

文章编号:1671-9352(2009)11-0016-05

化学计量学方法用于蛤青注射液色谱数据重叠峰的分辨

王康,李华

(长安大学理学院,陕西 西安 710064)

摘要:固定窗口因子分析-直观推导演进投影法被用来将中药色谱-光谱二维数据中重叠峰的部分解析为纯组分色谱曲线和纯组分光谱曲线。在本文中,这种方法都被用来解析复方蛤青注射液色谱-光谱数据中重叠现象非常严重的一部分组分簇。通过对该方法结果的分析,提出了一种改进的新方法——扣除数量级固定窗口因子分析。

关键词:化学计量学;数据解析;中药;直观推导式演进特征投影法;蛤青注射液;色谱

中图分类号:O657 **文献标志码:**A

Analysis of the compound Haqing injection with hyphenated chromatography and chemometric resolution

WANG Kang, LI Hua

(College of Science, Chang'an University, Xi'an 710064, Shaanxi, China)

Abstract: Sectional moving window factor analysis (SMWEA)-heuristic evolving latent projection (HELP), are used to resolve the overlapped two-dimensional chromatogram-spectrum data of the Traditional Chinese Medicine into estimated pure component chromatograms and spectrum profiles. In this work, the method was used to resolve a heavily overlapped components cluster of the compound Haqing injection. On the basis of results of the method, an improved deducted order of magnitude of the fixed size moving window evolving factor analysis (DOMFSMWEA) procedure was proposed and applied in the resolution of the compound Haqing injection, and the satisfactory results were obtained.

Key words: chemometric; resolution; traditional Chinese medicine (TCM); heuristic evolving latent projection (HELP); compound Haqing injection; chromatography

中药因其明确的药理学活性和较低的毒性^[1],在中国的使用已经有几千年的悠久历史。但是目前仍然没有一种简单、方便和有效的质量检测技术来测定那些包含几百种甚至几千种化合物的中药材及其产物。在处理复杂化学样品的分离和分析研究中,色谱技术是一种非常有力的工具。通常情况下,化学工作者总是希望通过最优化的色谱条件来得到所有组分都可以完全基线分离的色谱信息。但是对于复杂体系来说,找到这样的色谱条件往往是一件非常困难的事情。面对这样的问题,中药色谱指纹

图谱的建立为中药质量控制提供了一种非常有效的方法。中药色谱指纹图谱就是利用多组分化合物的色谱曲线特征来鉴别和评定色谱信息中化学组成的稳定性来对中药的质量进行鉴定和控制^[2-4]。除此之外,中药的一些局部信息或者某些组分簇详细的定性和定量分析也是非常必要的。

为实现这样的目标,发展了新的仪器和相关技术。随着近年来各种联用仪器的出现,例如:GC-MS, HPLC-MS, HPLC-DAD (diode array detection), HPCE-DAD 等等。可以同时分析得到色谱和光谱

收稿日期:2009-02-16

作者简介:王康(1977-),男,博士,讲师,主要从事化学计量学研究。Email:conquer741@163.com

的二维数据。通过解析二维数据,定性和定量分析不仅仅依靠保留时间、峰高、峰面积,而且可以利用紫外光谱或者质谱信息。为了更好地利用这样的信息数据,化学计量学在近年来也发展了一系列化学计量学分辨方法来处理二维数据^[5-15]。由上述仪器所得到的数据具有很多特性,如光谱曲线非负,色谱曲线非负、单峰等。Booksh 和 Kowalski^[8]首先提出了上述这些优化限制条件。这些特性使得对于复杂体系的解析有了成功的可能,并由此发展出了很多二维数据渐进分辨技术,如:直观演进投影法(heuristic evolving latent projections, HELP)^[5-7],渐进因子分析(evolution factor analysis, EFA)^[9-11],窗口因子分析(windows factor analysis, WFA)^[12-13],子窗口因子分析(sub-window factor analysis, SFA)^[14-15]和正交投影分辨(orthogonal projection resolution, OPR)^[16]。随着这些方法的发展,联用仪器和化学计量学分辨方法使得清晰和准确地分析“黑色分析体系”成为可能^[17]。这些方法被成功地应用在各个领域^[17-19]。在这些方法中,固定窗口因子分析-直观演进特征投影法(sectional moving window factor analysis-heuristic evolving latent projection, SMWEFA-HELP)被广泛地应用在中药、食品以及环境分析领域^[8-9]。

