal forma que $S_{w_j} = \lambda_j W_j$; y 3) las componentes principales observadas en el vector tn están dadas por el vector $X_n = wt(t_n - \bar{t})$, cuando $W = (w_1, w_2, \dots, w_q)$. Las variables x_j entonces no están correlacionadas, de tal forma que la matriz de varianzas es diagonal con elementos λ_j .

$$\text{Matriz de varianzas} = \sum \frac{(X_n X_n)}{N} \quad (2)$$

C. Cálculo de distancias

A priori, es importante mencionar que la medición de los tipos de disolventes con respecto a sus características físicas es continua, de acuerdo con su propia escala de medida. Adicionalmente, el cero no tiene ninguna connotación especial para el análisis y no se requiere eliminar redundancias. Por estas razones, se ha utilizado, para el presente estudio, la distancia euclídea y se han estandarizado los datos con la finalidad de disminuir la discrepancia por unidades.

La distancia euclídea es la más usada en el desarrollo de procesos de aprendizaje no supervisado, sin embargo, es sensible a cambios de escala y considera las T variables independientes [15]. En [16] se proponen transformaciones que permiten reducir y/o desaparecer estos problemas. Se propone la estandarización de las variables, intentando que el aporte de cada variable se encuentre dentro del intervalo (0,1).

$$\text{Distancia Euclídea} = \sqrt{\sum_{i=1}^T (x_{it} - x_{jt})^2} \quad (3)$$

D. Análisis de clúster

El análisis de clúster es una técnica de análisis multivariante cuyo propósito es crear grupos intentando que las observaciones pertenecientes a un grupo estén muy cercanas entre ellas y apartadas de las observaciones ubicadas en otro grupo. Existen cuatro categorías de algoritmos de clustering: 1) de particionamiento; 2) basados en densidad; 3) basados en redes; y 4) jerárquicos. Los algoritmos de partición, tal como k-means y de partición alrededor de los medoides, iterativamente refinan un conjunto de k clústeres, y, por lo general, no escalan bien para grandes conjuntos de datos.

$$\text{Criterio de grupos} = \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (4)$$

Los algoritmos basados en densidad son capaces de generar clústeres de tamaño arbitrario y de lidiar con valores extremos. Los algoritmos basados en red reducen el espacio de agrupamiento a celdas dentro de una red, permitiendo un agrupamiento eficiente de conjuntos de datos de gran tamaño. Los algoritmos jerárquicos pueden ser de aglomeración o divisivos; en la forma aglomerada repetidamente se fusionan dos clústeres, mientras que en la forma divisiva repetidamente se divide un clúster en dos.

V. RESULTADOS

Las relaciones existentes entre las variables de estudio no pueden ser vistas a priori según el enfoque de "efectividad", por lo que es necesario aplicar el análisis de componentes principales para reducir la dimensionalidad del conjunto de datos con la finalidad de poder darles una explicación más específica a las variables. En el desarrollo del proceso se ha utilizado el paquete Factominer [17].

Como se observa en la Tabla 2, los tres primeros componentes explican el 76,96% de la variabilidad total de la información, por lo que, para esta investigación, se lo considerará suficiente. No obstante, para efectos de gestión de la información e interpretación, los dos primeros componentes, como se aprecia en la Tabla 2, contienen un porcentaje alto (68,45 %), por lo que el análisis posterior se desarrollará utilizando dos componentes.

TABLA 2. AUTOVALORES POR COMPONENTE PRINCIPAL.

Componente	Auto valor	% de varianza	% de varianza acumulada
1	1,90	40,65	40,65
2	1,57	27,76	68,45
3	0,87	8,54	76,96
4	0,76	6,54	83,51
5	0,71	5,72	89,24
6	0,64	4,67	93,91
7	0,53	3,17	97,09
8	0,43	2,13	99,22
9	0,26	0,77	100

Fuente: Autores.

En el componente uno (1), la variable con mayor correlación al eje de forma negativa es LogP, y de forma positiva se encuentran: solubilidad, ET30, momento dipolar y dielectricidad (Tabla 3). Este eje se denomina de *transformación* o de *características no constantes*, ya que las variables son sensibles a las condiciones atmosféricas y a otros reactivos [18]-[19].

