

Investigation of Chemometrics Methods in pKa Determination Based on Multivariate Data Analysis

Han Cheng, Xianzheng Zhang

College of Chemistry and Molecular Sciences, Wuhan University, Wuhan Hubei
Email: chm128256@163.com

Received: Oct. 24th, 2018; accepted: Nov. 5th, 2018; published: Nov. 16th, 2018

Abstract

A new approach to teaching basic experimental inorganic and analytical chemistry offered by our college was presented. The introduction of chemometrics method based on multivariate data analysis was of great significance to improve the ability of undergraduate students in understanding and use of analytical technology and statistical tools. Based on this, we designed a set of experiments to determine the absorbance of the acid-base indicators with multi-spectral overlap. The *pH* values were also determined. The dissociation constants were calculated by chemometrics and classical analytical methods. Moreover, the advantages and shortages were also compared and discussed.

Keywords

Multivariate Data Analysis, Chemometrics, Teaching Research

基于多元分析的化学计量学方法用于pKa测定的研究

程 翰, 张先正

武汉大学化学与分子科学学院, 湖北 武汉
Email: chm128256@163.com

收稿日期: 2018年10月24日; 录用日期: 2018年11月5日; 发布日期: 2018年11月16日

摘 要

无机及分析化学基础实验是我院开设的一门本科实验课程, 在本课程中引入化学计量学中多元分析的方

法用于实验教学,培养学生了解和使用分析统计工具的能力,对学生的综合素质的提高是有重要意义的。我们据此设计了一组实验,在选定的具有多光谱重叠的酸碱指示剂质子平衡体系中,通过UV-Vis分光光度法来测定酸碱指示剂体系的吸光度,同时测定其pH值,使用化学计量学计算其解离常数,并与经典分析方法进行比较,探讨其优缺点。

关键词

多元分析, 化学计量学, 教学研究

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

化学平衡是在本科化学教学中广泛涉及的概念。随着时代的发展,在实际研究中发现,面对复杂的化学平衡体系时我们需要更加先进的分析处理工具。在上世纪 70 年代,有关化学计量学的概念出现了,这是一门结合化学和统计学的新学科。通过结合统计学的分析计算方法,为化学体系的数据处理分析打造了强有力的工具。自上世纪 90 年代以来,随着计算机技术的高速发展,化学计量学在科学研究和工业设计中开始得到广泛应用,涉及到本领域相关研究的论文逐渐增多[1][2]。起初化学计量学只是作为一种化学分析方法,为化学测量提供理论和方法,随着相关研究的推进,目前已经建立并发展出三大方向:实验方案的设计和优化,模式识别以及多元变量分析[3],涵盖了化学测量的全过程,包括采样理论与方法,试验设计与过程优化控制,化学信号处理,分析信号的校正与分辨,化学模式识别,化学过程的计算机模拟,化学数据库等。实验方案的设计和优化是通过评估变量对分析响应的影响,确定变量的重要程度。有效的实验规划和优化,可以避免大量的重复实验和试剂的浪费[4]。模式识别技术,它旨在检测和分析一组样品的大量数据(称为“训练”或“学习”),确定其趋势和规律,为以后相似性的样品分类和鉴定提供参考依据[5]。而多元变量分析技术,则以化学量测数据构造多元校正模型,能够同时综合处理分析样本的多个自变量以及相关的因变量,获取变量间的联系和数据结构等信息,是化学计量学的核心组成部分[6]。目前常用的多元变量分析方法主要有自适应卡尔曼滤波法(Kalman filtering, KF),多元线性回归(multiple linear regression, MLR) [7],主成分分析(principal component analysis, PCA) [8],偏最小二乘法(partial least squares, PLS) [9]等,其中主成分分析 PCA 是一种通过正交变换将一组可能存在相关性的变量转换为一组线性不相关的变量(称为主成分)来表示,使它们尽可能多地保留原变量的信息,以便找到原变量之间的关系的多元变量分析方法。

综上所述,化学计量学在本科化学教学和基础科研中具有重要的意义[10]。为了促进化学计量学的应用和推广,我们提出一种旨在将实验性课程纳入教学的设想,增强学生对化学计量学的理解,激发学生的学习兴趣[11]。同时也可在化学本科课程已经提供的实验学科中纳入涉及化学计量学的实践。此外,通过数据分析软件对化学体系的评估,给学生展示了一种先进的数据处理技术[12]。

