Commentry   **Open Access**

Chemometrics and its Application in Pharmaceutical Field

Harshal Ashok Pawar[1]* and Swati Ramesh Kamat[2]

[1]*Assistant Professor and HOD (Quality Assurance), Dr. L. H. Hiranandani College of Pharmacy, Ulhasnagar-421003, Maharashtra, India*
[2]*Research Scholar (M.Pharm.), Dr. L. H. Hiranandani College of Pharmacy, Ulhasnagar-421003, Maharashtra, India*

Abstract

Chemometrics is the application of statistical and mathematical methods to analytical data to permit maximum collection and extraction of useful information. It is a data-driven interdisciplinary science suitable for solving diverse applications. This review focuses mainly on various chemometric models used and their applications in the pharmaceutical sciences.

Keywords: Chemometrics; Bilinear models; Pharmaceutical science

Introduction

Chemometrics is a branch of science that is used for extraction of the data related to chemical and physical phenomena involved in the manufacturing process by the application of the statistical and mathematical methods. It can be applied in predictive issues solving like predicting the target properties, desired features. Also can be used for the descriptive issue solving like the model composition, identification and understanding. Chemometrics shows its application in the multivariate data collection and analysis. Various algorithms and analogous ways are available for processing and evaluating the data. They can be implemented to various fields, like medicine, pharmacy, food control, and environmental monitoring [1,2]. Some of the Chemometric models for analysis are mentioned below

Bilinear models

In this the data is arranged in data matrices in such a way that each vertical column has variables and each horizontal row has the samples [1]. Bilinear chemometric techniques includes following:

Principal Component Analysis (PCA): This is a simple and non parametric technique which is used for extracting the relevant information from the data sets. It can be used to express the data on the basis of their similarity and the differences. It is widely used in multivariate data analysis. PCA reduces the dimensionality and multivariate data compression in different fields of sciences. During process monitoring, it can be used to develop a correlation structure between variables and also examine the changes. Thus it reduces the number of variables in process.

If for a series of sites, or objects, or persons, a number of variables are measured, then each variable will have a variance, and usually the variables will be associated with each other; that is, there will be covariance between pairs of variables. In PCA, data is transformed to describe the same amount of variability [3,4]. Luciana et al. applied chemometric tools, such as principal component analysis (PCA), consensus PCA (CPCA), to a set of forty natural compounds, acting as NADH-oxidase inhibitors [5].

Partial Least Squares [PLS]: It is one of the widely implemented methods which describe the relationship between different sets of different observed variables by the means of latent variables. The basic assumption of this method is that it modifies relations between sets of the observed variables by a small number of latent variables (not directly observed or measured) by incorporating regression, dimension reduction techniques, and modeling tools [6-10]. The latent variables increase the covariance between the different sets of variables. PLS is

similar to canonical correlation analysis (CCA) and can be used as a discrimination tool and dimension reduction method like principal component analysis (PCA). PLS is widely used technique as it can process large chemical data [11-15]. Determination of flow properties of pharmaceutical powders by near infrared spectroscopy NIR spectroscopy was done using Partial least square technique [16].

Multiway models

Multiway models are basically used when the data is multivariate and linear in more than 2 dimension arrays. Bilinear techniques could not provide sufficient data which was provided by multivariate techniques. The methods like multiway principal component analysis

(MPCA) and multiway partial least squares (MPLS) improve the process understanding and summarizes its behavior in a batch-wise manner and are therefore recognized as tools for monitoring batch data. However, if the original data contains higher dimensions then it becomes difficult for the models to interpret the computed data and therefore multiway methods that work with three-way or higher arrays like parallel factor analysis (PARAFAC andPARAFAC-2, Tucker-3, and N-partial least squares (N-PLS))are the methods of choice It is used widely in extracting the data from spectra which is usually difficult in cases of overlapping. These multiway methods that work with three or higher ways are the methods of choice [17,18].

