第38卷，第8期
2018年8月

光谱学与光谱分析
Spectroscopy and Spectral Analysis

Vol. 38, No. 8, pp2451-2455
August, 2018

应用荧光相关光谱测定L-色氨酸的浓度

顾　颂[1,2]，朱焯炜[1,2]，马超群[1,2]，陈国庆[1,2*]

1. 江南大学理学院，江苏 无锡　214000
2. 江苏省轻工光电工程技术研究中心，江苏 无锡　214000

摘　要　应用FLS920P型荧光光谱仪对L-色氨酸溶液进行三维荧光光谱检测，从中发现：L-色氨酸的特征荧光峰位于270/350 nm。设定发射波长为350 nm，测量激发谱。由测量结果发现在250～260 nm区间，谱线斜率较大、线性度好。因此选取250，255和260 nm三个激发波长，在每个激发波长下分别测量相应的荧光发射谱。基于三条不同的荧光发射谱，构建以激发波长为外扰变量的自相关光谱；而以浓度为外扰变量的自相关光谱，是以超纯水在不同激发波长的平均谱作为参考光谱，通过参考光谱与样本平均谱的相关计算得到。在此基础上，将相关光谱数据分别与偏最小二乘回归(PLSR)和径向基神经网络(RBFNN)相结合，建立溶液中L-色氨酸含量的预测模型，研究结果表明：采用浓度为外扰变量构造的荧光相关光谱信噪比较高，建模的预测效果要好；而在外扰变量相同时，基于径向基神经网络建立的预测模型比基于偏最小二乘回归建立的预测模型对溶液中L-色氨酸浓度的预测结果更为准确。其中，以浓度为外扰变量时的径向基神经网络预测模型准确度最高，该模型的预测相关系数为99.91%，预测均方根误差为0.033 $\mu g \cdot mL^{-1}$。研究结果表明，使用该方法能够对溶液中的物质含量进行准确测定，可为食品安全监管提供帮助。

关键词　L-色氨酸；荧光相关光谱；偏最小二乘回归；径向基神经网络
中图分类号：O657.3　　**文献标识码**：A　　**DOI**：10.3964/j.issn.1000-0593(2018)08-2451-05

引　言

色氨酸是维持人体生命活动所必需的一种氨基酸。目前，测定色氨酸含量的方法有很多，主要包括：高效毛细管电泳法、高效液相色谱法、紫外光谱和荧光光谱。光谱法相对于前两者来说，具有操作简单、快速有效等优点；而荧光光谱相对于紫外光谱而言，具有信息量丰富、灵敏度高等优点。由于三维荧光光谱含有大量的数据信息，而二维相关光谱能凸显外部扰动引起的细微变化，提高光谱的分辨率。因此，本文将两者用于溶液中L-色氨酸浓度的测定。

目前，二维相关光谱在红外光谱、判别分析[1-4]等方面研究较多，但在荧光光谱和含量检测[5]方面报道较少，有待深入。本文以L-色氨酸溶液为研究对象，分别从外部扰动和化学计量学算法等方面，对使用荧光相关光谱建立定量模型的预测效果进行研究，可为使用二维相关光谱技术[1-9]进行含量检测的研究提供帮助。

1　实验部分

1.1　仪器与材料

仪器：英国FLS920P型稳态荧光光谱仪(Edinburgh instruments，Edinburgh)、微量电子天平(Mettler Toledo，Switzerland)、移液枪(0～5 mL)。

表1　校正集和预测集样本中L-色氨酸的浓度
Table 1　The calibration and prediction concentration of the of L-Tryptophan

校正集	浓度/($\mu g \cdot mL^{-1}$)	预测集	浓度/($\mu g \cdot mL^{-1}$)
1	0.375	7	0.50
2	0.75	8	1.00
3	1.25	9	1.50
4	1.75	10	2.00
5	2.50	11	2.25
6	3.00	12	2.75

收稿日期：2016-12-20，**修订日期**：2017-06-22
基金项目：国家自然科学基金项目(61178032，61378037)，中央高校基本科研业务费专项资金项目(JUSRP51628B)资助
作者简介：顾　颂，1991年生，江南大学理学院硕士研究生　　e-mail：2546998952@qq.com
*通讯联系人　　e-mail：cgq2098@163.com