基于此目的,我们选择使用化学计量学方法和经典分析化学方法分别对酸碱指示剂的电离平衡进行分析处理,并将结果进行比较。对于某一物质而言,去质子化是指该物质分子中脱去质子生成其共轭碱的过程,酸性越强的物质越容易去质子化,酸平衡常数 K 反映了该过程的趋势。用 pK_a ($pK_a = -\log K_a$) 来表示。我们通常使用电位测定法和 UV-Vis 分光光度法测定 pK_a 。然而在酸碱指示剂复杂的电离水解平

衡体系中, 存在多种平衡并且具有彼此接近的 pK_a 值, 并且酸式体和碱式体吸收峰还表现出较大程度的重叠的现象。在测定这类复杂的体系的 pK_a 时, 经典分析化学方法往往仅通过测定最大吸收波长处的吸光度与对应 pH 值来计算 pK_a , 这不能完全反映整个系统的真实情况。因此我们设计了一个使用化学计量学的多元变量分析方法来计算 pK_a 的实验方案, 并将化学计量学分析所得结果与经典分析化学方法和来自文献的数据进行比较以证明所提出的方法的可靠性。

2. 原理与方法

对于酸碱平衡体系, 物质浓度与 pH 值的关系根据亨德森 - 哈塞尔巴尔赫方程(Henderson-Hasselbalch equation) (公式(1))计算。

$$pH = pK_a + \log\left(\frac{[A^-]}{[HA]}\right) \quad (1)$$

本实验中所用溶液均为稀溶液, 可认为其活度等于浓度。由朗伯-比尔定律(Lambert-Beer law)可知浓度与吸光度成正比。通过在不同 pH 值的条件下测定已知样品的吸收光谱可计算得到其 pK_a 。

2.1. 经典分析化学方法

指示剂的酸式体和碱式体在可见光区均有较强的吸收, 随着 pH 的变化, $[A^-]/[HA]$ 也发生变化, 溶液的吸光度也必然变化, 故吸光度与 $[A^-]/[HA]$ 有着线性关系, 由公式(1)可知, 当 $[A^-]/[HA] = 1$ 时, $pH = pK_a$ 。设在某一波长处, 指示剂全以酸式体 HA 存在时吸光度为 A_a , 全以碱式体 A^- 存在时吸光度为 A_b , 以两者共存时吸光度为 A , 根据 Lambert-Beer law 可得 $\log([A_a - A]/[A - A_b]) = pH - pK_a$, 以吸光度与对应 pH 值作图, 其结果是具有 S 形曲线(“S”或“S”反转)的曲线, 由以上公式可知发生凹面反转的 S 形曲线的中点即为 pK_a 。

具体方法如下: 取两份等量指示剂, 加入缓冲溶液调节其 pH 值, 使之分别处于其变色区间的上下限之外, 测定其吸光度, 以吸光度为纵坐标, 波长为横坐标在同一张坐标图上作图, 选取两者吸光度相差最大的波长定为测定波长, 接下来在其变色区内配制一系列不同 pH 值的溶液, 加入 $NaCl$ 溶液调节离子强度使之相同, 测量的吸光度与对应的 pH 值, 作图得到 pK_a 。

2.2. 化学计量学方法

经典分析化学方法和化学计量学方法的主要区别在于前者以单一波长的单变量为研究对象而后者同时分析不同波长下吸光度与 pH 的关系, 即通过 PCA 法进行每个溶液样本中吸光度(光谱)组和其各自的 pH 值的多变量分析。主成分分析 PCA 是用于数据处理的最常见的多元校正方法, 在减少需要分析的数据量同时, 尽量减少原数据包含信息的损失, 以达到对所收集数据进行全面分析的目的。由于各变量间存在一定的相关关系, 因此有可能用较少的综合指标分别综合存在于各变量中的各类信息。以三变量之间的相互关系为例, 若用自然坐标系(x, y, z)这三个轴来表示这组数据的话, 需要使用三个维度, 通过把 x, y, z 坐标系旋转使数据所在平面与(x, y)平面重合, 则这组数据可转换为用 x' 和 y' 两个维度表示, 这两个轴的信息即为数据的主成分。若这组数据在 z'轴有一个很小的抖动则会导致数据在 z'轴不完全相关, 然而 z'轴上的分布与原点构成的夹角非常小, 即在 z'轴上仍有很大的相关性, 综上可以认为数据在(x' , y')轴上的投影构成了数据的主成分, 可用上述的二维坐标表示这些数据。

