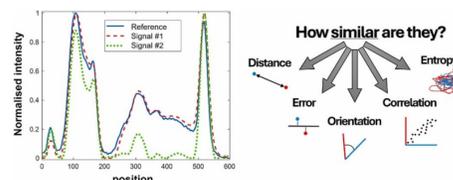


## Redescubriendo el análisis de similitud

19/02/2025

### Publicaciones

El grupo de investigación AnAMA (FQM-232) de la Universidad de Granada ha publicado un tutorial en la revista Trends in Analytical Chemistry que describe y revisa las estrategias para aplicar una poderosa herramienta del análisis de datos químicos: el análisis de similitud.



En la química moderna, la necesidad de extraer información de grandes conjuntos de datos ha impulsado el desarrollo de metodologías matemáticas y estadísticas avanzadas. Entre estas, el análisis de similitud es una herramienta clave que se define como el estudio del grado de semejanza entre dos o más sistemas materiales dados.

En esta publicación se describen los dos enfoques a seguir cuando el análisis de la similitud se aplica sobre cualquier tipo de señal instrumental 2D (espectros, cromatogramas, electroferogramas, voltamogramas, termogramas, curvas cinéticas, etc.): (1) similitud por pares de señales y (2) similitud de conjuntos de señales. El tutorial hace hincapié en el primer enfoque donde se describen las cinco estrategias matemáticas para su cálculo: basado en la distancia, el error, la orientación, la correlación y la entropía (Imagen 1).

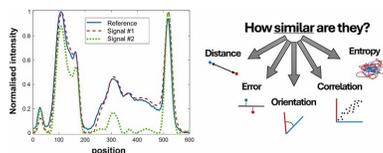


Imagen 1. Resumen gráfico de la publicación

Se revisan diversas aplicaciones en las que se puede hacer uso de este análisis de datos durante cualquiera de las etapas de un desarrollo experimental: (i) para asegurar las condiciones de adquisición adecuadas o evaluar la repetibilidad (o estabilidad) de las medidas; (ii) para identificar, determinar o incluso caracterizar la

muestra de acuerdo con la información química recopilada frente a la esperada (una referencia), aplicación común en diversos campos como la farmacología, la bioquímica y otras industrias químicas; (iii) para la formulación de hipótesis que respalden el desarrollo posterior de un modelo de aprendizaje automático (machine learning) cuyo objetivo puede ser la monitorización de procesos o la autenticación/control de calidad de productos, entre otros. Cualquier campo de la química, e incluso de cualquier otra disciplina que utilice señales instrumentales, podría beneficiarse del análisis de similitud.

La publicación se acompaña de una guía y ejemplos sencillos sobre cómo aplicar las citadas estrategias, así como las ecuaciones necesarias para calcular índices de similitud normalizados entre valores de 0 y 1 (siendo 1 la máxima similitud y 0 la máxima disimilitud), que proporcionan una forma fácil de interpretar y comparar los resultados.

Los investigadores implicados trabajan ahora en una segunda entrega en la que se amplíe esta primera, adaptando el cálculo a señales analíticas más complejas caracterizadas por una mayor dimensionalidad de la estructura de los datos.

El uso de herramientas propias de la inteligencia artificial, como es la minería de datos y métodos de aprendizaje automático, aplicados a datos químicos y conocido como quimiometría, viene siendo habitual por parte de los miembros del citado grupo. El desarrollo de índices de similitud como los presentados en esta publicación es original de los miembros del grupo desde hace más de una década, y a día de hoy continúan tratando de mejorar esta herramienta, en la búsqueda por ofrecer soluciones a los problemas de la actual química analítica moderna.

Referencia:

A. Arroyo-Cerezo, A.M. Jiménez-Carvelo, M., Medina-García, E. Roca-Nasser & L. Cuadros-Rodríguez (2025). Rediscovering similarity analysis of analytical signal: a not well-known mainstay of chemometrics. *Trends in Analytical Chemistry*, 185, 118166.

DOI : [10.1016/j.trac.2025.118166](https://doi.org/10.1016/j.trac.2025.118166)