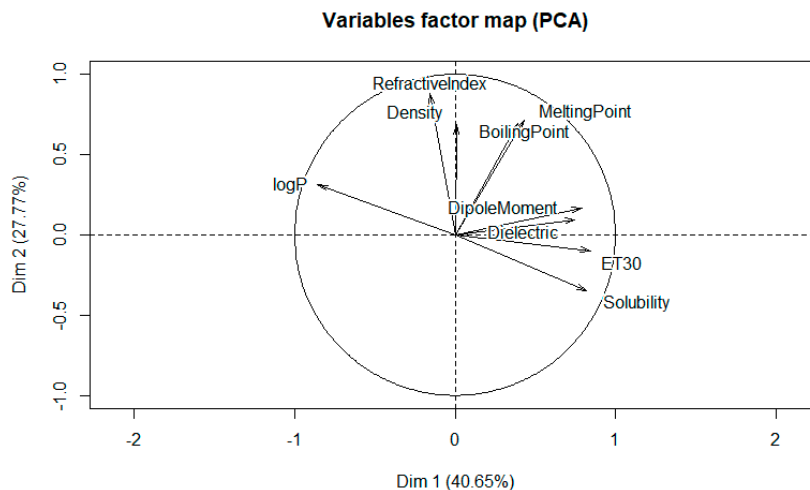


Fig. 2. Biplot para las dos primeras componentes principales.

Fuente: Autores.

En el componente dos (2) tienen una correlación positiva alta con el eje las siguientes variables: punto de fusión, punto de ebullición, densidad e índice refractario. Por lo que se reconoce a este eje como de *características propias o naturales de los disolventes*.

TABLA 3. CORRELACIÓN DE VARIABLES POR COMPONENTES PRINCIPALES.

Variable	Componente 1	Componente 2
Melting Point	0,39	0,69
Boiling point	0,43	0,71
Dielectric	0,75	0,09
Dipole moment	0,79	0,17
Refractive index	-0,16	0,88
ET30	0,84	-0,10
Density	0,01	0,69
Lop P	-0,86	0,31
Solubility	0,82	-0,35

Fuente: Autores.

A. Resultados del Análisis de clúster

Con la finalidad de identificar grupos dentro de los disolventes, se emplearon métodos jerárquicos y no jerárquicos para analizar la mejor opción de agrupamiento. Lo primero que se pretende es que los grupos sean lo más compactos posibles, y que la distancia entre los grupos sea máxima y la distancia intragrupal sea mínima, para ello se utiliza el método Ward [20].

Parece importante evaluar la posibilidad del método "average" por tratarse de grupos cuyo promedio se espera sea parecido (en principio deberían tener las mismas proporciones de características físicas). Posterior a la identificación de grupos, se procedió a identificar los patrones de comportamiento de los grupos.

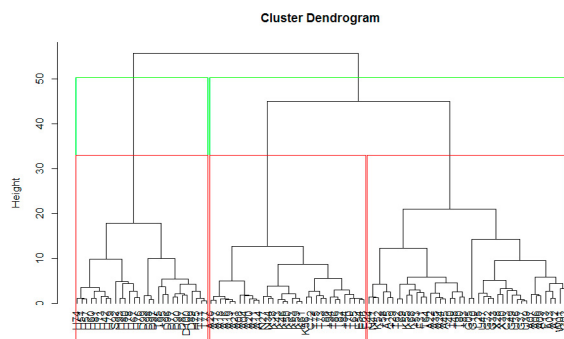


Fig. 3. Clúster jerárquico bajo el método Ward.

Fuente: Autores.

Se observa, en la Fig. 3, la formación de dos clústeres (línea verde) y tres clústeres (línea roja) superpuestos, la correlación entre los datos y la relación cofenética

con el método Ward es del 50%. A pesar de que la correlación es baja, es importante mencionar que con este método se garantizan "grupos compactos". A través del clúster jerárquico se tiene una idea de que el número de grupos serán dos o tres, porque a partir de cuatro grupos se generan clústeres con dos individuos. Por lo tanto, en el clúster no jerárquico se hará énfasis en estos dos números de grupos (2 y 3).

Se consideró importante desarrollar un proceso de clúster no jerárquico para confirmar el número ideal de grupos en el desarrollo posterior de la investigación, para esto se utilizarán las métricas de desempeño Tess, Gradiente y Pseudo-F [21]. En estas, se examinaron hasta siete grupos conformados, aunque a priori se conoce, por el clúster jerárquico, la existencia de dos o tres grupos generales.

TABLA 4. INDICADORES DE DESEMPEÑO EN EL CLÚSTER NO JERÁRQUICO.