在本实验中, 我们采集一系列不同 pH 值溶液的吸光度数据, 构建了一个 20×301 的数据矩阵, 其中前者为采集的 pH 值的数目, 后者为谱图的波长数目(350~650 nm)。

应用 PCA 方法分析所得数据, 得到每种物质不同酸式体, 碱式体对应的归一化浓度(相对浓度)与 pH 值的曲线, 其交叉点对应 pH 值即为 pK_a 。同时可以得到某一特定 pH 值条件下, 体系中每种物质的真实

分数。

3. 实验部分

3.1. 主要仪器与试剂

仪器:

PerkinElmer Lambda35 型 UV/VIS 分光光度计(美国珀金埃尔默仪器有限公司); Sartorius PB10 型数字式 pH 计(德国赛多利斯公司); DC-1015 恒温水浴槽(宁波海曙赛福实验仪器公司)。

试剂:

甲基红(Methyl red: 4-二甲氨基偶氮苯-2'-羧酸(CH_3)₂NC₆H₄N = NC₆H₄CO₂H), 乙基曙红(Ethyl eosin: 2',4',5',7'-四溴荧光素乙酯钾盐 C₂₂H₁₁Br₄KO₅), 曙红(Eosin: 2',4',5',7'-四溴荧光素 C₂₀H₈Br₄O₅), 氯化钠, 氢氧化钠, 盐酸, 磷酸钠, 柠檬酸, 乙醇, 去离子水。所有实验试剂均为分析纯, 购自 Sigma-Aldrich 公司。

3.2. 实验方法

3.2.1. 溶液配制

参照文献分别配制一系列不同 pH 值下的酸碱指示剂溶液[13] [14], 每种指示剂其总浓度保持一致, 吸光度范围均符合朗伯 - 比尔定律, 其中用于测定吸光度的甲基红溶液浓度为 $2.0 \times 10^{-5} \text{ mol}\cdot\text{L}^{-1}$, 乙基曙红, 曙红的浓度为 $8.5 \times 10^{-6} \text{ mol}\cdot\text{L}^{-1}$; 各种缓冲溶液均按分析化学标准配制。

3.2.2. pH 和离子强度

pH 小于 2.0 的溶液通过加入微量的盐酸溶液进行调节; pH 在 2.0~8.0 的范围内的溶液使用 McIlvaine 缓冲液来调节; 使用 $0.05 \text{ mol}\cdot\text{L}^{-1}$ 的氯化钠溶液来调节离子强度。

3.2.3. 吸收光谱和 pH 值得测量

通过恒温槽循环水泵, 调控至实验温度并恒温 15 min。分光光度计扫描范围 350~650 nm, 采样间隔为 1 nm, 分别测量上述系列样品溶液的吸收光谱。

用标准溶液标定并校正数字式 pH 计; pH 计的标定和待测液 pH 的测量均在恒温条件下进行, 取 3 次测量平均值。

3.3. 数据处理

实验数据使用 Origin, SPSS, ChemPattern 等软件进行分析处理。

4. 结果与讨论

本实验研究对象均为实验教学常规使用的指示剂, 通过使用化学计量学方法来分析和计算所测相关数据, 并与经典分析化学方法所得数据以及文献报道的数据进行比较来验证化学计量学方法的置信度。

4.1. 指示剂的结构特征

甲基红(图 1(I)): 甲基红在酸性条件(pH 值低于 4.4)下呈红色, 中性和碱性条件下(pH 值高于 6.2)呈黄色。曙红和乙基曙红属于氧杂蒽类且呈现出复杂的酸碱平衡体系(图 1(II)): 乙基曙红, 图 1(III): 曙红。对曙红和乙基曙红而言, 除乙基曙红由于酯键的存在而不存在 $\text{pK}_{\text{a,COOH}}$ 以外, 氧杂蒽类指示剂在其结构中除了羧基之外还具有两个酚基, 如图 1 所示, 其分子结构中的溴取代基可增加其酸度。对于甲基红和乙基曙红来说, 存在一个 pK_{a} 。而曙红存在两步电离, 因此具有两个 pK_{a} 。