Parallel Factor Analysis (PARAFAC): Parallel factor analysis (PARAFAC) is a decomposition method used for the modeling of three-way or higher data and is mainly intended for data having congruent variable profiles within each batch. The brief history to understand PARAFAC is as follows. Cattell reviewed seven principles for the choice of rotation in component analysis and concluded that the principle of "Parallel proportional profiles" the most fundamental principle. This principle means that the two data matrices with the same variables should contain the same components. By using this principle as a constraint Harshman proposed a new method to analyze

two or more data matrices that contain scores for the same person on the same variables and termed the method as PARAFAC [19-21].

Parallel Factor Analysis -2 (PARAFAC-2): PARAFAC-2 can handle data variable profiles that are shifted or/are in a different phase. In PARAFAC trilinearity is a fundamental condition whereas PARAFAC-2 enables trilinearity. However, it should be noted that PARAFAC may be used to fit nonlinearity to some extent in one mode only in cases where data shifts from linearity are regular. Both the techniques are mainly applied for analyzing chemical data from experiments that form a 3-way or higher data structure, for example, chromatographic data, fluorescence spectroscopy measurements, temporal varied spectroscopy data with overlapping spectral profiles, and process data [22,23].

Tucker-3 model: This can be used for exploring n way array data as it consists of n modes of loading matrices. The generality of the Tucker-3 model, and the fact that it covers the PARAFAC model as a special case, has made it an often used model for decomposition, compression, and interpretation in many applications [24].

N-Partial Least Square (N-PLS): For handling a multiway data extension of PLS method namely N-PLS was introduced. It basically uses dependent and independent variables for finding the latent variables which describes maximal covariance [25].

Techniques Used in Chemometric

Nowadays various spectroscopic techniques like HPLC, NIR, FTIR are being combined with various chemometric models like multivariate analysis methods, PLS, CLS, PCR, and so forth, for the evaluation of different pharmaceutical properties of tablets, powders, granules, and so forth. The most popular technique used is NIR spectroscopy.

Applications of Chemometrics in Pharmaceutical Field

Chemometric method is widely used in different areas of pharmaceutical fields like manufacturing, quality evaluation and quality assurance.

Powder flow properties

Sarraguca et al. determined flow properties of pharmaceutical powders using near IR spectroscopy. The experimental results obtained were correlated with the NIR spectrum. Partial least squares (PLS) method was used for the correlation [16]. Kim et al. determined density of polyethylene pellets using transmission Raman spectroscopy. Transmission Raman spectra were collected for 25 different grades of polyethylene pellets. The partial least squares method was used to determine the sample density [26]. Otsuka et al. developed a quick and accurate way to determine the pharmaceutical properties of granules and tablets in the formulation of pharmaceuticals by application of chemoinfometric NIR spectroscopy. To predict the pharmaceutical properties such as mean particle size, angle of repose, tablet porosity, and tablet hardness NIR spectra of the Antipyrine granules were measured. This was analyzed by principal component regression analysis. With the increase in the water amount, the mean particle size of the granules was found to increase from $81\mu m$ to $650\mu m$, and it was possible to make larger spherical granules with narrow particle size distribution using a high-speed mixer [27].

Water content determination

Water content of hygroscopic pharmaceutical excipients largely affects the manufacturing processes and the performance of the final product. And so it is necessary to determine the water content of such excipients. The water content of three commonly used tablet disintegrants namely crospovidone, croscarmellose sodium and sodium starch glycolate was studied by Szakonyi et al. [28].

Dissolution studies

The preparation parameters involved in increased solubility of poorly soluble Meloxicam drug was studied by Ambrus et al. The dissolution rate was improved by formulating the drug as a nanosuspension using different methods like emulsion diffusion, high-pressure homogenization, and sonication. SMCR method on the XRPD patterns of the nanosuspensions was used which revealed the crystalline form of the drug and the strong interaction between Meloxicam and the stabilizer [29].