复方蛤青注射液系蟾蜍、苦杏仁、五味子、黄芪、紫菀、附子、白果、前胡、黑胡椒等九味中药组成的中药制剂。具有散风、清热、祛痰、止咳之功效。临床用于感冒、咳嗽、头痛等症状^[20]。

在本文中,固定窗口因子分析-直观演进特征投影这种方法被用来对复方蛤青注射液二维色谱光谱数据进行分辨。通过该方法所得结果的分析,提出了扣除数量级固定窗口因子分析(deducted order of magnitude of the fixed size moving windows evolving factor analysis, DOMFSMW-EFA)的新方法。

1 理论部分

1.1 直观推导式演进特征投影分析(HELP)

直观推导式演进特征投影法(Heuristic Evolving Latent Projection, HELP)利用 HPLC-DAD 等联用仪器产生的二维数据,采用一种逐步的、推导式的方式进行解析^[10-11]。它结合了渐进因子分析(EFA)^[7]、固定尺寸移动窗口渐进因子分析(FSMWEFA)^[12]、PCA^[13]等多种化学计量学方法,同时又开创性地提出了特征投影图(Latent-Projective Graphs, LPG)、特征结构跟踪分析(Eigenstructure Tracking Analysis, ETA)等概念,具有从原始数据的

背景扣除一直到重叠峰分辨的所有解析步骤。

首先,需要使用演进特征投影(evolution latent projection graphs, ELPG)法对矩阵中包含的物种数进行确定。所谓演进特征投影图,是对二维数据作 PCA 分解,用第一主成分对第二主成分作图即得到 LPG。特征投影图中直接通过原点的直线或经延长可通过原点的直线(基于最小二乘意义),就代表色谱和光谱空间的选择性信息,用化学的语言来说就是所谓单组分。所以演进特征投影图可以找到没有色谱峰相重叠的单组分,或称选择性区域。通过图中单组分区域的数量可以确定矩阵中所包含的组分数。

其次,确定某组分的选择性区域和零浓度区域。特征结构跟踪分析(ETA)是对固定尺寸移动窗口渐进因子分析法的一种拓广。将移动窗口局部 PCA 所得的特征值连线,第一条特征值曲线由各个窗口子矩阵的最大特征值连线而成,第二条曲线由次大特征值连线而成,依此类推。在 ETA 图中,若某区域中第一条特征曲线大于噪声水平,第二条接近噪声水平,则说明该区域中只有一个组分,即为选择性区域;若某区域中前二条特征曲线大于噪声水平而第三条接近噪声水平,则说明该区域为二组分重叠区域。因此,ETA 图就是一个各组分在保留时间方向上的分布信息图,故又被称为“化学秩图”。

第三,纯组分光谱和色谱的满秩分辨。根据:

$$X = CS^t + E,$$

其中, X 为量测得到的二维矩阵, C 为真实色谱阵, S 为真实光谱阵, E 为残差阵。令 T 、 P 分别为对 X 进行主成分分解后得到的正交阵(又称为抽象色谱阵和抽象光谱阵),则有:

$$CS^t = TP^t = TRR^{-1}P^t,$$

$$C = TR \quad S^t = R^{-1}P^t,$$

其中, R 称为满秩旋转矩阵。由以上两式可知,求出 R 阵就可得到各组分的色谱、光谱信息。HELP 方法利用某组分选择性区域和零浓度区域,借最小二乘技术可惟一地求出 R 阵,进而得到各组分的纯色谱和纯光谱。

1.2 扣除数量级固定窗口因子分析(deducted order of magnitude of the fixed size moving window evolving factor analysis, DOMFSMW-EFA)

由于特征结构跟踪分析中 PCA 所得的特征值数量级差距较大,因此在传统的固定窗口因子分析中,使用对数法处理特征值信息。但是这种处理方法削弱了特征值曲线的特征,使得无法对选择性区域进行准确判断,本文中使用一种新的方法来处理