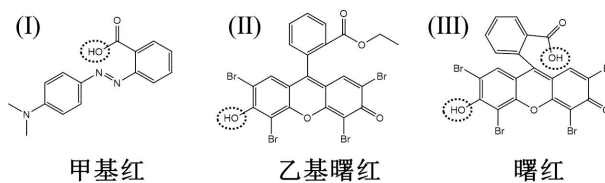


Figure 1. Structure of the pH indicators
 图 1. 指示剂分子结构

4.2. 数据分析

4.2.1. 经典分析化学方法

我们分别将甲基红, 乙基曙红和曙红配制成一系列不同 pH 值的溶液, 测定其吸收光谱。如图 2 中所示。甲基红(图 2(I))在酸性条件下其溶液呈红色($\lambda_{\max} = 530 \text{ nm}$), 并且随着 pH 增加, 溶液变成黄色($\lambda_{\max} = 427 \text{ nm}$)。乙基曙红在低 pH 值的条件下溶液几乎是无色的。随着 pH 值的增加, 溶液变为粉红色, 在 520 nm 处出现最大吸收。它的结构比甲基红复杂, 可观察到质子化和去质子化的谱带叠加。

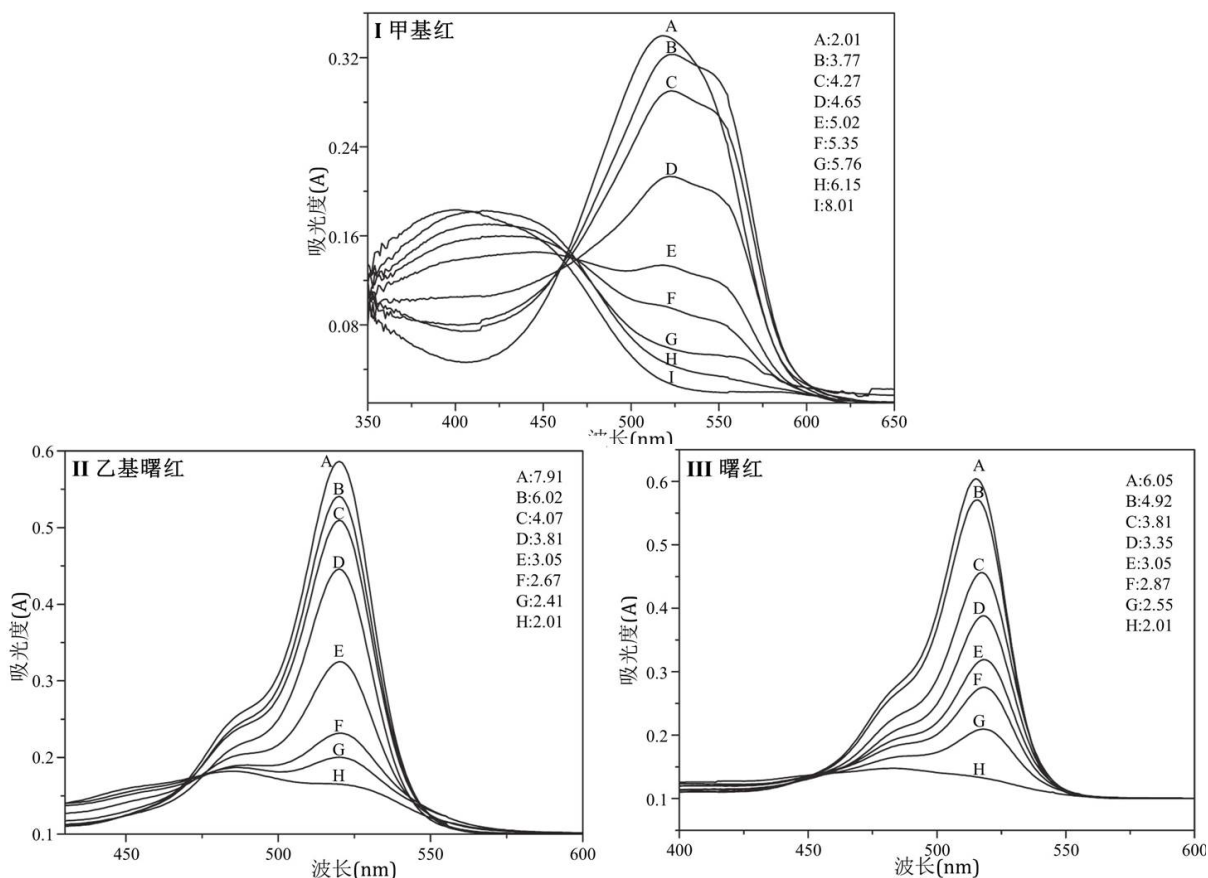


Figure 2. Absorption spectra of the indicators under conditions of different pH values. (I) Methyl red, (II) Ethyl eosin, (III) Eosin
 图 2. 不同 pH 条件下指示剂的吸收光谱。其中(I) 甲基红, (II) 乙基曙红, (III) 曙红。图中 A~H 分别表示待测液的 pH 值

