e-ISSN:2321-0812
p-ISSN:2347-2340

Research and Reviews: Journal of Pharmaceutical Analysis

Is Chemo-Metrics Showing the Pathway for Unsolved Questions in Chemical / Drug Analysis?.

Ritika, Navdeep K Gill, Ravindra K Rawal, and GS Sarma*

Department of Pharmaceutical Analysis, Indo-Soviet Friendship College of Pharmacy, Moga, Punjab – 142 001, India.

Review Article

Received: 05/07/2014
Revised: 19/08/2014
Accepted: 27/08/2014

*For Correspondence

Department of Pharmaceutical Analysis, Indo-Soviet Friendship College of Pharmacy, Moga, Punjab – 142 001, India.
Mobiel: +91 94630 95859

Keywords: Chemometrics, Multivariate Analysis, CLS, ILS, PCA.

ABSTRACT

Analytical chemistry is the science of modeling quantitative measurements. Basically, it is the quantification of analytes in combined dosage forms samples which exhibits as the crucial part of the subject. Mainly, it is the radical apprehension of practical aspects of analysis, and the ability to identify a problem which cannot be solved by the existing analytical methods/techniques that motivate an analyst to develop creative approaches or new analytical methods. The answer to the all problems faced in analysis is chemometrics. Chemometrics has been defined as elementary tool for the implementation of mathematical and statistical methods to chemical measurements. Chemometric methods are often applied in situations when no sufficient theory is available for describing or solving analysis problems. The aim of chemometrics is to find out hidden relationships exist between the available data and the desired information. The prominent problems in chemistry are: recognition of the chemical structure from spectral data (spectral elucidation), quantitative analysis of chemical substances in complex mixtures (multivariate calibration), and determination of the origin of samples (cluster analysis and classification), and prediction of properties or activities of chemical compounds or technological materials (quantitative structure-activity or structure-property relationships). This literary work includes the basics of chemometrics and its aid to put this novel analytical discipline into application to face the present challenges of analytical chemistry.

INTRODUCTION

Chemometrics is the methodology which is very well known in the world of analysis from last few years. This method was originated by S.Wold in 1972 who is physical organic chemist, jointly with B.R Kowalski (American analytical chemist), S.Wold combine to create the international chemometrics community or society [1]. Thus, chemometrics can be explained as the science which gives the information that how the chemical measurement can be done with the help of mathematical and statistical method. The term chemometrics originated from chemistry and measurement. Thus this method helps to obtain the hidden information from the chemical data as shown in Fig. 1 [2]. Now in this world of technology, the chemometrics method is the expansion as computer software that contribute to getting final recognition from the unsolved information that present in term of raw or crude data [3,4]. Some physicist or analyst reviewed about the chemometrics method as a sub field for modern analytical chemistry. But the chemometricans themselves consider chemometrics set the new directions for analysis [5].

Need For Chemometrics

There are several reason why chemometrics is needed to solve the unsolved problems in the world simultaneous analysis such as:

e-ISSN:2321-0812
p-ISSN:2347-2340

- Chemometrics helps in achieving the results of chemical signals (chromatograms, kinetic curves, titration curves, or in other formats with the wavelength, or frequency) obtained from analytical measurements in the form of the digital vector.
- Signal noiseratioin proper design have been found before the results were incorporated [6].
- Fig. 2 shows the graph where one can see that there is overlapping in the spectra of various drugs. Since, the conventional methods are not able to resolve the tedious overlapping of the drugs spectra. Thus, chemometrics assisted multipurpose methods are used to resolve such heavy data.