Grupo	TESS	Gradiente	Pseudo-F
1	918		
2	676,39	26,31	36,10
3	511,15	24,43	39,79
4	440,76	13,77	35,73
5	383,40	13,01	34,16
6	344,33	10,19	32,32
7	308,20	10,49	31,66

Fuente: Autores.

Según el Tess, el quiebre, que indica el número de clústeres, sería de dos, y según PseudoF, serían tres (Tabla 4). Así que, a criterio de los investigadores, se mantendrá entre dos y tres grupos.

B. Visualización del clúster

Es importante generar herramientas visuales para el análisis de los resultados en los procesos de clasificación. Según [22], el proceso de creación de imágenes mentales es fundamental para los procesos de pensamiento y creación de nuevo conocimiento en las personas. En [23], se plantea el enfoque visual como una materialización de la experiencia, considerando los conceptos y su representación como algo indivisible.

Así, de acuerdo con el resultado del análisis de clúster, se realiza la representación de los grupos utilizando los componentes (Fig. 4). En el grupo de la izquierda encontramos la mayor concentración de disolventes que se relacionan con los de tipo microesfera, que son sistemas que se han tratado de modificar apropiadamente para que el transporte de conjugados poliméricos mantenga una baja solubilidad y no causen problemas de síntesis; estos son utilizados para la elaboración de inyecciones para el flujo sanguíneo. Así, este grupo se rotulará como *microesféricos*.

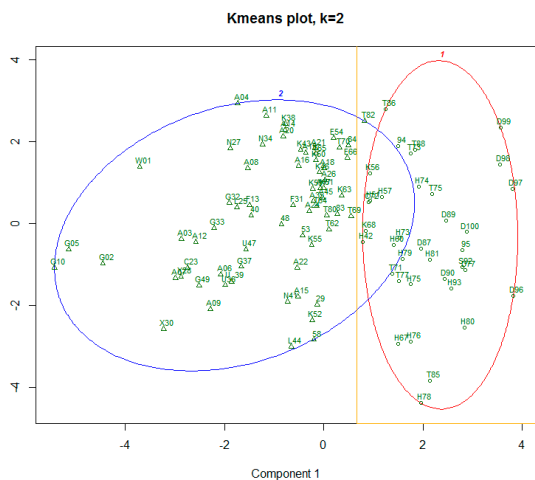


Fig. 4. Visualización para dos clústeres. Fuente: Autores.

En el grupo de la derecha (Fig. 4) se aprecia la mayor concentración de disolventes que se relacionan con materiales introducidos en los tejidos corporales con propósitos terapéuticos específicos de diagnóstico o con propósitos preventivos, los cuales tienen una base en biomateriales que es controlada por agentes bioactivos mediante el empleo de sistemas poliméricos biodegradables a prueba de bioerosión y bioabsorción [24]-[25]. Así, este grupo se rotulará como *biodegradables*.

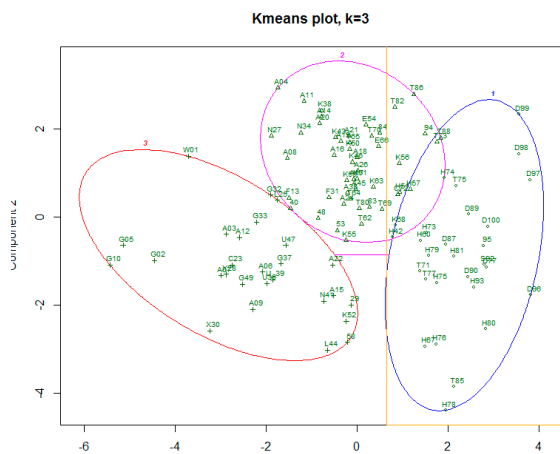


Fig. 5. Visualización para tres clústeres. Fuente: Autores.

En la visualización para tres clústeres observados en la Fig. 5, el grupo de la izquierda y de la derecha mantienen los patrones antes mencionados. Sin embargo, en el del medio se encuentra la mayor concentración de disolventes que se relacionan con fluidos supercríticos, que intervienen netamente en el proceso de cristalización. Una etapa muy importante

porque puede ejercerse un control en el tamaño y la forma de los cristales, características que influyen en la absorción del medicamento por el organismo [1]. Así, este grupo se rotulará como *supercrítico*.

Tabla 5. CENTRO DE GRAVEDAD PARA LAS VARIABLES DE LOS TRES GRUPOS.