对于曙红, 可以观察到两个波段(图 2(III)), 在碱性条件下(溶液显粉红色)在 517 nm 处有最大吸收, 而实际上在酸性条件下, 在 480 nm 区域也有一相对不明显的最大吸收, 溶液几乎没有无色。尽管它的质子化形式更多(图 1(III)中的结构), 由曙红(图 2(III))呈现的光谱轮廓与乙基曙红(图 2(II))获得的光谱轮廓

几乎相同, 因为两者结构类似, 生色团是相同的。

根据甲基红的分光光度数据, 在 530 nm 处构建吸光度与 pH 值的函数(图 3(I))。得到具有 S 形特征的曲线, 其拐点值即为甲基红的 pKa。以 520 nm 处吸光度与 pH 的函数曲线, 拐点即为乙基曙红的 pKa 值(图 3(II))。同样通过经典分析化学方法所得中曙红(图 3(III))的 pH-吸光度曲线却错误地表明仅存在一个质子化结构, 导致仅得到单一 pKa。

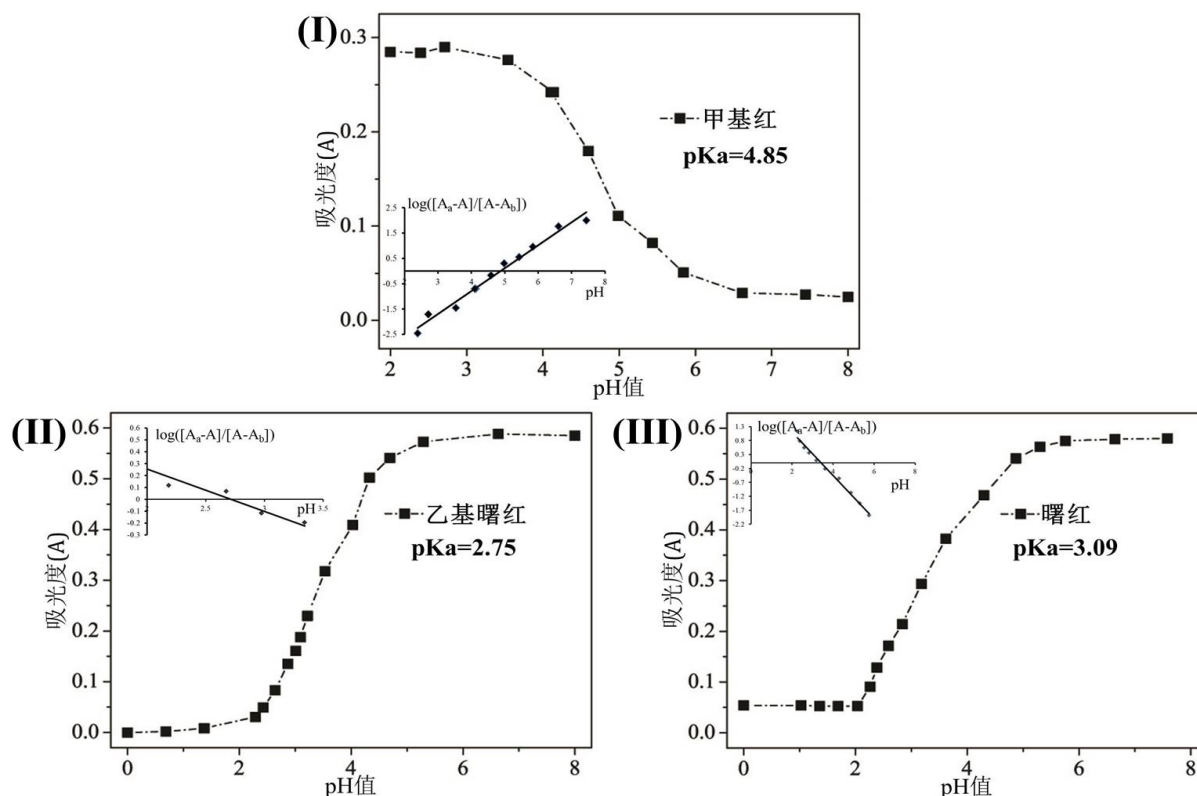


Figure 3. Absorption spectra of the indicators under conditions of max wavelength. (I) Methyl red, (II) Ethyl eosin, (III) Eosin
图 3. 不同指示剂的在最大吸收波长下 pH 值与吸光度的曲线。其中(I) 甲基红, (II) 乙基曙红, (III) 曙红

因此将经典分析化学中固定最大吸收波长的单变量分析法应用于存在两个或多个 pKa 的指示剂体系, 加之其强烈的光谱重叠, 会导致错误的计算结果。

4.2.2. 用于测定 pH 值的化学计量方法

对于甲基红, 在测定的 pH 范围内, 溶液中有酸式体和碱式体两种存在形式, 如前所述(图 2(I))。主成分分析的结果表明, 两个变量(对应于酸式体和碱式体的参与)解释的方差分别为 81.62%和 18.19%, 即主成分累计贡献率为 99.81%, 第三个变量描述的信息量只有 0.13%的贡献率。如表 1 中所列。根据甲基红的化学结构特征和第三个非常小的变量的贡献率, 我们可忽略了第三个变量, 以两变量体系来收集和处理数据。