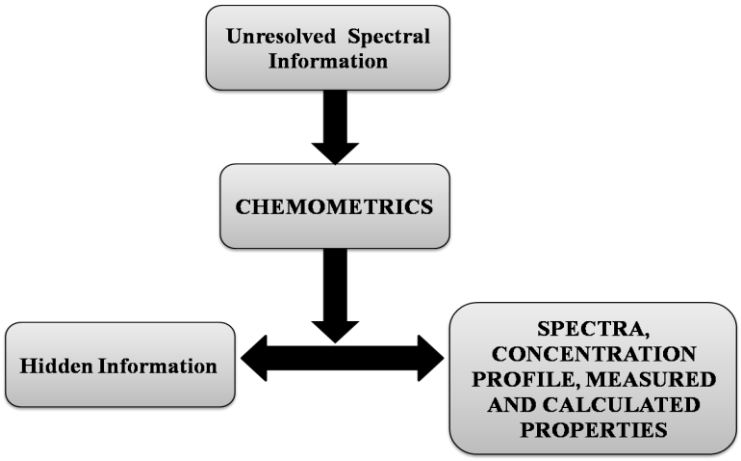

Figure 1: Chemometric Analytical Flow.

- Usually, analytical chemists try their best to adjust the experimental conditions to separate these overlapping peaks in the complex mixture. Yet, it is time-consuming and not easy to achieve this, especially for complicated systems such as herbal medicines, more than four drugs mixtures which have many constituents.

Thus, different multivariate analysis techniques have been developed to accomplish the target. The powerful computers currently available considerably help in this respect.

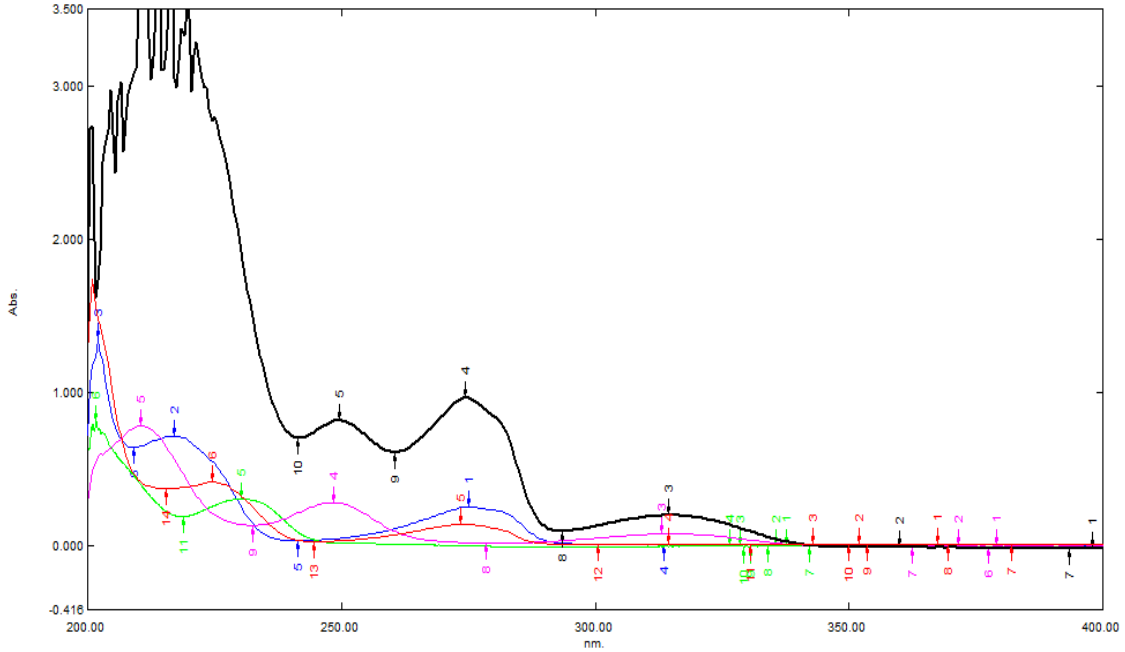

Figure 2: Zero order overlay spectra of drugs.

