

A Comprehensive Review of Advancing Pharmaceutical Analysis Technology Through Chemometrics Approaches

Vaishnavi Kadali¹, Prashant Unde^{2*} and Laxmikant Borse³

¹Department of Pharmaceutical Quality Assurance, Sandip Institute of Pharmaceutical Sciences, Nashik, Maharashtra, India.

²Department of Pharmaceutical Chemistry, Sandip Institute of Pharmaceutical Sciences, Nashik, Maharashtra, India.

³Department of Pharmacology, Sandip Institute of Pharmaceutical Sciences, Nashik, Maharashtra, India.

<http://dx.doi.org/10.13005/bbra/3450>

(Received: 12 July 2025; accepted: 13 November 2025)

In today's pharmaceutical world, where advanced instruments like NIR, FTIR, Raman, and HPLC generate huge volumes of complex data, chemometrics has become an essential tool. It combines statistics, mathematics, and computer science to help scientists understand and make use of this data in smarter ways. This review highlights how chemometric techniques are being used to improve pharmaceutical analysis, from raw materials to final products. These tools help overcome challenges in traditional methods, such as overlapping spectra or detecting ingredients present in very low concentrations. Chemometrics is also making routine tasks like calibration curve generation, impurity profiling, and dissolution studies faster, more accurate, and less dependent on complex separations. It supports techniques like UV-Vis, NIR, and NMR spectroscopy, and even helps with optimizing eco-friendly methods in line with green chemistry goals. Overall, chemometrics bridges the gap between complex analytical data and practical decision-making. It's not just about crunching numbers- it's about empowering researchers to ensure the safety, quality, and effectiveness of medicines with greater confidence and efficiency.

Keywords: Chemometrics; Multivariate analysis; Pharmaceutical analysis; Process analytical techniques (PCA); Spectroscopy (FTIR, NMR, UV-Vis).

The pharmaceutical sector is subject to high-quality and regulatory standards to ensure that drug products are safe, effective, and consistent. Analytical chemistry is essential for achieving these goals because it helps identify, measure, and track active pharmaceutical ingredients (APIs), excipients, and contaminants during the drug development process. The volume and complexity of analytical data have expanded

significantly as highly efficient instrumentation has evolved, specifically spectroscopic (e.g., NIR, FTIR, Raman) and chromatographic (e.g., HPLC, GC) techniques. Traditionally, single methods of analysis are frequently insufficient to handle this data complexity.

Chemometrics is a scientific field that uses statistical and mathematical techniques to extract significant data derived from chemical and physical

*Corresponding author E-mail: prashantunde86@gmail.com



processes that take place during the production process. Chemometrics is utilized in statistical process control, process modelling, pattern identification and classification, multivariate data collection and analysis methods, signal rectification and compression, and calibration. To identify, compose, and comprehend the model, descriptive concerns comprise features of the systems being studied. The most advanced automated laboratory equipment used in biological and medical research generates a huge amount of measurement data that is challenging to process and understand. The difficult work of absorbing the data and revealing the relevant information is made easier with the application of chemometrics. Chemometrics is a diverse and abstract field that uses statistical and mathematical approaches, particularly multivariate methods, to describe scientific fields. Algorithms and methods in Information Technology can be used for handling data and assessment. Different fields, including medical, pharmacy, food control, and environmental monitoring. This survey analyses how chemometric approaches affect pharmaceutical analysis.

Spectrophotometric analytical techniques are commonly employed in monitoring food and pharmaceutical product quality during stability testing or batch manufacture. The simplicity of sample preparation and execution, together with the quicker analysis time and lower cost when compared to alternative techniques for analysis, supports this decision. However, methods based on standard spectrophotometry are often unsuccessful when investigating complex mixes due to their limited resolution. Many drugs are multicomponent mixtures that are often challenging to evaluate. Many articles have described the analysis of complicated pharmaceutical combinations using various chemometric approaches on spectrum data. Alternating least squares with the use of multidimensional curve determination to examine complicated mixes & resolve various components in pharmaceutical formulation. Furthermore, typically, it is desirable to decrease down on records when treating complicated data sets chemometrically in order to avoid redundant or worthless information while simultaneously selecting those that contain valuable analytical information. The selection of the most valuable

data frequently affects the multivariate models' capacity to predict outcomes, and this process can be highly helpful.¹

This review is situated within the domain of Pharmaceutical Sciences, concentrating on how chemometric strategies can be integrated with spectroscopic methods such as UV-Vis, NIR, FTIR, and NMR, along with chromatographic tools like HPLC and GC, to advance pharmaceutical analysis.

Brief Introduction to Chemometrics

Origin

The field of chemometrics was founded in the early 1970s. The Swedish chemist S. Wold named a fund project in 1971 after taking into account three ideas: chemometrics, computers in chemistry, and chemical data analysis. When he opted for the latter, the field of chemometrics began to take shape. In Seattle, Washington, he and Professor Kowalski established the International Chemometrics Society (ICS) three years later. Traditional statistical methods served as the main foundation for early chemometrics. For instance, American statistician H. Hotelling developed and popularized the idea of principal component analysis (PCA) in 1933 after British statistician K. Pearson first introduced it in 1901. In the 1960s, Swedish econometric statistician H. Wold developed the well-known Partial Least Squares method for analysing economic data. It was later developed by Hinson S. Olds in the 1980s. In order to increase spectral resolution and decrease interference, Hammond *et al.*² suggested three derivative spectrophotometry as early as 1953. This technique is currently frequently employed in molecular spectroscopies. Chemometrics reached its zenith in the 1980s. Computer popularity, industrial interests, and the advancement of analytical tools all contributed to the exceptional depth and scope of chemometrics research. In actuality, many of the most cutting-edge techniques that are currently in widespread usage were developed or refined during that period. Chemometrics development may often be broken down into four phases.

Development History

Prior to the establishment

The use of statistical statistics in chemistry, particularly analytical chemistry, is what distinguishes this stage. The linear free energy structure-activity relationship, which is regarded

as the forerunner of the QSAR (Quantitative Structure-Activity Relationship), was investigated by organic chemists.

The birth of chemometrics

Analysts customized sorting, investigation, forecasting, and processing of information, procedures for chemistry-specific needs. Chemometrics has become a significant part of quantitative analysis. That evolution involves two aspects. Computers have become increasingly common, particularly in analytical chemistry, providing scientists with accurate and dependable data. The development of chemometrics stemmed from the need to efficiently turn instrument data into valuable information. The second is that, with the aid of quicker computing, a variety of potent mathematical techniques can be used in analytical chemistry. Chemometrics' ascent can be seen as the primary example of how contemporary technology advancements in chemistry are reflected in computer applications.