Table 1. Variance and cumulative contribution of Methyl red

表 1. 甲基红主成分分析方差贡献率及累计贡献率

成分	方差贡献率	累计贡献率
1	81.62%	81.62%
2	18.19%	99.81%
3	0.13%	99.94%

使用同样方法对曙红和乙基曙红进行分析, 计算得到 pK_a 值。图 4 中显示了各指示剂的不同质子化形式在额定 pH 范围内的组成分数变化。

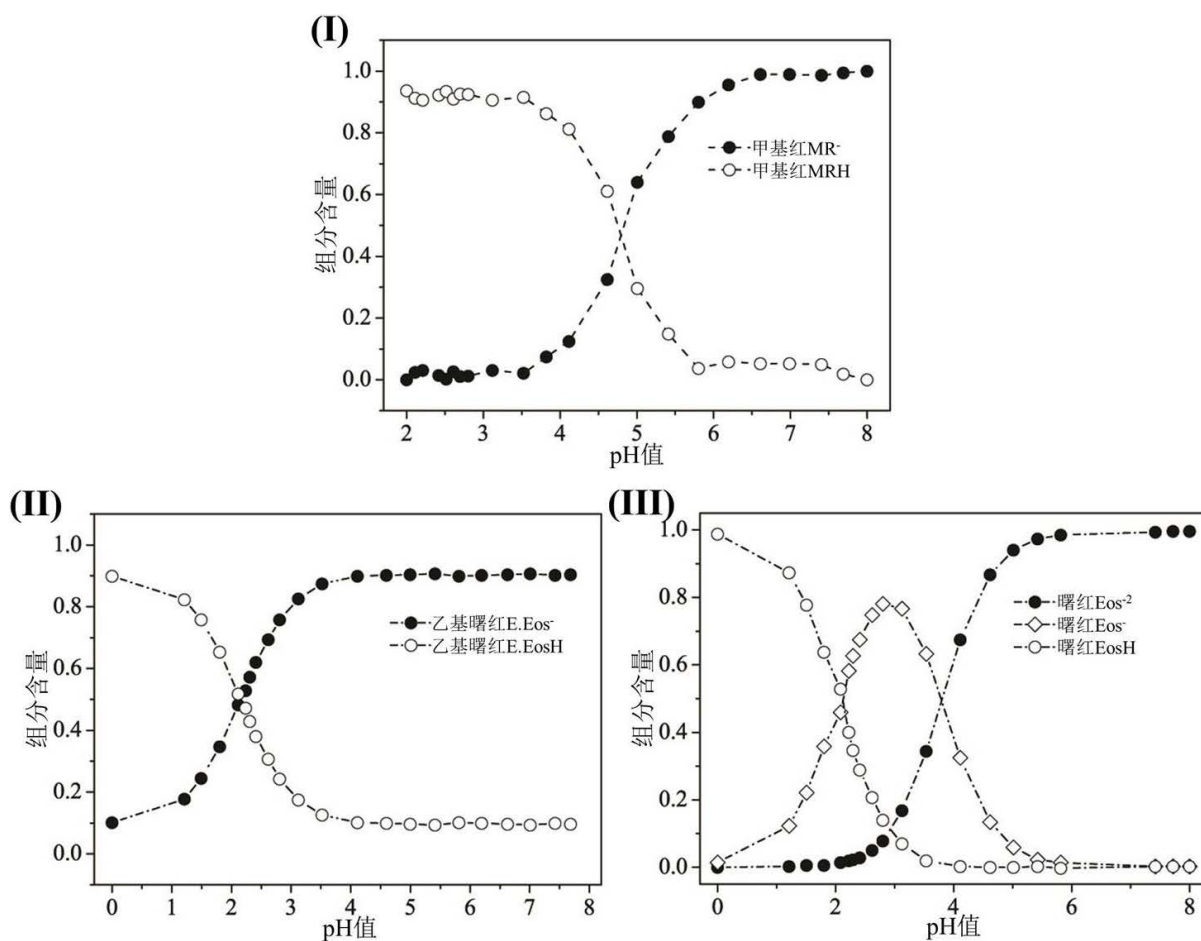


Figure 4. Component content of the pH indicators. (I) Methyl red, (II) Ethyl eosin, (III) Eosin

图 4. 指示剂组分含量。其中(I) 甲基红, (II) 乙基曙红, (III) 曙红

在表 2 中我们分别列出了通过化学计量学和经典分析化学方法计算所得指示剂的 pK_a 以及文献报道数据[15] [16]。对于甲基红而言, 其电离平衡相对简单, 通过经典分析化学和化学计量学方法所得 pK_a 值之间没有显著差异, 同时与文献报道数据也非常接近。从其分子结构可知, 由于甲基红只有两种质子化形式, 上述方法均可得到可靠的结果。

Table 2. pK_a obtained experimentally and comparison with the literature

表 2. 文献报道与实验所得 pK_a 的比较

成分	文献报道	经典分析方法	化学计量学方法
甲基红	5.02	4.85	4.95
乙基曙红	1.9	2.75	2.17
曙红	2.02	3.09	2.06/3.72