特征值数据。该方法通过扣除特征值的数量级的方式代替对数法对特征值进行处理,所以将该方法命名为扣除数量级固定窗口因子分析(deducted order of magnitude of the fixed size moving window factor analysis, DOMFSMWFA)。扣除特征值数量级的具体过程如下公式所示:

$$e_{\text{MAX}} = \text{MAX}(E) = n \times 10^m, \quad (6)$$

$$E_{\text{dm}} = E / 10^m, \quad (7)$$

其中, E 表示特征值的行矩阵, e_{MAX} 为 E 矩阵的最大值, n 为 0 到 10 之间的实数, m 是一个整数, E_{dm} 为扣除特征值数量级后得到的结果。通过扣除数量级的处理,使得特征值均分布在 0~10 之间,并且能够保留特征值变化差异,在对选择性区域判断过程中更加清晰和准确。

2 实验部分

2.1 试剂和原料

蛤青注射液由陕西省华禹制药公司提供。流动相:甲醇-水-冰醋酸(20:80:0.05);样品体积:10 μL 。色谱纯甲醇购自 Tedia 公司(俄亥俄,美国)。分析纯冰醋酸购自上海化学试剂公司(上海,中国)。使用去离子水(Milli-Q water system, Millipore Bedford, MA, USA)对样品进行预处理。

2.2 仪器

Beckman 高效液相色谱仪(125pump, 168PDA diode array detector, 32Karat work station, USA), Beckman 紫外检测器, C18 色谱柱(250 mm $\times$ 4.6 mm, 5 μm), Beckman 508 自动进样系统。

检测波长:168~327 nm, 流速:0.8 ml/min, 温度:25 $^{\circ}\text{C}$ 。

2.3 实验过程

将大约 8 mL 蛤青注射液放入水浴锅中脱水干燥,将得到的残留物研磨成粉,使用甲醇-0.5%冰醋酸(20:80, v/v, 2 mL)溶液溶解,超声处理 15 分钟,离心处理 5 分钟,重复两次。取出上清液置于 50 mL 容量瓶中。将溶液用 0.45 μm 滤网,取 2 mL 用于 HPLC 分析。

2.4 样本

通过高效液相色谱得到蛤青注射液色谱数据。其中,很多组分表现为重叠峰。为了验证化学计量学分辨方法,将其中的一部分数据截取出来进行分辨,图 1 显示了整体色谱-光谱曲线,并标明了该部分数据的位置(放大部分)。经过背景扣除处理,得到的数据大小为色谱方向从 1050~1300 检测时间点,

光谱方向检测波长为 200~300 nm,数据矩阵为 250 行、100 列。

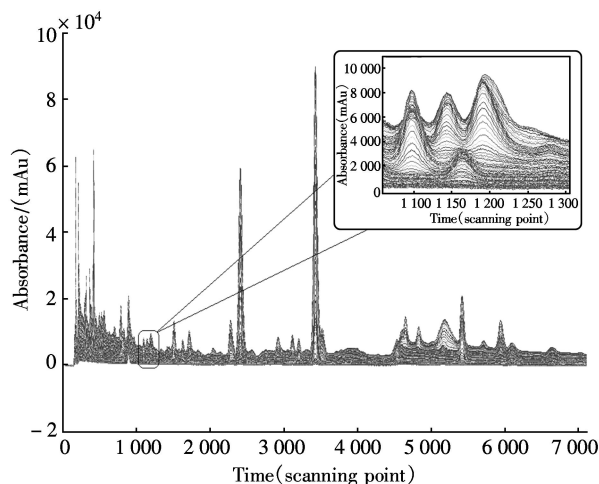


图 1 蛤青注射液样品 I 的 HPLC-DAD 二维数据图,其中保留时间为扫描点 1050~1300 的 A 部分被单独取出用于曲线分辨计算

Fig. 1 The two-dimensional data of Haqing injection by HPLC-DAD and the part A of scanning point 1050~1300 taken out from the whole data

2.5 数据分析

数据分析使用联想奔腾 4 个人计算机完成。所有使用的程序是在 Windows XP 环境下的 MATLAB 7.0 软件中实现的。

3 结果与讨论

3.1 数据预处理

因为数据中 A 部分(1050~1300 扫描点)重叠峰现象非常严重,所以选择该部分作为分析目标来进行方法比较。检测波长范围在 200~300 nm 的两组数据三维图绘制在图 2 中。