材料：L-色氨酸由北京北纳创联生物技术研究院提供，纯度为 98.56%；溶剂为超纯水。

1.2 样品的制备

使用标准品和超纯水配置 L-色氨酸浓度范围为 0.375～3 $\mu g \cdot mL^{-1}$ 的 12 个溶液样本，从中抽取 6 个作为校正集，其余作为预测集，样本浓度如表 1 所示。

1.3 L-色氨酸的荧光光谱测定

通过对该溶液体系的三维荧光光谱检测，可以看出：L-色氨酸的荧光峰位于 270/350 nm；从发射波长为 350 nm 的激发光谱中，我们可以发现：在激发波长为 250～260 nm 区间内，曲线斜率大而平稳，荧光强度变化较快。因此，选取该波段的激发波长作为计算荧光相关光谱的外扰变量。L-色氨酸溶液的三维荧光光谱和激发光谱分别如图 1(a)和 1(b)所示。采集超纯水及所有样品在激发波长为 250，255 和 260 nm 的荧光发射谱，发射波长的测量范围为 300～450 nm，步长为 1 nm。激发和发射的分辨率均设置为 5 nm^{-1}，积分时间设置为 0.1 s。

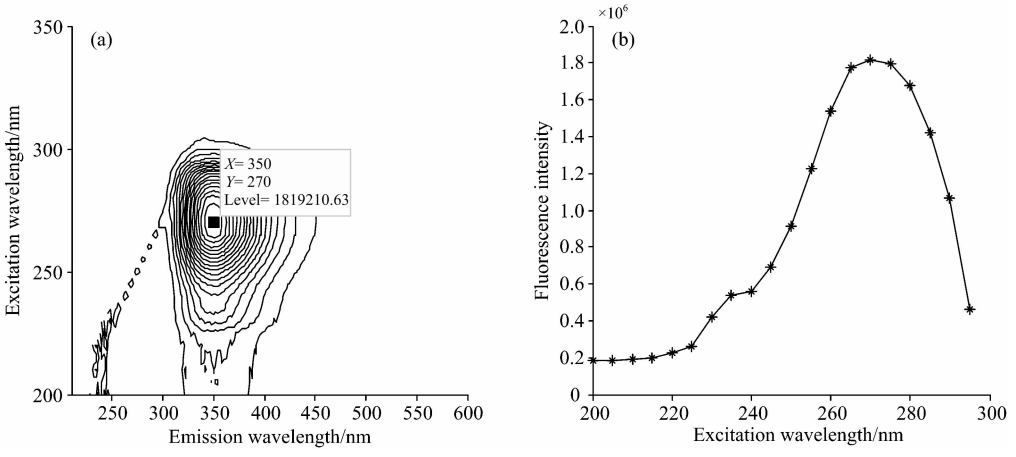

图 1 (a) L-色氨酸溶液的三维荧光光谱；(b) 发散波长为 350 nm 的激发谱

Fig. 1 (a) Three-dimensional fluorescence spectrum of L-Tryptophan solution; (b) Excitation spectrum with emission wavelength at 350 nm

1.4 数据处理

使用由日本 Kwansei-Gakuin 大学 Daisuke Adachi 编写的 2Dshige 软件，将不同激发波长的荧光发射谱作为输入信息进行相关计算，得到以激发波长为外扰变量的自相关光谱；而以浓度为外扰变量的自相关光谱，是以超纯水的平均谱作为参考光谱，通过背景光谱与样本平均谱的相关计算得到。文中采用的平均谱为不同激发波长的平均谱，而非同一激发波长多次测量的平均谱，这是因为前者相对于后者不仅可以有效避免荧光漂白现象的干扰，同时也可以节约样品，缩减工作量。图 2(a)和 2(b)分别表示以激发波长和浓度作为外扰变量，通过相关计算得到的自相关光谱。