Extensity of Analytical Chemical Data

Extensity of analytical chemical data is limited in the chemometrics method. The chemometrics method helps to resolve the very complex data or three-dimensional data into very simple or one-

e-ISSN:2321-0812
p-ISSN:2347-2340

dimensional data.A sample detected in the UV-Visible spectrophotometer at definite wavelength having some absorbance of the definite concentration. Therefore, there is only one variable and that is absorbance. The resulting data is the one-dimensional. For example, the sample is scan in the range of 275-245 nm with the resolution of 1 nm gives the 30 resulting absorbencies [7].This is implemented also on HPLC-DAD in which the absorbance is a function of both wavelength and retention time. So, the absorbance is the variable and provides1D-data. In the second case, when the sample is again determined with the two variables in UV-Visible spectrophotometer i.e. absorbance and wavelength that means different absorbencies were recorded at more than one wavelength with the constant concentration. The resulting data is 2D-matrix. The absorbance of a sample is determined as a function of wavelength and retention time in HPLC-DAD.In third case, followed by the same procedure of second case, there are three variables i.e. absorbance, wavelength, and concentration. The resulting data so obtained knownas 3D-matrix or dataset. The absorbance is determined as a function of wavelength, retention time and samples in HPLC-DAD.

This extensity of chemical data can be anticipated by the chemometrics method now. But as it describe above the chemometrics method cannot be applied on one-dimensional data because this method is useful in foresee of multivariate data. But one can resolve three-dimensional data into one-dimensional data. There are most common or simple chemometrics methods that can be classified into Least Square Regression (CLS), Inverse Least Square Regression (ILS), Net Analytical Signal (NAS).But for very complex calculations the most useful methods that come into consideration are: PrincipalComponent Analysis(PCA), Partial Least Squares Regression (PLS), and Parallel Factor Analysis (PARAFAC) etc[7]. Three way data unwind into two way data. This three way data cut the segments of 3D-data into 2D-data. Then by simplification of 2D-data, it forms the large 2D-data or matrix [8]. The 3D-data unwind into three different directions along the row space, along the column space and along the third direction.

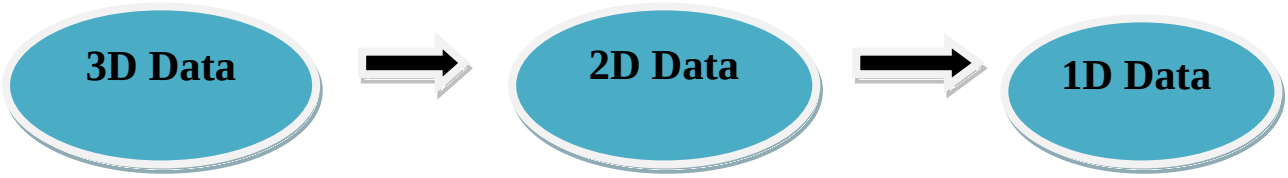

Figure 3: Unfolding of data.

METHOD VALIDATION

Method validation is the process that gives the conformation that the method adopted for any test whether it is working properly or not. Positive results obtained after validation confirms the uniformity or trueness of the analytical method. Thus, the validation is the procedure which provides high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality characteristics [9].

Prediction set [10]

The prediction set is prepared for validation purpose. It is also known as cross validation. Various parameters are applied on prediction set. In chemometrics, validation parameters are termed as figure of merits (FOM). These validation parameters are:

- Sensitivity
- Selectivity
- Signal to noise ratio
- Analytical sensitivity
- Limit of detection
- Limit of discrimination

Sensitivity

A method is referring to be sensitive when it predicts the small variation in the concentration of interested analytes. The sensitivity of method is depending upon the concentration of the analytes. The

sensitivity can be anticipated by using the formula that is division of alteration in the instrumental respond to the alteration in the concentration of interested analyte [11-12]. This parameter can be applied on both the quantitative and qualitative analysis.

Selectivity

Selectivity refers to the method by which the particular analyte can be determined even in the apparition of another material or components. The selectivity generally gives the data about single analyte in a complex mixture [13].

Signal to noise ratio

It is defined as the rate of signal power to the noise power. For example, if the ratio between them exceeding more than 1:1 that means signal strength is more than noise. During the measurement of signal to noise ratio, the special care should be taken that the system must be at the same identical/parallel or alternate or equivalent points.