Tablet parametric test

The determination of disintegration time of theophylline tablets was studied by Donoso and Ghaly using NIR reflectance spectroscopy. Laboratory disintegration time was compared to near infrared diffuse reflectance data [30]. Donoso et al. researched the use of the near infrared diffuse reflectance method to evaluate and quantify the effects of hardness and porosity on the near infrared spectra of tablets. The results demonstrated that an increase in tablet hardness and a decrease in tablets porosity produced an increase in near infrared absorbance [31]. Ebube et al. evaluated a method using NIR spectroscopy as a nondestructive technique to differentiate three microcrystalline cellulose forms in powdered form and in compressed tablets. Avicel grades PH-I0l, PH-102, and PH-200 were evaluated in their study. The developed technique was able to identify both in powdered form and in compressed tablets the three different Avicel grades. The result was not affected by the presence of a lubricant, magnesium stearate. They also successfully developed a method for the determination of magnesium stearate by a multiple linear regression method [32]. Chen et al. used artificial neural network and partial least squares models for predicting the drug content in theophylline tablets. A better prediction of drug content was observed with a partial least squares model than with an artificial neural network model for drug content greater than or equal to 5%w/w whereas the artificial neural network model showed better results than the partial least squares model at less than or equal to 2% w/w theophylline content [33].

Formulation development

Formulation and evaluation of protein-loaded solid dispersions by nondestructive methods. Technique used was Powder X-ray diffraction (PXRD), near infrared chemical imaging (NIR-CI) and the method used was Principal component analysis and partial least square regression [34].

Pharmaceutical analysis

Simultaneous determination of ambroxol hydrochloride and guaifenesin was carried out by HPLC, using principle component regression (PCR), and partial least squares (PLS) [35].

Conclusion

Chemometrics and its methods are versatile and there is a high level of abstraction as it characterizes the scientific disciplines extensively by the application of the statistical and mathematical methods. Various chemometrics models have been applied for the analysis of data of a particular manufacturing process, quality control test, or an instrumental output data with an aim to achieve maximum accuracy,

precision, and robustness. Chemometric is having wide applications in the pharmaceutical and medical field. The implementation of chemometric techniques with a view of ensuring overall production process control entails the use of analytical techniques capable of providing accurate results in a simple and rapid manner.

References

1. Matero S (2010) Chemometrics Methods in Pharmaceutical Tablet Development and Manufacturing Unit Operations. Publications of the University of Eastern Finland Dissertations in Health Sciences.

2. Mocak J (2012) Chemometrics in medicine and pharmacy. Nova BiotechnologicaetChimica 11: 11–25.

3. Jolliffe IT (2002) Principal Component Analysis, Springer, New York, NY, USA, 2nd edition.

4. Wentzell PD, Andrews DT, Hamilton DC, Faber K, Kowalski BR (1997) Maximum likelihood principal component analysis. Journal of Chemometrics 11: 339–366.

5. Luciana S, Elizabeth IF, Sobral SM, Marcus TS (2010) Chemometric studies on natural products as potential inhibitors of the nadh oxidase from trypanosoma cruzi using the volsurf approach. Molecules 15: 7363–7377.

6. Wold H (1982) Soft Modeling: The Basic Design and Some Extensions, North Holl and Press, Amsterdam, The Netherlands.

7. Wold H (1985) Encyclopedia of the Statistical Sciences, JohnWiley and Sons, New York, NY, USA.

8. Wold S, Ruhe H, Wold H, Dunn WJ (1984) The collinearity problem in linear regression: the partial least squares (PLS) approach to generalized inverse. SIAM Journal on Scientific Computing 5: 735–743.

9. Hoskuldsson A (1988) PLS regression methods. Journal of Chemometrics 2: 211–228.

10. Barker M, Rayens W (2003) Partial least squares for discrimination. Journal of Chemometrics 17: 166–173.

11. Garthwaite PH (1994) An interpretation of partial least squares. Journal of the American Statistical Association 89: 122–127.

12. Worsley KJ (1997) An overview and some new developments in the statistical analysis of PET and fMRI data. Hum Brain Mapp 5: 254-258.

13. Wold S (2001) Personal memories of the early PLS development. Chemometrics and Intelligent Laboratory Systems 58: 83–84.

14. Frank IE, Friedman JH (1993) A statistical view of some chemometrics regression tools. Technometrics 35: 109– 147.

15. Martens H (2001) Reliable and relevant modelling of real world data: a personal account of the development of PLS Regression. Chemometrics and Intelligent Laboratory Systems 58: 85–95.