Propiedad	Micro-esférico	Bio-degradable	Super-crítico
Melting point	-49,50	-68,17	-5,04
Boiling point	121,43	101,42	204,62
Dielectric	4,08	12,73	46,62
Dipole moment	2,81	6,05	11,15
Refractive index	1,47	1,39	1,47
ET30	34,62	41,92	47,25
Density	1,09	0,86	1,05
LogP	3,02	0,60	0,13
Solubility	-2,53	0,20	0,56

Fuente: Autores.

Los centros de gravedad para cada una de las variables expresadas en la tabla 5 proveen evidencia de que cada grupo tiene la proporción de las cantidades físicas correspondientes a lo que previamente fue descrito como 3 grupos. En conclusión, se tomarán 3 grupos debido a que la formación de éstos, en términos teóricos, es coherente, es un número apropiado para la gestión de la información y los estadísticos han orientado hacia un número similar de grupos. Para los siguientes efectos del estudio vamos a llamar al “biodegradable” grupo 1, al “supercrítico” grupo 2 y al “microesférico” grupo 3.

C. Análisis discriminante

Como es el objetivo de la empresa farmacéutica probar nuevos disolventes para los fármacos en proceso de diseño y pre-producción, la idea es que, una vez establecidas las características del fármaco, se pueda determinar el tipo de disolvente y a qué patrón recogido por el grupo pertenecería. En este caso, la variable dependiente es creada con valores de 1 a 3 según los grupos previamente creados. A través del *Box’s M test*, para verificar el supuesto de igualdad de varianzas, se rechaza la hipótesis nula de varianzas homogéneas al 95% de nivel de confianza. En este caso, se sugiere realizar el discriminador cuadrático, no obstante, se tomará en cuenta también el lineal para efectos de interpretación.

TABLA 6. PREDICCIONES PARA CLASIFICADOR LINEAL Y CUADRÁTICO.

	Lineal			Cuadrático		
	Grupo 1	Grupo 2	Grupo 3	Grupo 1	Grupo 2	Grupo 3
Grupo 1	23	0	0	26	0	2
Grupo 2	4	27	2	0	28	2
Grupo 3	0	1	46	1	0	44

Fuente: Autores.

Utilizando el discriminante lineal, el 85% (23/27) de los disolventes del grupo 1 estarán bien clasificados en él, el 96% en el grupo 2 y el 96% en el grupo 3 (Tabla 6). Así que se podría decir que es una clasificación efectiva. Al proceder con el discriminante cuadrático por la no homogeneidad de varianzas, se pone en evidencia una mejor clasificación de los grupos con respecto a la lineal con un 96% de los individuos del grupo 1 clasificados efectivamente en dicho grupo, un 100% para el grupo 2 y un 92% para el grupo 3.

D. Inferencia multivariante

El estadístico T^2 de Hotelling será utilizado para contrastar las medias grupales de los clústeres creados.

TABLA 7. TEST DE PERMUTACIONES DE HOTELLING PARA LOS GRUPOS

Parameter	Permutación 1		
	1	2	3
Test stat	29,57	24,32	33,79
Numerator df	9	9	9
Denominator df	66	65	45
Permutation P-value	0	0	0
Number of permutations	1000	1000	1000

Fuente: Autores.

El estadístico T^2 de Hotelling será utilizado para contrastar las medias grupales de los clústeres que fueron previamente creados. Para el caso no paramétrico, con el test de Hotelling de permutaciones, se obtiene que las medias de los grupos no son iguales (Tabla 7).

V. CONCLUSIONES

La formación de grupos con un enfoque de efectividad (según un propósito) fue posible con un número de 3 conglomerados caracterizados por disolventes de tipo supercrítico, microesférico y biodegradable, los cuales han mostrado claramente un patrón de comportamiento y conforme a lo mencionado en el

Marco Conceptual fácilmente pueden adaptarse a cualquiera de las fases de elaboración de fármacos, o, a su vez, incorporar nuevos diseños de disolventes que sean actualmente investigados.

El discriminador cuadrático resultó ser una buena alternativa con menor probabilidad de error de clasificación. La prueba de T de Hotelling muestra que las medias de los conglomerados son efectivamente diferentes, y, en efecto, el centro de gravedad da valores distintos para cada grupo, con lo que efectivamente se puede reconocer la formación acertada de grupos, con ello se cumple el objetivo de la investigación. Como valor fundamental de esta investigación se aporta a la comunidad científica un modelo de clasificación o metodología, en donde se articulan y sistematizan diferentes enfoques de análisis multivariante para identificar y analizar nuevos patrones de efectividad en los disolventes y así facilitar la producción de fármacos.