Progressive 1980s

Theoretical and algorithmic research has significantly advanced the unique methods for chemical pattern recognition, multidimensional discrimination, and multivariate calibration, such as evolving factor analysis and rank annihilation factor analysis. Along with many of the classic monographs on chemometrics, "Journal of Chemometrics" & "Chemometrics and Intelligence Laboratory Systems" had been founded at this time. These Publications were crucial in establishing development trends, directing scientific study themes, and spreading knowledge of chemometrics. The American Mathwork Company formally introduced MATLAB software in 1984. It allows for the realization of numerous intricate mathematical computations used in chemometrics using. It is essentially a standard programming language for chemometrics research because it only requires one coding expression. MATLAB scripts were typically coupled to newly published algorithms, which significantly aided in the discipline's advancement.

By the 1990s, chemometrics had really advanced to the point of useful applications in fields including sensors, medicine and pharmacy, near-infrared spectroscopy, etc. The chemometrics software was installed on a computer or microprocessor in practically every contemporary analytical

device. Chemometrics was quickly emerging as a vital tool for analytical or daily chemistry work. Additionally, a number of novel techniques were used by analysts as new instruments for resolving chemical issues, including support vector machines, wavelet transforms, genetic algorithms, and artificial neural networks.³

Chemometrics is, in many ways, a product of the "information age," computers, and statistics. In the previous three decades, the discipline of Chemometrics has grown phenomenally due to rapid technical advancements, particularly in the field associated with analytical chemistry's computerized tools.⁴

Chemometrics: - What is it?

Chemometrics is the study of chemical (and biological) measurements and is a subfield of analytical chemistry. This definition is clear from the term, where chemo- means chemical and metrics means measurement. Chemometric applications include,

- ensuring that the data we gather is relevant to our goals
- improving the quality of an analytical signal by reducing the contribution of noise
- reporting on an experiment in a manner that gauges the degree of uncertainty in its findings and our confidence in those findings.
- creating practical models that forecast the results of subsequent experiments. Identifying underlying patterns in chemical data to extract information that is concealed yet analytically valuable.⁵

chemometrics, we can employ them to solve predictive problems, like forecasting desired features or target properties. Additionally, it can be applied to descriptive challenges such as model design, recognition, and comprehension. Chemometrics is used in the collection and analysis of multivariate data. The data can be processed and evaluated using a variety of algorithms and comparable techniques. They can be applied to a number of industries, including environmental surveillance, nutrition management, pharmaceuticals, and medical treatment. The following lists a few chemometric algorithms for analysis.⁶

Different aspects of chemometrics

Chemometric experiments should be conducted with all experimental variables varied, which can be accomplished through the use of experimental

design. To gain more and deeper information
The Chemometric Principle

1. To integrate, rationalize, and explain chemical data gained during study.
2. Incorporated Divergence in the algorithm and address it using distribution.
3. When preparing sets of trials to modify conditions or optimize, always use statistical designs rather than changing only one component while leaving the others constant.
4. employ multiple-variate assessment techniques and display the findings in charts.⁷

Bilinear model

This model arranges data in matrices, with variables in vertical columns and samples in horizontal rows.⁹ The following categories apply to bilinear models:

PCA or Principal Component Analysis

Principal component analysis is a straightforward, nonparametric technique for data extraction, pattern recognition, and data presentation that emphasizes differences and similarities. In a number of scientific fields, it reduces scale and investigates data with multiple variables compression. Because of its many uses in resolving multivariate issues, it is the most widely used multivariate approach. PCA lowers a process's number of variables. by creating a correlation structure for the variables and analysing changes in the correlations between them in process monitoring. To describe the same level of variability, the data will be processed via PCA.¹⁰

PLS or Partial Least Square

Partial least square is a well-known technique for modelling relationships between observed data using latent variables.¹¹ used PLS to project observable data to latent structures. This method uses regression, dimension reduction, and modelling methods to modify the relationship between observed and latent variables. Latent variables tend to make sets of statistical parameters more covariant with one another. Partial Least Square, like CCA, can be used to discriminate and reduce dimensions, similar to PCA.¹² These techniques can be combined into a single, cohesive strategy called continuum regression.¹³ Because it can handle large volumes of chemical data, the Partial Least Squares approach is widely used in the field of chemometrics. The partial least squares method evolved ascertain the flow characteristics of medicinal powders.¹⁴

Multway model

In general, multivariate strategies were able to supply more data than bilinear techniques could. Multiway principal component analysis is one of the techniques acknowledged as instruments for monitoring because they enhance process comprehension and provide a batch-wise summary of its behaviour. It is frequently used to extract data from spectra, which is typically challenging when there is overlap. The preferred multiway techniques are those that use three or more ways.¹⁶⁻¹⁷

Classical Least Square, or CLS¹⁸

subjected to Multiple Linear Regression:

$$A = KC$$

Or

$$dA/d\lambda = KC$$

where,

A - zero-order absorbance

dA/dλ – derivative of absorbance matrices C-

concentration of matrix

K- calibration coefficient of matrix

Only systems where all of the sample's constituents are known can use this technique. When contaminants that were absent from the calibration mixtures may be present in the unknown sample,

Inverse Least Square, or ILS

ILS is sometimes called P-matrix calibration because it applies Multiple Linear Regression (MLR) to the inverse expression of the Beer-Lambert law of spectroscopy.

$$C = PA \text{ or } (C = P \times dA/d\lambda)$$

where,

A-Zero-order absorbance

dA/dλ- derivative of absorbance matrices C-

concentration coefficient of matrix

P-calibration coefficient of matrix

This approach works well for more intricate studies that CLS is unable to handle. The ILS method's drawbacks include the potential for challenging and time-consuming wavelength selection.¹⁹

Parallel Factor Analysis (PARAFAC)

Cattell (1944) proposed "parallel proportional profiles" as the main concept for choosing rotation in component analysis, after reviewing seven other principles. A technique for

analysing multiple data matrices with rating based on same individual on the same factors was created by Harshman (1970) and is called PARAFAC.²⁰

Parallel Factor Analysis [PARFAC-2]

Profiles of data attributes that have evolved or are in a new stage can be handled by PARAFAC-2. Trilinearity is a basic need of PARAFAC, while PARAFAC-2 makes it possible. It should be mentioned, nevertheless, that PARAFAC may only be utilized when data deviations from linearity are regular in order to partially fit nonlinearity in a single mode.^{18,21-22}

The Tucker - 3 model

Simply because the Tucker-3 model has multiple loading matrices, it can be used to analyse n-way array data. The model was put forth in 1966 by psychometrician Ledyard R. Tucker. Due to its universality and the inclusion of the PARAFAC model as a unique instance, the Tucker-

3 model is widely utilized in many applications for decomposition, compression, and interpretation.²³

N - Partial Least Square

The N-Partial Least Square method was developed to deal with multiway data extensions of the PLS approach, which finds the latent variables for representing maximal covariance using variables that are both independent and dependent.