Analytical sensitivity

It is more effective than the simple sensitivity, as it gives the determination between sensitivity of analytes and instrumental noise [14, 15]. It generally requires the less concentration range. The unit of analytical sensitivity is concentration^{-1}.

Limit of detection(LOD)

It is the detection limit that detects the limited amount of interested analyte. Practically, it can be determined on blank by taking mean of blank and the standard deviation of the blank. There are different detection limit that can be used such as method detection limit (MDL) and practical quantification limit (PQL) etc [16-19].

Limit of discrimination

It is related to the analytical sensitivity or the limit of detection. It gives intervene in the concentration of prediction sample it is indicative different from the original analyte concentration in term of analytical signal [20].

Multivariate Analytical Methods for the Analysis

Hyphenation of spectroscopy with multivariate analysis combined to give the qualitative and quantitative information about the analytes. Involvement of chemometrics method with spectroscopic techniques gives the rapid, simple and efficient analytical method that is independent of the variables.The spectroscopic analysis involves the spectroscopic method in combination with the chemometrics data evaluation which in turn result the elevated automation. By using this method, a new analytical method is high capacity analysis (HCA) comes into consideration [21- 24]. The main chemometrics method implement on:

- Experimental design,
- Classical least square (CLS)
- Inverse least square (ILS)
- Principal component analysis (PCA),
- Net analyte signal- standard addition method (NAS-SAM)
- Partial least squares regression (PLS) and parallel factor analysis (PARAFAC).

These above analytical methods can be performed by scanning ultraviolet-visible spectroscopy, HPLC with diode array detector (HPLC-DAD) and fluorescence spectroscopy [25].

Classical Least Squares (CLS) [30, 31]

Classical Least Square technique is also considered as K matrix and employed for qualitative and quantitative determinations. It works on the basis of Beer-Lambert law. It is one of the advantageous method used in complex mixtures. Simple least regression method is helpful in the calculation of the

absorptivity coefficients. It includes variables such as absorbances (A), Molar absorptivity (ε), path length of the sample (b) and concentration of the compound (C).

Pathlength of the sample and Molar absorptivity kept constant during the quantitative experiments.

$$A = C \times K \quad \text{..... (1)}$$

Where K is the constant of molar absorptivity and path length.Concentration of unknown sample is determined by this equation with the help of absorbancies having known concentration. The equation for classical least square is given below.

$$K = (C^T \times C)^{-1}. C^T \times A \text{..... (2)}$$

Inverse Least Squares [32-34]

This method is also known as P- matrix and multiple linear regressions employed for quantitative determinations. By rearrange the Beer Lambert law the equation can be written as:

$$C = A \times P \text{..... (3)}$$

It includes variables such as:absorbance (A), molar absorptivity (ε), path length of sample (b), and concentration of the compound in solution (C).For the calibration the inverse least squares solution to above equation:-

$$P = (A^T \times A)^{-1} A^T \times C \text{..... (4)}$$

The advantage of the method is that the equations or unsolved matrix can be easily calculated. On the basis of recognition of constituents of interest, this model helps to predict the very complex mixture. Thus, those complex mixtures that cannot be resolved by CLS method can easily predict by ILS method.

Principle Component Regression [35, 36]

It is the combination of inverse least square and principal component analysis. It is divided into following steps.

By considering data matrix which is performed by PCA for different variables, components were obtained. On the suitable basis selection of subset was done and PCA considered for further use.

- By the application of linear regression on co-variable estimated regression coefficient get vector and the observed vector is the outcome of principal components.
- With the help of PCA loading, principal components the vector now transformed into actual co various to get final PCR estimation which give characterization of original model.

The highest values of Eigen values give the optimal numbers of principal components. The concentration of unknown sample can be easily determined by the application of linear regression equation:

$$C = a + b \times A \text{..... (5)}$$

We calculate the coefficients a and b coefficient

$$b = P \times Q \text{..... (6)}$$

Where P is the matrix of eigenvectors and Q is C-loadings given by

$$Q = D \times T^T \times A_0 \text{..... (7)}$$

Where, T^T is the transpose of the score matrix T, D is a diagonal matrix having on the diagonal components the inverses of the selected Eigenvalues.