Como proyección de esta investigación se invita a la comunidad científica a abordar el mismo objeto de esta investigación utilizando otras herramientas de análisis multivariante, especialmente la técnica de redes neuronales, con lo cual se puedan contrastar los resultados de esta investigación con otro enfoque multivariante.

VI. FINANCIACIÓN

Artículo de investigación científica derivado del proyecto de investigación titulado "Comportamiento de los indicadores de productividad y financieros en el sector de extracción y explotación de minerales en Colombia, mediante análisis discriminante" y financiado con recursos del plan de fortalecimiento del grupo de investigación Calidad y productividad organizacional integral (CYPROI) avalado por la universidad de Cartagena.

REFERENCIAS

- [1] V. Sáenz, E. Hernández, L. Sanz y I. Katime, «Liberación controlada de fármacos. Micropartículas.» *Revista Iberoamericana de Polímeros*, vol. 5, no. 2, pp. 87-101, 2004.
- [2] C. Capello, U. Fisher y K. Hungerbühler, «What is a green solvent? A comprehensive framework for the environmental assessment of solvents.» *Green Chemistry*, vol. 9, no. 9, pp. 927-934, 2007. <http://dx.doi.org/10.1039/b617536h>

- [3] S. C. Shen, W. K. Ng, J. Hu, K. Letchmanan, J. Ng y R. B. H. Tan, «Solvent-free direct formulation of poorly-soluble drugs to amorphous solid dispersion via melt-absorption.» *Advanced Powder Technology*, vol. 28, no. 5, pp. 1316-1324, 2017. <https://doi.org/10.1016/j.apt.2017.02.020>
- [4] O. S. Hammond, D. Bowron, A. J. Jackson, T. Arnold, A. Sanchez-Fernandez, N. Tsapatsaris y K. J. Edler, «Resilience of Malic Acid Natural Deep Eutetic Solvent Nanostructure to solidification and Hydration.» *The Journal of Physical Chemistry B*, vol. 121, no. 31, pp. 7473-7483, 2017. <https://doi.org/10.1021/acs.jpcc.7b05454>
- [5] R. Minphimai, S. Piriyaprasarth y P. Sriamornsak, «Application of Artificial Neural Networks for Prediction of Ibuprofen Yield from Ultrasound-Assisted Anti-Solvent Crystallization.» *Advanced Materials Research*, vol. 1060, pp. 145-148, 2015. <https://doi.org/10.4028/www.scientific.net/AMR.1060.145>
- [6] A. Parker, «Protic-dipolar aprotic solvent effects on rates of bimolecular reactions.» *Chemical Reviews*, vol. 69, no. 1, pp. 1-32, 1969. <https://doi.org/10.1021/cr60257a001>
- [7] A. de Juan, G. Fonrodona y E. Casassas, «Solvent classification based on solvatochromic parameters: a comparison with the Snyder approach.» *TraC trends in Analytical Chemistry*, vol. 16, no. 1, pp. 52-62, 1997. [https://doi.org/10.1016/S0165-9936\(96\)00079-9](https://doi.org/10.1016/S0165-9936(96)00079-9)
- [8] A. Katritzky, D. Fara, M. Kuanar, E. Hur y M. Karelson, «The classification of solvents by combining classical QSPR methodology with principal component analysis.» *The Journal of Physical Chemistry A*, vol. 109, no. 45, pp. 10323 - 10341, 2005. <https://doi.org/10.1021/jp050395e>
- [9] L. Snyder, «Classification of the solvent properties of common liquids.» *Journal of Chromatography A*, vol. 92, no. 2, pp. 223-230, 1974. [https://doi.org/10.1016/S0021-9673\(00\)85732-5](https://doi.org/10.1016/S0021-9673(00)85732-5)
- [10] N. Bondarev, «Exploratory analysis in thermodynamics of equilibria. Classification and prediction of benzoic acid strength in aqueous organic solvents.» *Russian Journal of General Chemistry*, vol. 86, no. 6, pp. 1221-1228, 2016. <https://doi.org/10.1134/S1070363216060025>
- [11] K. Skalicka-Wozniak y I. Garrard, «A comprehensive classification of solvent systems used for natural product purifications in countercurrent and centrifugal partition chromatography.» *Natural Product Reports*, vol. 32, no. 11, pp. 1556-1561, 2015. <https://doi.org/10.1039/C5NP00061K>