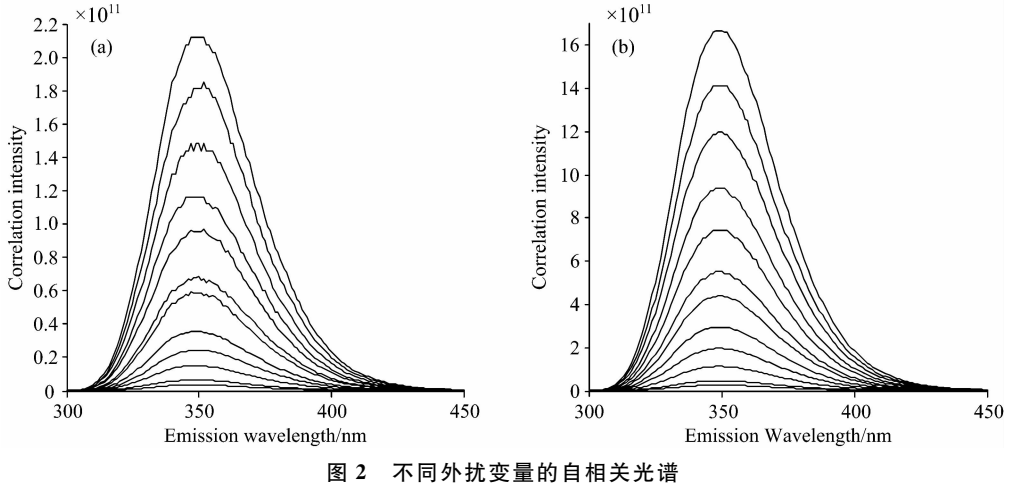

图 2 不同外扰变量的自相关光谱

(a)：激发波长；(b)：浓度

Fig. 2 Auto correlation spectra with the different disturbance variable

(a): by excitation wavelength; (b): by concentration

2　结果与讨论

2.1　应用偏最小二乘回归测定 L-色氨酸的浓度

自相关光谱是由外部扰动引起经动态信号自身相关得到的，能够有效提取同步相关光谱中的特征信息，避免信息的冗余。偏最小二乘回归[10](partial least squares regression, PLSR)采用非线性迭代偏最小二乘算法对变量信息进行分解，同时集结主成分分析、多元线性回归、典型相关分析于一身，从而保证自变量和因变量两者的最大相关性。因此，本文将采用不同外扰方式计算得到的自相关光谱数据与偏最小二乘回归算法相结合，建立对溶液中 L-色氨酸浓度的预测模型，模型的预测结果如表 2 所示。

表 2　PLSR 模型对样中 L-色氨酸浓度的预测结果

Table 2　Prediction results of PLSR model for the concentration of L-Tryptophan

预测集	实际值 /(μg·mL^{-1})	激发波长		浓度	
		预测值 /(μg·mL^{-1})	相对误差 /%	预测值 /(μg·mL^{-1})	相对误差 /%
7	0.5	0.647	29.413	0.555	11.093
8	1.0	0.865	−13.459	0.879	−12.091
9	1.5	1.321	−11.908	1.321	−11.928
10	2.0	1.954	−2.275	1.862	−6.917
11	2.25	2.335	3.790	2.314	2.832
12	2.75	2.679	−2.599	2.878	4.646
r	—	—	98.844	—	98.364
RMSEP	—	—	0.120	—	0.276

采用激发波长构建的自相关光谱结合偏最小二乘回归算法建立的预测模型，在主成分为 2 时，预测结果最好：对未知样品的预测相关系数(r)为 98.844%，预测均方根误差(RMSEP)为 0.120 μg·mL^{-1}；而使用 L-色氨酸浓度构造的自相关光谱建立的预测模型，在主成分为 3 时，预测效果最好：对未知样品的预测相关系数为 98.364%，预测均方根误差为 0.276 μg·mL^{-1}。图 3(a)和(b)分别是在不同外扰方式下，偏最小二乘回归模型对溶液中 L-色氨酸浓度预测值和实际值的线性拟合。从表 2 和图 3 中我们不难看出：两种预测模型的相对误差具有相同的规律，在浓度较低时，预测样本的误差相对较大，而在浓度较高时，预测样本的误差相对较小。这说明，荧光相关强度与溶液的浓度在预测的浓度区间内，并不完全满足线性关系。