N-PLS decomposition occurs by optimizing the covariance between the two matrices and developing a unique model for the contingent response variables that resembles PARAFAC.²⁴

Locally Weighted Regression

The locally weighted regression uses conjunction to assign a larger weight to calibration samples that are most similar to the sample that needs to be forecasted. Numerous variations could exist, and the one with PLS for every sample with equal weights should be the favored one. Thus, the local approach turns into a global approach. Although LWR was not as thoroughly explored as the previous approaches, it was demonstrated to function effectively when evaluating nonlinear or clustered data; as a result, it is less recommended and has fewer diagnostics accessible.²⁵

Chemometrics in Pharmaceuticals Analysis

To assess the pharmaceutical aspects of powders, granules, and tablets, spectral methods like UV, FTIR, etc., are often utilized with chemometric models like multidimensional analysis, PCR, PLS, and CLS. The most commonly utilized technology is NIR spectroscopy.²⁶

Assessing Spectroscopy During Integration with Chemometrics Calibration curve analysis

In recent years, spectroscopy and chemometrics analysis techniques primarily

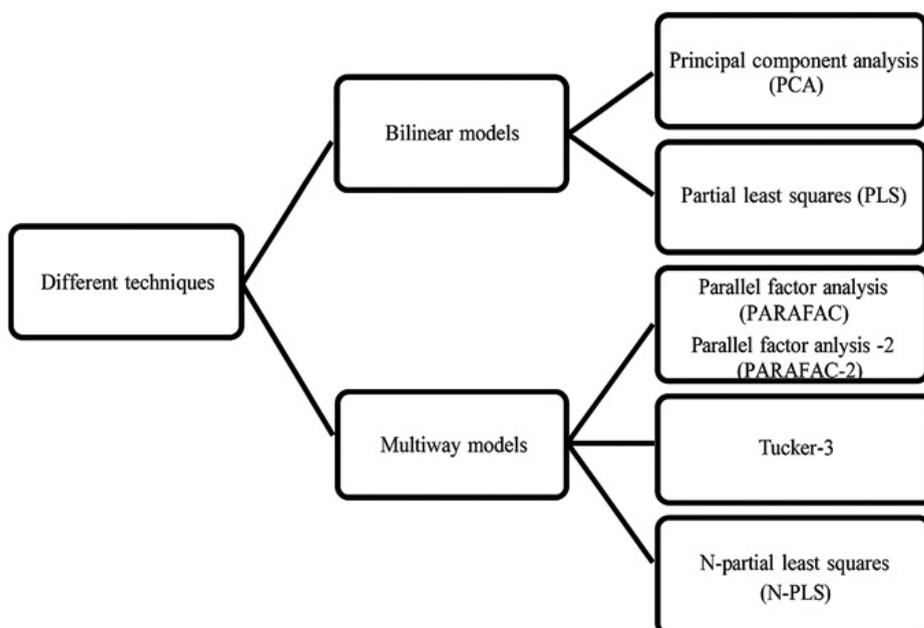


Fig. 1. Chemometric Techniques⁸

employ the same methodology, which is a collection of known samples known as calibration samples or training samples. A calibration or recognition model is created using the spectra of these references and the samples' data that goes with them. Measuring the sample's spectrum is all that is required for testing, and the developed model will yield quantitative or qualitative results.³

Every quantitative analytical method requires calibration, and carrying out this step correctly is crucial to the validation and development of the method. All of the method's results could contain a crude error due to poor statistical analysis or design. Advanced statistical tools are now accessible thanks to the rapid advancement of open-source software and computational technology. As a result, there are now more requirements for statistical evaluation of calibration. It helps detect and correct issues like curvilinearity, heteroscedasticity, outliers, and

non-normal residuals, which can otherwise lead to major errors. Chemometric tools enable better model selection, error estimation, and overall method validation, making calibration more robust and scientifically valid.²⁷

Visible Spectrophotometry or UV

UV is a cost-effective and efficient method for testing pharmaceuticals with UV-absorbing components. UV-Visible absorption lacks specificity, making it ineffective when bands from distinct components overlap. UV-Visible spectrophotometry can now be used to analyse complicated mixtures without the requirement for separation, thanks to advancements in MLR and factor-based approaches.²⁸

NIR Spectroscopy

NIR spectroscopy is widely used in various industries, it provides both qualitative and quantitative analysis.³⁰ A commonly used analytical method is near-infrared spectroscopy (NIR).

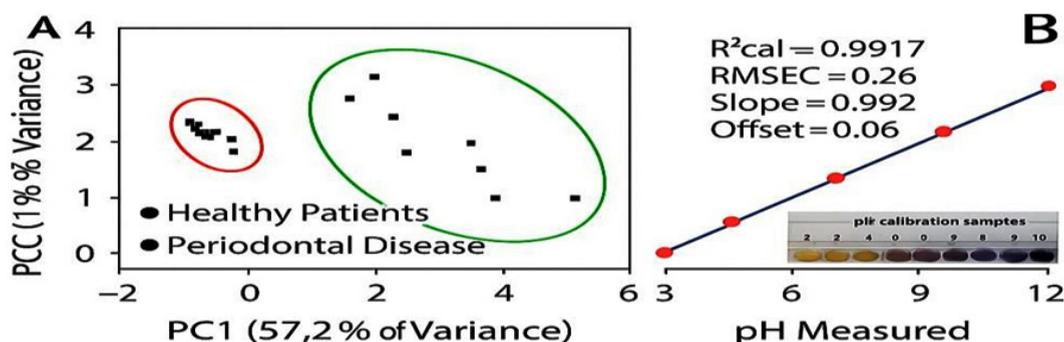


Fig. 2. A) The PCA score plot shows how the behavior of healthy patients (circle) and those with periodontal disease (square) differs. (B) PLS correlations using a portable electronic nose equipped with ten metal oxide sensors between the measured and predicted saffron samples