Knowing b, we found a using formula:

$$a = C_{mean} - A^{T}_{mean} \times b \text{..... (8)}$$

Where, C_{mean} represents the mean concentration of the calibration set and A^{T}_{mean} is the transpose of the matrix having the entries of mean absorbance values.

Net Analyte Signal [37]

The net analyte signal is unique signal which is termed as NAS and defined as the part of the signal that is orthogonal to the signal of the interferences (other analytes) present in the sample.NAS vector for any component in a mixture after each standard addition, can be found by the following equation.

$$NAS_{drug} = (I - P^{+}P)\, Ai \text{..... (9)}$$

This NAS can be directly relates to the unknown concentrations of components in the sample mixture and to the other Figures of Merits (FOM).

CONCLUSION

As we discussed above the chemometrics method is used for the multivariate data analysis. Chemometrics is the mathematical application on the chemical data to obtain the related information. Currently, the chemometrics can be applicable on various fields. The data with one or two variable can predict by the classical methods such as simultaneous equation method and Vierodt's method etc. But when the data contain more than two variables or the complex data the methods such as CLS, ILS, NAS, and PCAetc. Comes into consideration. The advantage of chemometrics is that not only it can predict all the variables simultaneously but also it doesnot require expensive solvents like HPLC grade unless if we perform the experiment through the HPLC technique. Moreover UV based chemometrics is widely used because ease of solvents like AR grade. Therefore, from above description, the basis of chemometrics can be explained as the data or matrix can easily simplify by the chemometrics to get the important relevant information. In modern era, it gives the more accurate results and the validation is performed at each and every step and more importantly the results were found in graphical manner that can be easily understood. Chemometrics is useful not only in analytical chemistry but also in food analysis and drug design and herbal drug standardization. Finally, we conclude that chemometrics a way for some unresolved questions in drug/chemical analysis.

REFERENCES

1. Ortiz MC and Sarabia L. Quantitative determination in chromatographic analysis based on n-way calibration strategies. J Chromatogr A. 2007; 1158: 94-110.
2. Xu L, Tang LJ and Cai CB,. Chemometric methods for evaluation of chromatographic separation quality from two-way data—A review. Anal Chim Acta. 2008; 613: 121-134.
3. Siouffi A and Phan-Tan-Luu R. Optimization methods in chromatography and capillary electrophoresis. J Chromatogr A. 2000; 892: 75-106.
4. Vanbel PF. Development of flexible and efficient strategies for optimizing chromatographic separations. J Pharm Biomed Anal. 1999; 21: 603-610.
5. Liang YZ, Xie P and Chan KC. Quality control of herbal medicines. J Chromatogr B. 2004; 812: 53-70.
6. Kiralj R and Ferreira M. The past, present, and future of chemometrics worldwide: some etymological, linguistic, and bibliometric investigations. J Chemometr. 2006; 20: 247-272.
7. Jing D and Linfang H. Application of chemometrics in quality evaluation of medicinal plants. J Medicinal Plants Res. 2006; 5: 4001-4008.
8. Hopke PK. The evolution of chemometrics. Anal Chim Acta. 2003; 500: 365-377.
9. Dinç E. Linear regression analysis and its application to the multivariate spectral calibrations for the multiresolution of a ternary mixture of caffeine, paracetamol and metamizol in tablets. J Pharm Biomed Anal. 2003; 33: 605-615.
10. CO Alejandro, MF Nicolaas, F Joan, B Ricard, HK John, and M Howard. Uncertainty Estimation and Figures of Merit for Multivariate Calibration. Int Union Pure App Chem. 2006;78:633-661.

11. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. International Vocabulary of Basic and General Terms In Metrology (Vim), International Organization For Standardization (Iso), Geneva. 1993: 2.
12. M Galea-Rojas, MV de Castilho, H Bolfarine and M de Castro. Detection of Analytical Bias. Analyst. 2003;128: 1073.
13. M Thompson, SLR Ellison and R Wood. Harmonised Guidelines for single laboratory validation of method analysis. Pure Appl Chem. 2002; 74: 835.
14. J. Mandel, RD Stiehler. J Res Natl Bur Stand. 1954; 53: 155.
15. L Cuadros Rodríguez, AM GarcíaCampaña, C Jiménez Linares, M Román Ceba. Estimation of Performance characteristics of an analytical method using the data set of the calibration experiment. Anal Lett. 1993; 26 : 1243.
16. LA Currie. Nomenclature in Evaluation of Analytical Methods Including Detection and Quantification Capabilities. Pure Appl. Chem. 1995; 67: 1699.
17. P Wilrich (Chairman). ISO 11843 series of standards about capability of detection; Part 1: Terms and definitions (1997) + Cor 1:2003, IDT; Part 2: Methodology in the linear calibration case (2000 IDT); Part 3: Methodology for determination of the critical value for the response variable when no calibration data are used (2003, IDT); Part 4: Methodology for comparing the minimum detectable value with a given value (2003, IDT); Part 5: Methodology in the linear and non-linear calibration cases (2005, AWI ISO TC 69/SC6).
18. LA Currie. Chemom. Limits of detection and decision. Intell Lab Syst. 1997; 37: 151.
19. R Boqué, and FX Rius. Computing detection limits in multicomponent spectroscopic analysis. Trends Anal Chem. 1997; 16: 432.
20. R Ferrús, and MR Egea. Anal Chim Acta. 1994; 287: 119.
21. Wiberg K. Quantitative impurity profiling by principal component analysis of high-performance liquid chromatography–diode array detection data. J Chromatogr A. 2006; 1108**:** 50-67.
22. Haaland DM, Thomas EV. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Anal Chem. 1988; 60**:** 1193-1202.
23. Musumarra G and Fichera M. Chemometrics and cultural heritage. Chemometr Intell Lab Syst. 1998; 44**:** 363-372.
24. Wold S, Sjöström M and Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab Syst. 2001; 58**:** 109-130.
25. Swierenga H, De Groot P, De Weijer A, Improvement of PLS model transferability by robust wavelength selection. Chemometr Intell Lab Syst. 1998; 41**:** 237-248.
26. Bro R. PARAFAC. Tutorial and applications. Chemometr Intell Lab Syst. 1997; 38**:** 149-171.
27. Marx BD, Eilers PHC and Li B. Multidimensional single-index signal regression. Chemometr Intell Lab Syst. 2011; 109**:** 120-130.
28. Carroll JD, Chang JJ. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika. 1970; 35**:** 283-319.
29. Kruskal JB. Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algebra App. 1977; 18**:** 95-138.
30. Lorber A. Features of quantifying chemical composition from two-dimensional data array by the rank annihilation factor analysis method. Anal Chem. 1985; 57**:** 2395-2397.
31. Faber N, Buydens L, Kateman G. Generalized rank annihilation method. I: derivation of eigenvalue problems. J Chemometr. 2005; 8**:** 147-154.
32. Sanchez E and Kowalski BR. Generalized rank annihilation factor analysis. Anal Chem. 1986; 58**:** 496-499.
33. Maeder M. Evolving factor analysis for the resolution of overlapping chromatographic peaks. Anal Chem. 1987; 59**:** 527-530.
34. Xu CJ, Jiang JH and Liang YZ. Evolving window orthogonal projections method for two-way data resolution. Analyst. 1999; 124**:** 1471-1476.
35. Kvalheim OM and Liang YZ. Heuristic evolving latent projections: Resolving two-way multicomponent data. 1. Selectivity, latent-projective graph, datascope, local rank, and unique resolution. Anal Chem. 1992; 64**:** 936-946.
36. Liang YZ, Kvalheim OM and Keller HR.. Heuristic evolving latent projections: resolving two-way multicomponent data. 2. Detection and resolution of minor constituents. Anal Chem. 1992; 64**:** 946-953.
37. Manne R, Shen H and Liang Y. Subwindow factor analysis. Chemometr Intell Lab Syst. 1999; 45**:** 171-176