对于乙基曙红而言, 化学计量学所得结果与经典分析化学方法有显著区别, 化学计量学结果与文献报道更为接近。尽管它仅存在一个质子平衡, 但乙基曙红在其质子形式(图 3(III))之间展现出比甲基红(图

3(I)更强烈的光谱重叠。因此化学计量学方法更加可靠, 该方法综合考虑了各种波长下的溶液各组分的贡献。

对于曙红体系而言, 经典分析化学法仅得到一个 pK_a 值, 而化学计量学可得两个值。由于曙红体系除了具有多个质子化平衡之外, 指示剂本身呈现出强烈的光谱重叠。化学计量学方法在面对此类复杂的体系进行 pK_a 测定时明显优于经典分析方法, 利用化学计量学方法, 可以解决经典分析方法中存在的可质子化基团数目与 pK_a 值不一致问题。此外, 在表 2 中通过与乙基曙红($pK_{a,OH} = 2.17$)的比较可以发现, 曙红酚基的与之类似($pK_{a,OH} = 2.06$), 而它结构中羧基酸性较弱($pK_{a,COOH} = 3.72$)。这是由于氧杂蒽环上溴离子的强吸电子效应导致了酚基更强的酸度。曙红的高光谱重叠使经典分析化学中单变量法对曙红的 pK_a 测定无能为力。多变量分析的重要性更为明显, 化学计量学方法更能体现不同 pH 条件下所有波长的原始吸收光谱的数据信息。

5. 结论

虽然化学计量学已成为工业和研究领域大量使用的重要分析工具, 但在本科阶段的理论和实验教学中还没有得到广泛应用, 这是由于相对于经典分析化学方法, 化学计量学方法在数据处理环节相对复杂, 需要学生具备相关的数学分析, 计算机处理等理论知识, 这也限制了其应用的开展。然而通过测定酸碱指示剂 pK_a 的实验表明类似体系适合使用化学计量学方法, 特别是在具备以下特点的体系中: 1) 谱图中谱带间重叠度高; 2) 分析信号质量低(信噪比小于 1, 或基线对光谱的影响); 3) 化学平衡之间的干扰(例如, 接近 pH 值存在各种平衡)。我们发现在此类复杂体系中, 化学计量学方法比经典分析化学测定方法效率更高。