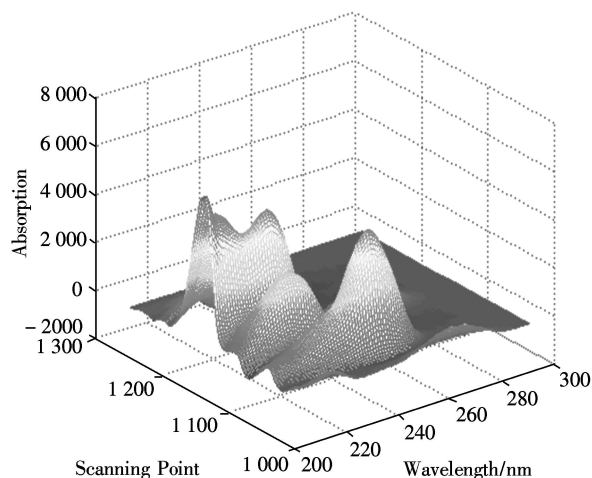


图 2 由 HPLC-DAD 得到的波长范围在 200~300 nm 的蛤青注射液样品中 A 部分的三维色谱光谱曲线

Fig. 2 The part A of Haqing injection three-dimensional graphics by HPLC-DAD of data in 200~300 nm

3.2 组分数的确定

首先使用演进特征投影法对矩阵中包含的物种数进行确定。在 ELPG 图中,按照色谱方向,直线通过原点的部分为单组分的纯物质选择性区域,曲线部分表示至少有两个组分重叠区域。图 3 显示了上述该数据的 ELPG 图,在图中标出了纯物质选择性区域。

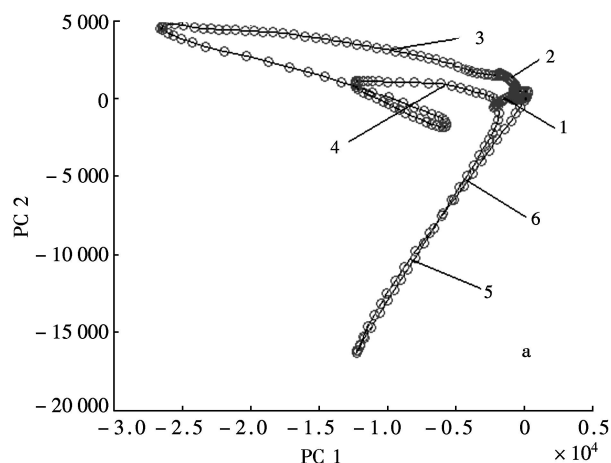


图 3 上述数据的 ELPG 图,直线并且其延长线通过原点的部分为纯物质区域,曲线部分为重叠物质区域

Fig. 3 The Evolving Latent Projection Graphs (ELPG), the straight lines and those prolongations can track the origin denote the pure region of the components while the curve represents the overlapping region

通过上述方法的计算,数据分辨所使用的组分数被确定为 6。

3.3 数据分辨方法

使用扣除数量级固定窗口因子分析方法 (Deducted order of magnitude of the fixed size moving window evolving factor analysis, DOMFSMW-EFA) 对数据进行第一次处理作为例子绘制于图 4-a 中。如图 4-a 所示,曲线被分为组分 1 的纯物质浓度区域、重叠区域和零浓度区域。应用传统的固定窗口因子分析 (FSMW-EFA) 所得结果绘于图 4-b,比较可见,因为使用了扣除数量级方法代替对数法来处理特征值,使得对于纯物质浓度区域和零浓度区域的选择更加清晰和准确。

在图 4-a 中,曲线被分为四个部分: I 组分 1 的零浓度区域; II 组分 1 的纯物质浓度区域; III 组分 1 和组分 2 的重叠区域; IV 组分 1 的零浓度区域。通过在图中选择出组分 1 的纯物质浓度区域和零浓度区域,可以得到组分 1 纯物质浓度曲线(C_1)和光谱曲线(S_1)的唯一解,然后再在第一组数据中扣除组分 1 的吸收数据 $C_1 \times S_1$ 得到第一组数据的剩余部分。重复以上步骤直至选择结束。使用 DOMFSMW-EFA 方法处理该数据的分辨结果列于图 5。