Table 1. Summary of Comparative Method Performance

Method	Primary Use	Advantages	Limitations
PCA	Exploratory data analysis	Fast, easy to interpret	Limited to linear variance
PLS / PCR	Quantitative prediction	Handles collinearity, predictive	Sensitive to overfitting
MCR-ALS	Curve resolution	Provides interpretable pure spectra	Needs component number input
PARAFAC / PARAFAC2	Multiway data analysis	Unique decomposition, preserves structure	Computationally intensive
SVM / ANN / CNN	Nonlinear modeling	High accuracy, flexible	Requires large datasets, less interpretable

method that uses less material for examination. Associated with chemometrics, NIR spectra can be analysed using multivariate techniques to retrieve analytical information. Multivariate analysis is also applied to qualitative analysis.³¹

NIR spectra of the various meloxicam powder mixes were created and examined. chemometrics and visible and near-infrared spectroscopy to determine the orange juice's pH and soluble solids concentration. 104 orange juice samples were collected, pre - processed, and their spectra were captured using a wavelet packet. Partial least squares regression analytical chemometrics was chosen for processing spectral data, and further measurements of orange juice's pH and SSC demonstrated that chemometrics and NIRS improve data analysis assessment. A strong and dependable method for developing pharmaceutical products and guaranteeing the quality of the finished product is near-infrared spectroscopy. This technique's capacity to capture a large amount of spectral information quickly is one of its key features.³²

The spatial and qualitative data regarding the substance utilized in pharmaceutical formulations might be analysed using the multivariate curve resolution model and the classical least squares approach. Here chemometric applications provide a useful way to characterize and estimate drug and carrier.³³

FT-IR

A brand-new FTIR spectroscopic method with chemometric support was developed and validated. The suggested FTIR method does not require reagents and uses fewer solvents. The developed approach was simple, cost-effective, and precise, and it met the majority of validation criteria in a concentration range appropriate for monitoring the quality of both pure and solid dosage forms.³⁴ Tablets were analysed using the developed chemometrics-assisted FTIR and Raman spectroscopy techniques in contrast to HPLC. PQ and DHA in tablets were calculated using suggested models based on the PLSR, which made use of latent factors and optimal wavenumbers. The results were compared statistically using the t-test

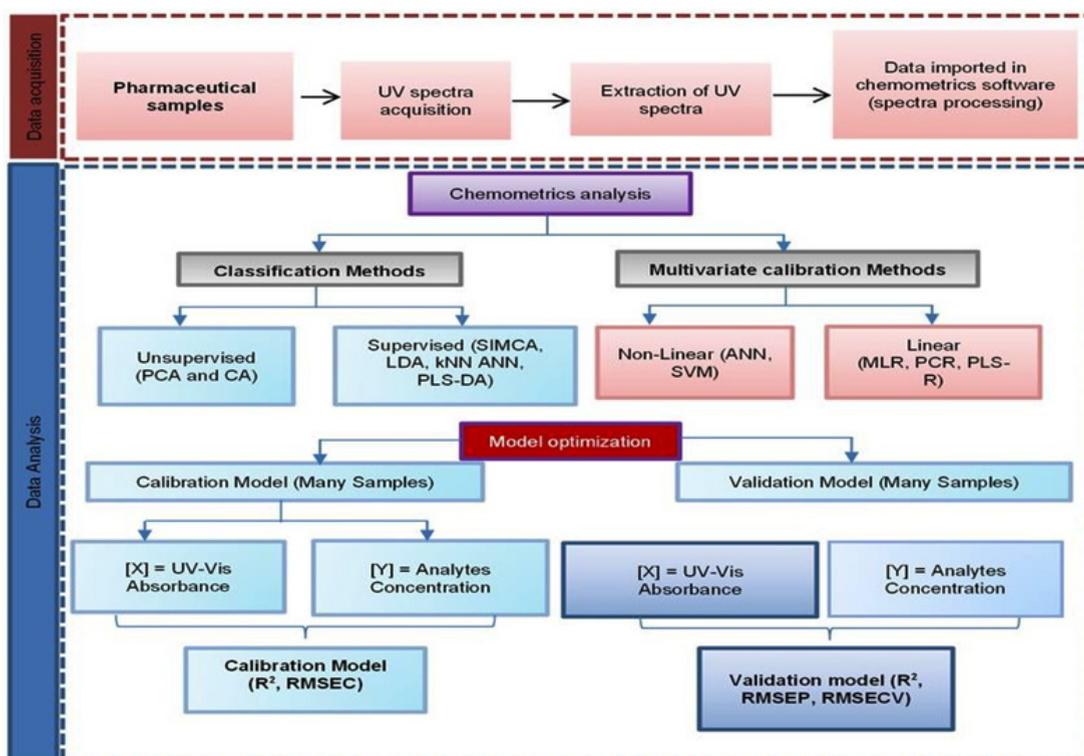


Fig. 3. A diagrammatic illustration of the implementation of the chemometrics approach using the UV-Vis spectrum.²⁹

from each method at ninety-five percent confidence levels.³⁵

NMR

rapidly increasing as a result of the growing interest in analytical NMR spectroscopy, in 1983, chemometrics was applied to NMR spectra for the first time. PCA was introduced in the early 1990s. the field of study known as “metabolomics,” which is defined as “metabolic processes studied by NMR spectroscopy of biofluids” or “understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data.” The extremely complicated metabolic system in bodily fluids, which results in similarly complex NMR spectra, was undoubtedly the driving force.³⁶

NMR signals are gathered over time, and the chemical shift spectrum is generated using a Fourier transformation. Quantitative information in spectra depends on external magnetic field quality, precise tuning, temperature, spectral preprocessing, including line broadening, zero filling, and phasing. Applying chemometrics or empirical research to huge NMR data sets poses a difficulty for the analytical NMR platform’s long-term stability of parameters.³⁷

Chemometric in LC (Liquid Chromatography) Separations Chemometrics for Aiding the Analysis of Sustainable APIs

The contribution of chemometrics to the improvement of LC methods, particularly

with regard to boosting greenness, has been covered in numerous published references. For example, an LC technique has been developed through experimental design for the separation of glibenclamide, gliclazide, glipizide, glimepiride, and metformin. Numerous chemometrics methods address issues that arise during LC procedures. Without first separating the active ingredients from the commercially available tablet formulations, the spectrometric content was proceeded using the PCR, CLS, MLR, and PCR techniques. Selectivity, sensitivity, LOD, LOQ, and analytical sensitivity were identified as analytical figures of merit. Following screening, four important parameters were identified in HPLC-QbD, which was employed for efficiency improvement.