- [12] K. Dunn, «Solvents-OpenMV.net Datasets.» 14 Diciembre 2017. [En línea]. Available: [www.http://openmv.net/info/solvents](http://openmv.net/info/solvents). [Último acceso: 14 Diciembre 2017].
- [13] M. Aulton, *La ciencia del diseño de las formas farmacéuticas*, Leicester: Elsevier, 2004.
- [14] S. F. Sneddon, D. J. Tobias y C. L. Brooks III, «Thermodynamics of amide hydrogen bond formation in polar and apolar solvents.» *Journal of molecular biology*, vol. 209, no. 4, pp. 817-820, 1989. [https://doi.org/10.1016/0022-2836\(89\)90609-8](https://doi.org/10.1016/0022-2836(89)90609-8)
- [15] C. Cuadras, *Nuevos métodos de análisis multivariante*, Barcelona: CMC Editions, 2014.
- [16] H. Breu, J. GIL, D. Kirlpatrick y M. Werman, «Linear Time Euclidean distance transform algorithms.» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 529-533, 1995. <https://doi.org/10.1109/34.391389>
- [17] S. Le, J. Josse y F. Husson, «FactoMiner: A package for Multivariate Analysis.» *Journal of Statistical Software*, vol. 25, no. 1, pp. 1-18, 2008. <https://doi.org/10.18637/jss.v025.i01>
- [18] P. Trodler, R. D. Schmid y J. Pleiss, «Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase.» *BMC Structural Biology*, vol. 9, no. 1, pp. 1-13, 2009. <https://doi.org/10.1186/1472-6807-9-38>
- [19] F. Kerton y R. Marriot, *Alternative Solvents for green chemistry*, vol. 77, London: RSC Publishing, 2014. <https://doi.org/10.1186/1472-6807-9-38>
- [20] J. H. Ward Jr, «Hierarchical grouping to optimize an objective function.» *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236-244, 1963. <https://doi.org/10.1080/01621459.1963.10500845>
- [21] U. Maulik y S. Bandyopadhyay, «Performance evaluation of some clustering algorithms and validity indices.» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650-1654, 2002. <https://doi.org/10.1109/TPAMI.2002.1114856>
- [22] F. Tang, T. Hess, J. Valacich y J. Sweeney, «The effects of visualization and interactivity on calibration in financial decision-making.» *Behavioral Research in Accounting*, vol. 26, no. 1, pp. 25-28, 2013. <https://doi.org/10.2308/bria-50589>
- [23] R. Swedberg, «Can You Visualize Theory? On the Use of Visual Thinking in Theory Pictures, Theorizing Diagrams, and Visual Sketches.» *Sociological Theory*, vol. 34, no. 3, pp. 250-275, 2016. <https://doi.org/10.1177/0735275116664380>
- [24] B. P. Whim y P. G. Johnson, *Directory of solvents*, London: Springer Science & Business Media, 2012.
- [25] B. Trémillon, «Chemistry in non-aqueous solvents,» *Reidel*, p. 289, 1971.

Enrique De La Hoz Domínguez es ingeniero industrial de la Universidad Libre, magíster en Investigación de operaciones por la Universitat de Barcelona, candidato a doctor en Redes y sistemas de información en la Universitat Oberta de Catalunya. Profesor de la Universidad Tecnológica de Bolívar. Sus áreas de interés son los sistemas de recomendación contextuales y Learning Analytics <https://orcid.org/0000-0003-2531-6389>

Tomás Fontalvo Herrera es ingeniero químico de la Universidad del Atlántico, cuenta con una maestría en administración de empresas de la Universidad Nacional de Colombia. En 2013, obtuvo su PhD por la Atlantic International University. Actualmente trabaja como profesor e investigador en la Universidad de Cartagena. Sus áreas de interés son la minería de datos del sector empresarial. <https://orcid.org/0000-0003-4642-9251>

Adel Mendoza Mendoza es ingeniero químico de la Universidad del Atlántico, cuenta con una maestría en Ingeniería Industrial de la Universidad del Norte. Actualmente trabaja como profesor e investigador en la Universidad del Atlántico. Sus áreas de interés son la medición de la eficiencia operacional a través del análisis envolvente de datos y optimización. <https://orcid.org/0000-0002-4278-1226>