基金项目

本文研究项目经费来自国家自然科学基金面上项目(51573142)。

参考文献

- [1] 梁逸曾, 吴海龙, 沈国励, 等. 分析化学计量学的若干新进展[J]. 中国科学 B 辑, 2006, 36(2): 93-100.
- [2] 潘荣荣, 曲刚莲, 马果花, 等. 化学计量学在分析化学中的应用[J]. 化学分析计量, 2007, 16(2): 76-78.
- [3] Poppi, R.J. and Pasquini, C. (1993) Spectrophotometric Determination of a Mixture of Weak Acids Using Multivariate Calibration and Flow Injection Analysis Titration. *Chemometrics and Intelligent Laboratory Systems*, **19**, 243-254. [https://doi.org/10.1016/0169-7439\(93\)80108-T](https://doi.org/10.1016/0169-7439(93)80108-T)
- [4] Leardi, R. (2009) Experimental Design in Chemistry: A Tutorial. *Analytica Chimica Acta*, **652**, 161-172. <https://doi.org/10.1016/j.aca.2009.06.015>
- [5] Rencher, A.C. and Christensen, W.F. (2012) *Methods of Multivariate Analysis*. 3rd Edition, Wiley, New York. <https://doi.org/10.1002/9781118391686>
- [6] Leardi, R. (2001) Genetic Algorithms in Chemometrics and Chemistry: A Review. *Journal of Chemometrics*, **15**, 59-569. <https://doi.org/10.1002/cem.651>
- [7] Preacher, K.J., Curran, P.J. and Bauer, D.J. (2006) Computational Tools for Probing Interactions in Multiple Linear Regression, Multilevel Modeling, and Latent Curve Analysis. *Journal of Educational and Behavioral Statistics*, **31**, 437-448. <https://doi.org/10.3102/10769986031004437>
- [8] Jolliffe, I.T. (2002) *Principal Component Analysis*. 2nd Edition, Springer, New York.
- [9] 赵静. 化学计量学多元校正技术结合紫外分光光度法在药品检测中的应用研究[D]: [硕士学位论文]. 南宁: 广西大学, 2009.
- [10] 房靖. 化学计量学在分析化学中的应用及发展前景[J]. 科技资讯, 2010(21): 242-242.
- [11] 许珑耀. 化学计量学在分析化学中的应用研究[J]. 城市建设理论研究, 2016(9): 1885-1885.
- [12] 万鹏. 化学计量学在分析化学中的应用研究[J]. 中国新技术新产品, 2016(2): 54.
- [13] Foulona, C., Duhal, N., Lacroix-Callens, B., et al. (2007) Determination of pK_a Values of Benzoxa-, Benzothia- and Benzoselena-Zolinone Derivatives by Capillary Electrophoresis: Comparison with Potentiometric Titration and Spectrometric Data. *European Journal of Pharmaceutical Sciences*, **31**, 165-171. <https://doi.org/10.1016/j.ejps.2007.03.002>

-
- [14] Barbosa, J., Barrón, D., Jiménez-Lozano, E. and Sanz-Nebot, V. (2001) Comparison between Capillary Electrophoresis, Liquid Chromatography, Potentiometric and Spectrophotometric Techniques for Evaluation of pKa Values of Zwitterionic Drugs in Acetonitrile-Water mixtures. *Analytica Chimica Acta*, **437**, 309-321.
[https://doi.org/10.1016/S0003-2670\(01\)00997-7](https://doi.org/10.1016/S0003-2670(01)00997-7)
- [15] Srour, R.K. and McDonald, L.M. (2008) Determination of the Acidity Constants of Methyl Red and Phenol Red Indicators in Binary Methanol- and Ethanol-Water Mixtures. *Journal of Chemical & Engineering Data*, **53**, 116-127.
<https://doi.org/10.1021/jc700413x>
- [16] Vanzin, D., Freitas, C.F., Pellosi, D.S., *et al.* (2016) Experimental and Computational Studies of Protolytic and Tautomeric Equilibria of Erythrosine B and Eosin Y in Water/DMSO. *RSC Advances*, **6**, 110312-110328.
<https://doi.org/10.1039/C6RA12198E>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2163-1557, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: aac@hanspub.org