Chemometrics for Aiding Longevity Impurity Characterization

Impurity profiling is a vital method used by the QC department in the healthcare sector. Through the examination of both inorganic and organic contaminants present in an active pharmaceutical ingredient, producers guarantee the goods’ safety, effectiveness, and uniformity. Additionally, such a tool provides information on the limitations of impurity identification and quantification in APIs and/or final products. Pharmaceutical contaminants can be separated, examined, and identified. However, the application of UV-visible spectrophotometry is cost-effective and advantageous for the environment. The chemometric method is unquestionably required

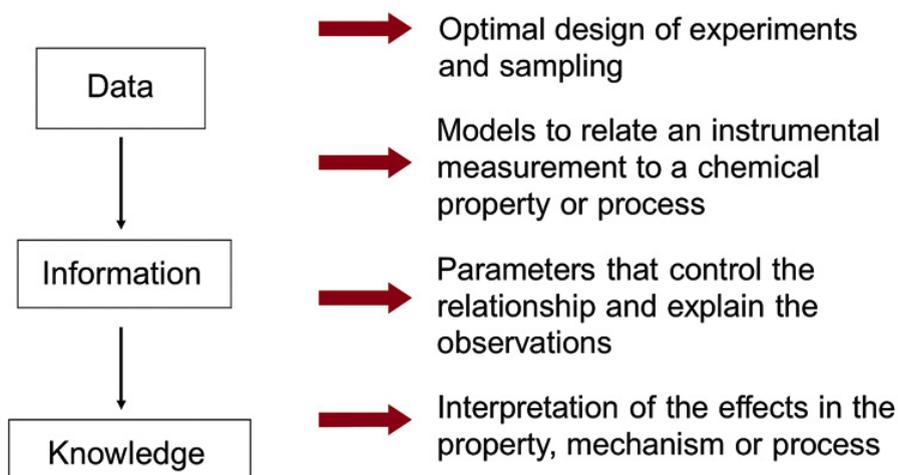


Fig. 4. Chemometric Process⁴³

for the characterization of pharmaceutical contaminants. PCA has been used to quickly sort simvastatin tablet batches to determine their impurity profiles.³⁸

Chemometric-Aided Quantification of Synthetic and Marketed Drug Formulations

Multivariate calibration is a useful technique for analysing multicomponent mixtures because it eliminates the need for draw-out separation processes. Thanks to modern instrumentation for acquiring and digitizing spectral information and powerful computers for processing vast amounts of data.

The drug dicyclomine hydrochloride (DICY) has anticholinergic and antispasmodic properties. DICY has antispasmodic properties and is used to treat irritable bowel syndrome, an intestinal condition. Clidinium bromide is an anticholinergic and antisecretory substance that works by preventing parasympathetic innervations from working, which lowers stomach acid production. It also has minor antispasmodic properties. One kind of benzodiazepine is chlordiazepoxide.

Additionally, it has GABA facilitator activity. It has skeletal muscle relaxants, hypnotics, sedatives, and anxiolytic properties. For these ternary combinations, there are various commercially available formulations, such as tablets of Arvin Fort, Equital, Curemaxine, and Normalxine. Peptic ulcers, gastritis, irritable spastic colon, mucous colitis, and nervous dyspepsia are all treated with this highly successful medication combination.

Chlordiazepoxide, Dicyclomine hydrochloride, and Clidinium bromide have all been evaluated separately or in combination with other medications using a few analytical techniques. According to the US Pharmacopoeia, chlordiazepoxide and clidinium bromide are evaluated using a non-aqueous technique. A few further procedures for estimating the amounts of chlordiazepoxide and clidinium bromide in mixed dose forms have been published, including spectrophotometry and derivative spectroscopy using multivariate calibration methods and UV/VIS spectrophotometry, HPLC derivative spectrophotometry, flow-injection potentiometry, and voltammetry. There are very few techniques available for estimating dicyclomine hydrochloride

and its combination with chlordiazepoxide and cyclinium bromide.³⁹

Solid State Analysis

The pharmaceutical industry and academia have recently seen dosage forms change from being merely scientific nuances and curiosities to issues that require serious and in-depth attention. The same concerns must be examined in formulations, which also need to look into the impact of excipients. The latter's complexity, however, foreshadows that it would be far more challenging to analyse specific chemical species in dose forms.

A well-defined lattice is formed by the regular repetition of structural units found in crystalline materials. Without changing their chemical makeup, pharmaceutical solids can have two or more distinct structural orientations that result in various crystalline solid forms, each with unique physical properties. This well-known and acknowledged phenomenon is called structural or crystal polymorphism, and it is typically referred to as "polymorphism" in the relevant literature.

Unexpected polymorphic transformations can lead to major pharmaceutical problems at the industrial level, which could ultimately cause development delays, production halts, or the cancellation of commercialization. Therefore, in order to guarantee the quality, safety, and performance of their products, pharmaceutical businesses are required by current legislation to look into and control medication substance polymorphism. Unexpected polymorphic transformations can lead to major pharmaceutical problems at the industry level, which can ultimately cause delays in development, halts in manufacturing, or the cancellation of commercialization. Therefore, in order to guarantee the quality, safety, and performance of their products, pharmaceutical businesses are required by current legislation to look into and control medication substance polymorphism.⁴⁰

Electrochemical Method Optimization

As compared to spectroscopic approaches, its usage in electroanalytical chemistry is still in its infancy. The connection between mathematics and electroanalytical chemistry is linked to the sparse application of chemometrics in this field. The basic corpus in this instance consists of:

(i) a theoretical physicochemical depiction of the processes, transport phenomena, and measurement

characteristics; (ii) the numerical resolution of the mathematical formulation; and (iii) calculation of concentrations, constants, or anything else. This method, which is commonly referred to as “hard modelling” by electrochemists and used in electrochemical research, is thought to be the authentic methodology. It is challenging to postulate a theoretical physicochemical model due to the complexity of the electrode process, the transport phenomena, and excitation signal perturbation.⁴¹

The analysis findings of validation samples were evaluated in order to validate the optimized voltametric method. The suggested technique was then used to quantify cefdinir in three distinct pharmaceutical preparations: powder for oral solution, effervescent tablets, and film-coated tablet.⁴²

Dissolution studies

The dissolution behavior of hydrochlorothiazide (HCT) and bisoprolol fumarate (BIS) tablets was evaluated using USP Apparatus II (paddle method) under standard conditions in 0.1 N HCl. An on-line circulation system coupled with a UV-visible detection unit enabled continuous monitoring of the dissolution process. The dissolution medium was pumped through a UV flow cell using a peristaltic pump, and spectral data were recorded every 30 seconds across the 210–350 nm wavelength range.