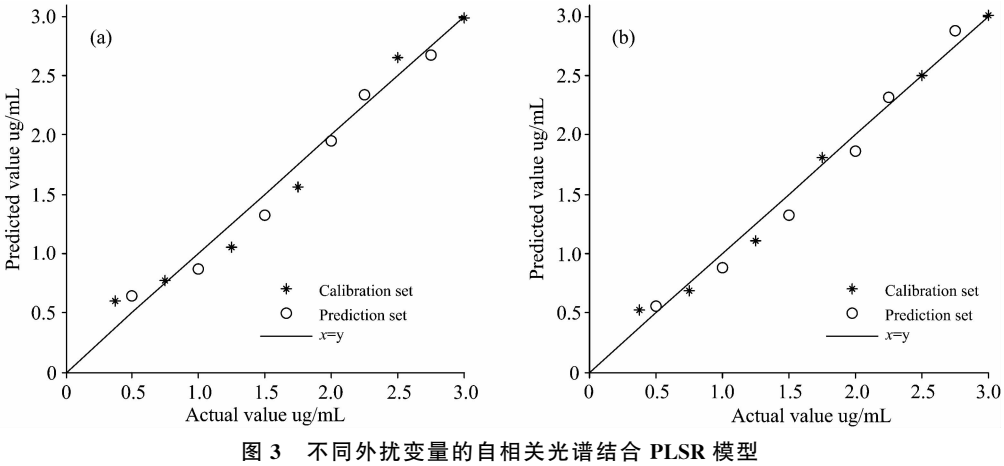

图 3　不同外扰变量的自相关光谱结合 PLSR 模型

(a)：激发波长；(b)：浓度

Fig. 3　Auto correlation spectra with the different disturbance variable combined with PLSR

(a)：by excitation wavelength；(b)：by concentration

2.2　应用径向基神经网络测定 L-色氨酸的浓度

径向基神经网络[11](radial basis function neural network, RBFNN)是一种三层静态前馈型神经网络，相对于 BP 神经网络而言，具有局部逼近的特点，能以任意精度逼近任意连续函数，是解决非线性问题的常用方法之一。在使用径向基神经网络建模的过程中发现：采用 350 nm 发射波长处的自相关强度比采用自相关光谱建立的定量模型对溶液中 L-色氨酸浓度的预测结果更为准确。径向基神经网络输入层和输出层神经元的个数均为 1；评价函数采用均方根误差，目标值 goal 为 10^{-4}，采用神经元个数递增的方式，确定隐含

层神经元的数目；经多次反复尝试后，扩散系数 spread 为 5 时，模型准确度较高，预测结果如表 3 所示。两种外扰方式下，径向基神经网络定量模型对溶液中 L-色氨酸浓度预测值和实际值的线性拟合分别如图 4(a)和(b)所示。

表 3 RBFNN 模型对样本中 L-色氨酸浓度的预测结果

Table 3 Prediction results of RBFNN model for the concentration of L-Tryptophan

预测集	实际值 /(μg·mL^{-1})	激发波长		浓度	
		预测值 /(μg·mL^{-1})	相对误差 /%	预测值 /(μg·mL^{-1})	相对误差 /%
7	0.5	0.480	−3.952	0.471	−5.742
8	1.0	1.005	0.537	0.999	−0.139
9	1.5	1.627	8.497	1.563	4.176
10	2.0	2.021	1.035	1.992	−0.403
11	2.25	2.203	−2.076	2.208	−1.855
12	2.75	2.797	1.692	2.741	−0.333
r	—	—	99.742	—	99.906
RMSEP	—	—	0.060	—	0.033

由表 3 和图 4 我们可以看出：两种外扰方式下径向基神经网络模型的预测浓度和实际浓度都具有良好的线性关系，对未知样品的预测相关系数均在 99%以上，预测均方根误差较偏最小二乘回归预测模型而言，减小了一个数量级。这是因为偏最小二乘回归算法多用于线性区间，对非线性明显的区间适用性不强；而人工神经网络可以模拟人脑的思维方式，具有很强的学习和记忆能力，因此对于线性区间和非线性区间并存的情形同样适用。对比两种外扰下径向基神经网络模型的预测结果，能够发现：采用浓度为外扰方式构造的自相关光谱，能够有效地抑制噪声，预测结果更为准确，这与我们从图 2 中获取的信息相一致。采用激发波长作为外扰变量时，动态光谱由三条不同的发射谱组成；而在采用浓度作为外扰变量时，动态光谱由背景光谱以及样本光谱两条平均谱组成。动态变化是建立在背景光谱的基础上，可以有效消除背景的干扰；经过均值计算，可以起到降低系统噪声，抑制噪声干扰的作用。