16. Sarraguça MC, Cruz AV, Soares SO, Amaral HR, Costa PC, et al. (2010) Determination of flow properties of pharmaceutical powders by near infrared spectroscopy. J Pharm Biomed Anal 52: 484-492.

17. Bro R (1996) Multiway calibration: multilinear PLS. Journal of Chemometrics 10: 47–61.

18. Bro R, Kiers HAL (2003) A new efficient method for determining the number of components in PARAFACmodels. Journal of Chemometrics 17: 274–286.

19. Andersen CM, Bro R (2003) Practical aspects of PARAFACmodeling of fluorescence excitation-emission data. Journal of Chemometrics 17: 200–215.

20. Bro R (1997) PARAFAC: tutorial and applications. Chemometrics and Intelligent Laboratory Systems 38: 149–171.

21. Rinnan A, Riu J, Bro R (2007) Multi-way prediction in thepresence of uncalibrated interferents. Journal of Chemometrics 21: 76–86.

22. Bro R, Andersson CA, Kiers HAL (1999) PARAFAC2. PartII: modeling chromatographic data with retention time shifts. Journal of Chemometrics 13: 295–309.

23. Kiers HAL, Ten Berge JMF, Bro R (1999) PARAFAC2. PartI: a direct fitting algorithm for the PARAFAC2 model. Journal of Chemometrics 13: 275–294.

24. Smilde A (2001) Comments on three-way analyses used for batch process data. Journal of Chemometrics 15: 19–27.

25. Olivieri AC (2008) Analytical advantages of multivariate data processing. One, two, three, infinity? Anal Chem 80: 5713-5720.

26. Sarraguça MC1, Cruz AV, Soares SO, Amaral HR, Costa PC, et al. (2010) Determination of flow properties of pharmaceutical powders by near infrared spectroscopy. J Pharm Biomed Anal 52: 484-492.

27. Kim M, Chung H, Jung YM (2011) Accurate determination of polyethylene pellet density using transmission Raman spectroscopy. Journal of Raman Spectroscopy 42: 1967–1976.

28. Otsuka M, Mouri Y, Matsuda Y (2003) Chemometric evaluation of pharmaceutical properties of antipyrine granules by near-infrared spectroscopy. AAPS Pharm Sci Tech 4: E47.

29. Szakonyi G, Zelkó R (2012) Water content determination of superdisintegrants by means of ATR-FTIR spectroscopy. J Pharm Biomed Anal 63: 106-111.

30. Ambrus R, Kocbek P, Kristl J, Sibanc R, Rajk´o R, et al. (2009) Investigation of preparation parameters to improve the dissolution of poorly water-soluble meloxicam. International Journal of Pharmaceutics 381: 153–159.

31. Donoso M, Ghaly ES (2005) Prediction of tablets disintegration times using near-infrared diffuse reflectance spectroscopy as a nondestructive method. Pharmaceutical Development and Technology 10: 211–217.

32. Donoso M, Kildsig DO, Ghaly ES (2003) Prediction of tablet hardness and porosity using near-infrared diffuse reflectance spectroscopy as a nondestructive method. Pharmaceutical Development and Technology 8: 357–366.

33. Ebube NK, Thosar SS, Roberts RA, Kemper MS, Rubinovitz R, et al. (1999) Application of near-infrared spectroscopy for nondestructive analysis of Avicel powders and tablets. Pharm Dev Technol 4: 19-26.

34. Chen Y, Thosar SS, Forbess RA, Kemper MS, Rubinovitz RL, et al. (2001) Prediction of drug content and hardness of intact tablets using artificial neural network and near-infrared spectroscopy. Drug Development and Industrial Pharmacy 27: 623–631.

35. Dinc E, Ozdemir A, Aksoy H, Ustundag O, Baleanud D (2006) Chemometric determination of naproxen sodium and pseudoephedrine hydrochloride in tablets by HPLC. Chemical andPharmaceutical Bulletin 54: 415–421.

Figure 1: Flow chart depicting steps involved in MD simulation of a protein.