A second-order multivariate chemometric approach, as described by Maggio *et al.* (2013),⁴⁴ was applied to resolve the overlapping UV spectra of the two active ingredients. In their study, Maggio and colleagues emphasized that second-order chemometric methods had previously been used in only two reports, both of which analyzed single-drug systems. i. Wiberg K.H., Hultin U.-K. (2006)⁴⁵ This study applied second-order chemometric modeling to dissolution testing using fiber-optic UV data, but only for a single active drug. ii. Rajkó R., Nassab P.R., Szabó-Révész P. (2009)⁴⁶ This work also used second-order chemometric analysis (self-modeling curve resolution) but focused solely on one drug component (meloxicam) in a binary mixture.

Unlike these earlier studies, Maggio *et al.* developed a second-order multivariate curve resolution-alternating least squares (MCR-ALS)

model capable of simultaneously quantifying HCT and BIS in real time. The MCR-ALS algorithm accurately separated the contributions of each analyte, despite strong spectral overlap. Method validation, following ICH Q2 (R1) guidelines, confirmed excellent linearity ($R^2 > 0.999$), precision, and robustness, with results closely matching those obtained by reference HPLC analysis.

This approach demonstrates a reliable, rapid, and environmentally efficient technique for monitoring the dissolution profiles of multicomponent pharmaceutical formulations under continuous flow conditions.

AI and Chemometrics Integration in Spectroscopy

The merging of artificial intelligence (AI) with chemometric techniques is transforming how spectral data is processed and interpreted. Chemometrics traditionally relies on multivariate statistics to extract information from complex datasets, whereas AI contributes automated pattern recognition, adaptive learning, and high-level feature extraction. In pharmaceutical Raman and infrared spectroscopy, AI models improve the accuracy of classification, signal denoising, and anomaly detection, surpassing many conventional multivariate methods. This partnership moves analytical chemistry toward self-learning and real-time decision-making systems, extending the role of chemometrics from statistical interpretation to intelligent data analytics

Chemometrics as the Foundational Framework for AI-Driven Analysis

Chemometrics forms the conceptual and computational basis for today's AI-assisted analytical strategies. It introduced the principles of multivariate modelling, dimensionality reduction, and data-pattern interpretation that underpin machine-learning algorithms. As noted by Mantsch (2021), chemometrics can be viewed as the first generation of data-mining science in vibrational spectroscopy, paving the way for AI's nonlinear and large-scale analytical capabilities. Within pharmaceutical research, AI builds on these fundamentals to produce more adaptive and predictive models, capable of handling the growing complexity of spectroscopic datasets.⁴⁷

Gap Analysis

Despite substantial progress, several gaps remain in applying chemometrics to pharmaceutical analysis. Most studies still focus on traditional models such as PCA, PLS, and MCR-ALS, with limited attention to model optimization, real-time adaptability, and cross-instrument standardization. The integration of chemometrics with artificial intelligence (AI) and machine learning (ML) is underdeveloped, restricting the ability to manage nonlinear and high-dimensional data for predictive and automated process control. A lack of standardized protocols for data preprocessing, validation, and benchmarking also hampers reproducibility and regulatory acceptance. Furthermore, chemometric applications remain largely confined to offline laboratory analysis, with minimal real-time integration into Process Analytical Technology (PAT) frameworks. Insufficient interdisciplinary training often leads to misuse or misinterpretation of models, while limited studies on electroanalytical and hyphenated techniques such as LC-MS and GC-MS highlight the need for broader methodological expansion. Addressing these issues through AI integration, automation, and harmonized validation standards will be essential for achieving intelligent and regulatory-compliant analytical systems in the pharmaceutical industry.

DISCUSSION

The study emphasizes the growing importance of chemometrics in pharmaceutical analysis, particularly its ability to interpret complex data from techniques such as Raman, NIR, and FTIR spectroscopy. Among the compared models, MCR-ALS and PMF showed superior accuracy in identifying and mapping chemical components, while SMMA effectively estimated the number of ingredients. These results confirm the reliability of chemometric tools for both qualitative and quantitative evaluations. The integration of artificial intelligence (AI) and machine learning (ML) into chemometric workflows offers further potential for handling nonlinear and high-dimensional datasets. To achieve consistent and regulatory-compliant outcomes, future work should focus on standardized validation, real-time

model adaptation, and cross-platform calibration. Overall, combining chemometrics with AI-driven spectroscopy represents a crucial step toward intelligent, automated, and efficient pharmaceutical quality control.

Future direction

Future advances in chemometrics for pharmaceutical analysis will center on integrating AI and deep learning for better prediction and real-time process control, particularly within Process Analytical Technology (PAT) frameworks. The creation of mathematical models aided by artificial intelligence (AI) is another aspect of future directions for chemometric study. There have been reports of supervised models created using machine learning methods. The accuracy and sensitivity of this method in evaluating intricate drug formulations, such those used in traditional medicine, may potentially be improved by advanced modelling.⁴⁸

Advanced multiway and nonlinear models will help analyze data from complicated, high-dimensional datasets. Chemometrics will also help with green analytical chemistry, solid-state characterisation, and customized medicine via metabolomic profiling. As huge data from hyphenated approaches expands, cloud-based and open-source chemometric platforms will become more accessible. Furthermore, its contribution to Quality by Design (QbD) and regulatory standards will grow, assuring efficient, environmentally friendly, and compliant pharmaceutical processes.

CONCLUSION

In pharmaceutical analysis, chemometrics has proven to be a strong and essential tool, enabling the extraction of meaningful insights from complex spectroscopic and chromatographic data. Its application improves accuracy, simplifies multicomponent analysis, and supports various analytical tasks including impurity profiling, dissolution studies, and solid-state characterization. When integrated with techniques like NIR, FTIR, UV-Vis, and NMR, it enhances method development, validation, and real-time quality control. Overall, chemometrics is crucial for achieving efficient, precise, and regulatory-compliant pharmaceutical analysis.

ACKNOWLEDGEMENT

We thank the management of Sandip Institute of Pharmaceutical Sciences for their continuous support and encouragement.

Funding Sources

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflict of interest

The authors do not have any conflict of interest.

Data Availability Statement

This statement does not apply to this article.

Ethics Statement

This research did not involve human participants, animal subjects, or any material that requires ethical approval.

Informed Consent Statement

This study did not involve human participants, and therefore, informed consent was not required.

Clinical Trial Registration

This research does not involve any clinical trials.

Permission to reproduce material from other sources

Not Applicable.

Author contributions

Prashant Unde: Conceptualization, Methodology, Administration & Supervision; Vaishnavi Kadali: Writing – Original Draft, Visualization; Laxmikant Borse: Analysis, Writing, Supervision, Review & Editing.