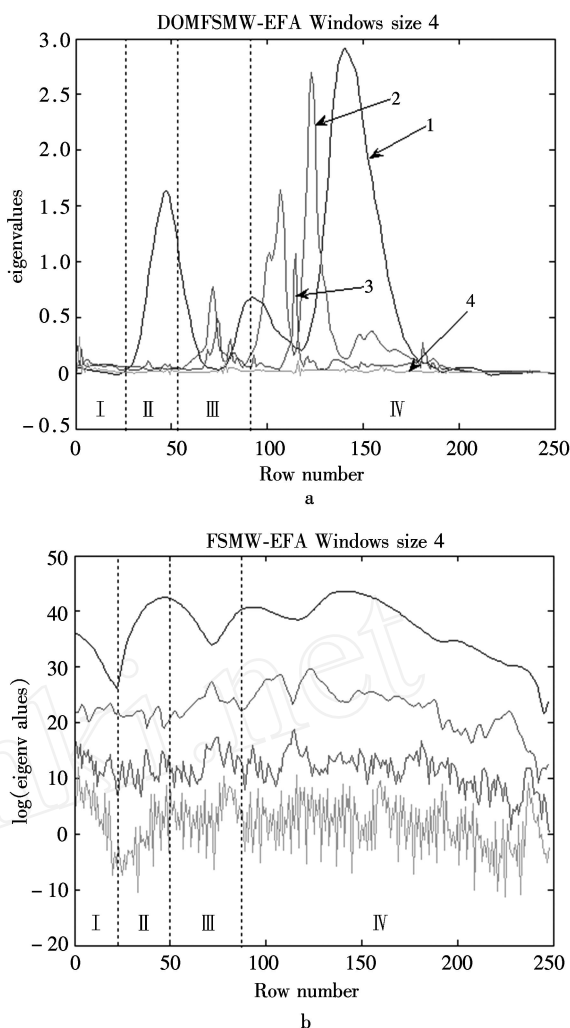


图 4 固定窗口因子分析秩图
a:扣除数量级固定窗口因子分析法;b:固定窗口因子分析法
Fig. 4 The rank profile of fixed size moving window evolving factor analysis
a: DOMFSMW-EFA; b: FSMW-EFA

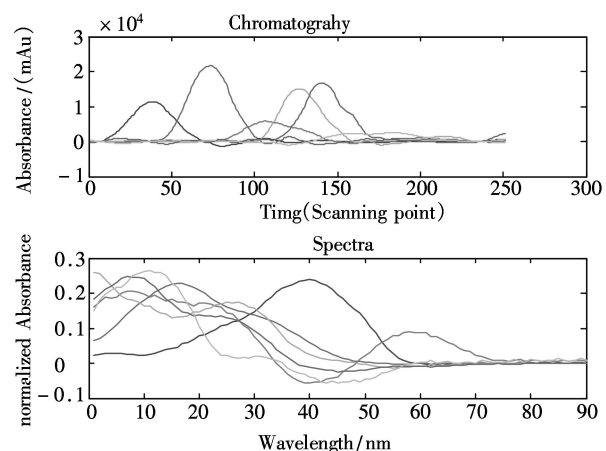


图 5 使用 DOMFSMW-EFA 计算得到的样品的纯物质色谱曲线和光谱曲线
Fig. 5 The pure components' chromatogram and spectra profiles of data obtained by DOMFSMW-EFA

使用数据解析得到的样品中各组分纯物质色谱曲线可以用于定量分析各组分含量;光谱曲线可以

结合标准品光谱信息进行定性分析。从而实现中药等多组分复杂体系的详细定性定量研究。

4 结论

使用固定窗口因子分析 - 直观演进特征投影方法来处理蛤青复方注射液二维色谱 - 光谱数据,可以得到其中纯物质组分的色谱和光谱信息。通过比较可以看出,使用扣除数量级固定窗口因子分析处理的蛤青复方注射液的局部高效液相色谱数据的过程更加简明直观,结果也更加准确。这表明,通过使用联用仪器结合化学计量学方法是一种处理真实未知复杂体系分析的快捷有效方法。使用这些方法,尤其是本文介绍的方法,不但可以提高色谱的分离能力,而且可以很大程度的提高色谱的定性识别能力。使用该方法可以更好的解决复杂体系的分析问题。在中药分析中,该方法可以被用来对体系中的每一个组分进行定性和定量分析。这将使得该方法成为药效学和构效关系研究中的有力工具。并且,该方法不仅可以完成中药色谱局部信息数据的分辨,也可以完成全谱图信息的分辨工作。