REFERENCES

- Michele D. L., Giuseppina I., Claudia S., Gaetano R. Optimization of wavelength range and data interval in chemometric analysis of complex pharmaceutical mixtures *Journal of Pharmaceutical Analysis*. 2016;6(1):64-6 doi:10.1016/j.jpha.2015.10.001
- Hammond, V. J., Price W. C. Derivative spectrophotometry. *Journal of the Optical Society of America*, 1953;43(10):924–929. DOI: 10.1364/JOSA.43.000924
- Bakeev K. A. Chemometric in Process Analytical Technology In: Bakeev KA, ed. *Process Analytical Technology*. Vol 1. 2nd ed. Hoboken, NJ: John Wiley & Sons Ltd; 2010:353-360. <https://onlinelibrary.wiley.com/doi/10.1002/9780470689592.ch12>.
- Gemperline P. Introduction. In: Gemperline P, ed. *Practical Guide to Chemometrics*. 2nd ed. Boca Raton, FL: CRC Press; 2006:2-3. <https://doi.org/10.1201/9781420018301>
- Harvey D. T. Introduction. In: Harvey DT, ed. *Chemometrics Using R*. LibreTexts; 2021:1. [https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Chemometrics_Using_R_\(Harvey\)](https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Chemometrics_Using_R_(Harvey)). doi:10.1201/9781420018301.
- Pawar H.A., Kamat S.R., Patel K.N. Chemometrics and its application in pharmaceutical field. *Journal of Physical Chemistry and Biophysics*. 2014;4(6):1. doi:10.4172/2161-0398.1000169.
- Prajapati P. R., Patel N. R., Patel R. P., Patel K. N. Chemometrics and its Applications in UV Spectrophotometry. *Int J Pharm Chem Anal*. 2016;3(1):43-48. doi:10.5958/2394-2797.2016.00005.8
- Bhalodia D. N., Baria D. A. A Review on Chemometrics in Pharmaceutical Analysis. *Int. J. of Pharm. Sci*. 2024; 2(9): 58-78 DOI: 10.5281/zenodo.13625694
- Matero S. Chemometrics Methods in Pharmaceutical Tablet Development and Manufacturing Unit Operations. Publications of the *University of Eastern Finland Dissertations in Health Sciences*. 2010;16:31-55 <https://erepo.uef.fi/bitstreams/77fcf9c2-b5e2-4562-b858-35218ba6ff84/download>.
- Scotti L., Ferreira E. I., Silva M., Scotti M. Chemometric studies on natural products as potential inhibitors of the NADH oxidase from *Trypanosoma cruzi* using the volsurf approach. *Molecules* 2010;(15): 7363–7377. doi:10.3390/molecules15107363
- Wold S., Ruhe H., Wold H., Dunn W. J. The co linearity problem in linear regression: the partial least squares (PLS) approach to generalized inverse. *SIAM Journal on Scientific Computing*, 1984; (5):735–743. doi:10.1137/0905052
- Barker M., Rayens W., Partial least squares for discrimination. *Journal of Chemometrics*, 2003: 17(3):166–173. <https://doi.org/10.1002/cem.785>
- Stone M., Brooks R. J., Continuum regression: cross validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression. *Journal of the Royal Statistical Society*, 1990;(52):237–269. doi:10.1111/j.2517-6161.1990.tb01786.x
- Sarraguca M. C., Cruz A.V., Soares S. O., Amaral H.R., Costa P.C. Determination of flow properties of pharmaceutical powders by near infrared spectroscopy. *J Pharm Biomed Anal* 2010;(52):

- 484-492. doi:10.1016/j.jpba.2010.01.038.
15. Tortorella S, Cinti S. How can chemometrics support the development of point-of-need devices? *Anal Chem.* 2021;93(5):2713–2722. doi:10.1021/acs.analchem.0c04151
 16. Bro R. Multiway calibration multilinear PLS. *Journal of Chemometrics.*1996; (10): 47–61. doi:10.1002/(SICI)1099-128X(199601)10:1<47::AID-CEM400>3.0.CO;2-C.
 17. Bro R., Kiers H. A new efficient method for determining the number of components in PARAFAC models. *J Chemometrics.* 2003;(17): 274–286. doi:10.1002/cem.801.
 18. Kiers H. A., Berge J. M., Ten Berge J.M.F., Bro R. PARAFAC2. Part I: a direct fitting algorithm for the PARAFAC2 model. *Journal of Chemometrics.* 1999(13): 3-4. 275–294. doi:10.1002/(SICI)1099-128X(199905/08)13:3/4<275::AID-CEM543>3.0.CO;2-2.
 19. Kramer R. Chemometric techniques in quantitative analysis. In: Marcel Dekker, ed. *Chemometrics in Analytical Chemistry.* 2nd ed. New York, NY: Marcel Dekker; 1998:51-97. <https://www.chemometrics.com/books/book001.html>
 20. Smilde A, Bro R. Multi-way analysis with applications in the chemical sciences. In: John Wiley & Sons, ed. *Multi-Way Analysis with Applications in the Chemical Sciences.* Vol 3. New York, NY: John Wiley & Sons; 2005:57-88. <https://onlinelibrary.wiley.com/doi/book/10.1002/0470863622>
 21. Bro R., Andersson C.A., Kiers H. PARAFAC2. Part II: modelling chromatographic data with retention time shifts. *Journal of Chemometrics.*1999;(13):295–309. doi:10.1002/(SICI)1099-128X(199905/08)13:3/4<295::AID-CEM550>3.0.CO;2-P
 22. Kiers H., Ten B., J.M.F., Bro R. PARAFAC2. Part I: a direct fitting algorithm for the PARAFAC2 model. *Journal of Chemometrics.*1999; (13): 275–294. doi:10.1002/(SICI)1099-128X(199905/08)13:3/4<275::AID-CEM549>3.0.CO;2-9
 23. Smilde A. Comments on three-way analyses used for batch process data. *Journal of Chemometrics.*2001; (15):19–27. doi:10.1002/1099-128X(200101)15:1<19::AID-CEM599>3.0.CO;2-F.
 24. Olivieri A. C. Analytical advantages of multivariate data processing one, two, three, infinity?" *Analytical Chemistry.* 2008; 80(15); pp. 5713–5720. doi:10.1021/ac800692c.
 25. Bhavana V., Srinivasa Rao Y. A review on chemometrics in pharmaceutical analysis. *Int J Sci Res.* 2020;9(8):1010-1016. <https://www.ijsr.net/getabstract.php?paperid=SR20808115241>
 26. Pawar H., Kama S. Chemometrics and its Application in Pharmaceutical Field. *J Phys Chem Biophys.* 2014;4(6):169. doi:10.4172/2161-0398.1000169.
 27. Lukasz K. Chemometric and Statistical Evaluation of Calibration Curves in *Pharmaceutical Analysis Journal of AOAC International* 2012; 95(3):669-672. doi:10.5740/jaoacint.SGE_Komsta.
 28. Bhavana, Srinivasa R.Y. *International Journal of Science and Research (IJSR)* ISSN:2020;9(8): 2319-7064. <https://www.ijsr.net/archive/v9i8/SR20808115241.pdf>
 29. Nurani L. H., Edityaningrum C.A., Irnawati I., Putri A. R., Guntarti A., Rohman A. Chemometrics-assisted UV-Vis spectrophotometry for quality control of pharmaceuticals: a review. *Indones J Chem.* 2023;23(2):542–567. doi:10.22146/ijc.74329.
 30. Gemperline P. *Practical Guide to Chemometrics.* 2nd ed. Boca Raton, FL: CRC Press; 2006:107. ISBN-13: 978-1574447835. doi:10.1201/9781420018301.
 31. Pino T. Development and assessment of spectroscopy methodologies and chemometrics strategies to detect pharmaceuticals blend endpoint in a pharmaceutical powder blend. *Journal of the Chilean Chemical Society.*2021 66(4):5387-5397. doi:10.4067/S0717-97072021000405387.
 32. Gu, C., Xiang, B., Su, Y. and Xu, J. Near-Infrared Spectroscopy Coupled with Kernel Partial Least Squares-Discriminant Analysis for Rapid Screening Water Containing Malathion. *American Journal of Analytical Chemistry.*2013;(4):111-116 doi:10.4236/ajac.2013.43015.
 33. Singh I., Juneja P., Kaur B., Kumar P. Pharmaceutical applications of chemometric techniques. *ISRN Anal Chem.* 2013;2013:1–13. doi:10.1155/2013/795178.
 34. Rahman A., Sravani G. J., Srividya K., Priyadharshni A. D. R., Narmada A., Sahithi K., Krishna Sai T., Padmavathi Y. Development and validation of chemometric-assisted FTIR spectroscopic method for simultaneous estimation of valsartan and hydrochlorothiazide in pure and pharmaceutical dosage forms. *J Young Pharm.* 2020;12(2):s51–s55. doi:10.5530/jyp.2020.12s.46
 35. Pruksapha P., Khongkaew P., Suwanvecho C., Nuchtavorn N., Phechkrajang C., Suntornsuk L. Chemometrics-assisted spectroscopic methods for rapid analysis of combined anti-malarial tablets. *J Food Drug Anal.* 2023;31(2):338–357.

- doi:10.38212/2224-6614.3449.
36. Winning H., Larsen F. H., Bro R, Engelsen S. B. Quantitative analysis of NMR spectra with chemometrics. *J Magn Reson.* 2008;190(1):26–32. doi:10.1016/j.jmr.2007.10.005.
37. Engelsen S. B., Savorani F, Rasmussen M. A. Chemometric exploration of quantitative NMR data. In: *eMagRes.* Vol 2. Wiley; 2013:267–278. <https://doi.org/10.1002/9780470034590.emrstm1304>
38. Aboushady D., Samir L., Masoud A., Elshoura Y., Mohamed A., Hanafi R. S., El Deeb S. Chemometric approaches for sustainable pharmaceutical analysis using liquid chromatography. *Chemistry.* 2025;7(1):11. doi:10.3390/chemistry7010011
39. Naveen K., Bansal A., Lalotra R., Sarma G. S., Rawal R. K. Chemometrics-assisted quantitative estimation of synthetic and marketed formulations. *Asian J Biomed Pharm Sci.* 2014;4(34):21–26. doi:10.15272/ajbps.v4i34.510 <https://www.alliedacademies.org/articles/chemometrics-assisted-quantitative-estimation-of-synthetic-and-marketed-formulations.pdf>
40. Calvo N.L., Maggio R.M., Kaufman T. S. Chemometrics-assisted solid-state characterization of pharmaceutically relevant materials. Polymorphic substances. *Journal of Pharmaceutical and Biomedical Analysis.* 2018;147:518-537. doi:10.1016/j.jpba.2017.06.018
41. Jalalvand A. R. Application of chemometrics-assisted voltammetric analysis. In: Stoytcheva M, Zlatev R, editors. *Applications of the Voltammetry.* InTech; 2017:1–4. <https://doi.org/10.5772/67310>
42. Dinç E., Dermi^o S. Ç., Akçasoy S. C., Ertekin Özkan Z. C. A New Chemometric Strategy in Electrochemical Method, *Electroanalysis, an international journal devoted by electroanalysis, sensor and bioelectronic Devices.* 2020;32(3):613-619 doi:10.1002/elan.201900574
43. Brereton R. G., Jansen J., Lopes J., Marini F., Pomerantsev A., Rodionova O., Roger J. M., Walczak B., Tauler R. Chemometrics in analytical chemistry-Part I: history, experimental design and data analysis tools. *Anal Bioanal Chem.* 2017;409(25):5891-5899. doi:10.1007/s00216-017-0517-1
44. Maggio R. M., Rivero M. A., Kaufman T. S. Simultaneous acquisition of the dissolution curves. *journal of pharmaceutical and biomedical analysis.* 2013;72(18):51-58. doi:10.1016/j.jpba.2012.09.022.
45. Wiberg K.H., Hultin U.K. Multivariate Chemometric Approach to Fiber-Optic Dissolution Testing. *Anal. Chem.* 2006;78:5076-5085. <https://pubs.acs.org/doi/pdf/10.1021/ac0602928>
46. Rajkó R., Nassab P.R., Szabó-Révész P. “Self-modeling curve resolution method applied for the evaluation of dissolution testing data: a case study of meloxicam–mannitol binary systems.” *Talanta.* 2009;79:268-274. <https://doi.org/10.1016/j.talanta.2009.03.068>
47. Mantsch H.H. Biomedical Vibrational Spectroscopy in the Era of Artificial Intelligence. *Molecules.* 2021;26:1439. <https://doi.org/10.3390/molecules26051439>
48. Zulkifli B., Fakri F., Odigie J., Nnabuife L., Isitua C. C., Chiari W. Chemometric-empowered spectroscopic techniques in pharmaceutical fields: A bibliometric analysis and updated review. *Narra X.* 2023;1(1):80. doi:10.52225/narrax.v1i1.80.