但是,尽管已经对该方法进行了改进,在具体使用过程中,对于组分数和选择性区域的判断仍然需要结合多种方法和经验才能达到准确真实。需要更加深入的研究以实现更加方便快捷的数据解析方法。

参考文献:

- [1] WEN K C, HUANG C Y, LU F L. Determination of baicalin and puerarin in traditional Chinese medicinal preparations by high-performance liquid chromatography[J]. J Chromatogr A, 1993, 631:241-250.
- [2] 国家药典委员会. 中华人民共和国药典[M]. 北京:中国医药科技出版社,2000,22:671-672.
- [3] GONG F, LIANG Y Z, XIE P S, et al. Information theory applied to chromatographic fingerprint of herbal medicine for quality control[J]. J Chromatogr A, 2003, 1002:25-40.
- [4] 谢培山. 中药质量控制模式的发展趋势[J]. 中药新药与临床药理,2001,12(3):188-191.
- [5] KVALHEIM O M, LIANG Y Z. Heuristic evolving latent projections: resolving two-way multicomponent data. 1. Selectivity, latent-projective graph, datascope, local. rank, unique resolution[J]. Anal Chem, 1992, 64:936-946.
- [6] LIANG Y Z, KVALHEIM O M, KELLER H R, et al. Heuristic evolving latent projections: resolving two-way multicomponent data. 2. Detection and resolution of minor constituents[J]. Anal Chem, 1992, 64:946-954.
- [7] LIANG Y Z, KVALHEIM O M, RAHMANI A, et al. Resolution of strongly overlapping two-way multicomponent data by means of heuristic evolving latent projections[J]. J Chemom, 1993, 7:15-43.
- [8] BOOKSH K S, KOWALSKI B R. Theory of analytical chemistry[J]. Anal Chem, 1994, 66:A782-A791.
- [9] LI H, HUO J R, WANG K, et al. Resolution of multicomponent overlapped peaks: a comparison of several curve resolution methods[J]. Talanta, 2006, 70(2):336-343.
- [10] MAEDER M. Evolving factor analysis for the resolution of overlapping chromatographic peaks[J]. Anal Chem, 1987, 59:527-530.
- [11] MAEDER M, ZILIAN A. Multivariate technique in chromatography[J]. Chemom Intell Lab Syst, 1988, 3:205-215.
- [12] MALINOWSKI E R. Window factor analysis: theoretical derivation and application to flow injection analysis data[J]. J Chemom, 1992, 6:29-40.
- [13] DEN W, MALINOWSKI E R. Factor analysis of ultraviolet spectra[J]. J Chemom, 1993, 7:89-98.
- [14] MANNE R, SHEN H L, LIANG Y Z, et al. Subwindow factor analysis[J]. Chemom Intell Lab Syst, 1999, 45:171-176.
- [15] SHEN H L, MANNE R, XU Q S, et al. Local resolution of hyphenated chromatographic data[J]. Chemom Intell Lab Syst, 1999, 45:323-328.
- [16] LIANG Y Z, KVALHEIM O M. Diagnosis and resolution of multiwavelength chromatograms by rank map, orthogonal projections and rank analysis[J]. Anal Chim Acta, 1994, 292:5-15.
- [17] LIANG Y Z. White, grey and black multicomponent systems and their chemometric algorithms[M]. Hunan: Hunan Publishing House of Science and Technology, 1996: 177-182.
- [18] LI H, ZHANG F, HAVEL J. Quantification of analytes in overlapping peaks from capillary electrophoresis using multivariate curve resolution-alternating least squares methods[J]. Electrophoresis, 2003, 24:3107-3115.
- [19] ZHANG F, LI H. Resolution of overlapping capillary electrophoresis peaks using chemometrics analysis: quantification of the components in compound reserpine tablets[J]. Electrophoresis, 2005, 26:1692-1702.
- [20] 王志红,张丽. 复方蛤青注射液抗菌作用的研究[J]. 中国药事,2005,19(1):28-32.

(编辑:胡春霞)