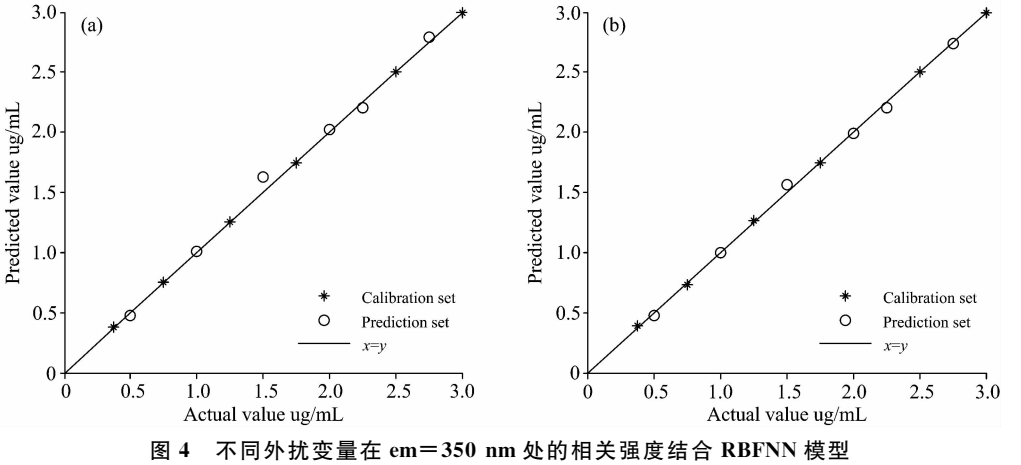

图 4 不同外扰变量在 em=350 nm 处的相关强度结合 RBFNN 模型

(a)：激发波长；(b)：浓度

Fig. 4 Auto correlation intensity of em=350nm with the different disturbance variable combined with RBFNN

(a)：by excitation wavelength；(b)：by concentration

3 结 论

以预测 L-色氨酸溶液的浓度为例，研究了荧光相关光谱在采用激发波长和浓度两种外扰方式以及偏最小二乘回归和径向基神经网络两种算法下的建模效果，研究结果表明：采用浓度为外扰构造的自相关光谱，能有效地抑制噪声；而在采用同种外扰数据时，径向基神经网络模型比偏最小二乘回归模型对 L-色氨酸浓度的预测更为准确，可为食品安全的含量检测提供借鉴。

References

[1] Yang Renjie, Liu Rong, XU Kexin. Proceedings of SPIE, 2012, 8229: 822918.

[2] Liu Xin, Lü Hua, Xu HuaChen. Chinese Chemical Letters, 2015, 26(2): 205.
[3] BO Gui-jun, YU Jing, DI Hui-hui, et al(卜贵军，于　静，邸慧慧，等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2015, 35(2): 362.
[4] Dong Yongjiang, Lu Hao, Yong Zhengdong, et al. Applied Optics, 2015, 54(23): 7032.
[5] LIU Rong, YANG Ren-jie, XU Ke-xin(刘　蓉，杨仁杰，徐可欣). The Journal of Light Scattering(光散射学报), 2013, 5(1): 92.
[6] SUN Yan-hui, CAI Hua-zhen, JIA Xiao-li, et al(孙艳辉，蔡华珍，贾小丽，等). Chinese Journal of Analytical Chemistry(分析化学), 2013, 41(9): 1373.
[7] Feng Juan, Han Ting, Zhang Miqing, et al. Chinese Chemical Letters, 2015, 26(2): 210.
[8] HE Jing-wei, ZHAO Xiu-ying, ZHENG Wei, et al(贺经纬，赵秀英，郑　玮，等). Acta Polymerica Sinica(高分子学报), 2016, 2(2): 234.
[9] LI Dan, YANG Zhong-qiang, XI Long, et al(李　丹，杨中强，奚　龙，等). Acta Polymerica Sinica(高分子学报), 2016, 2(2): 164.
[10] NI Li-jun, ZHANG Li-guo(倪力军，张立国). Basic Chemometrics and Its Application(基础化学计量学及其应用). Shanghai: East China University of Science and Technology Press(上海：华东理工大学出版社), 2011. 56.
[11] WANG Xu-dong, SHAO Hui-he(王旭东，邵惠鹤). Control Theory and Applications(控制理论与应用), 1997, 14(1): 59.

Detection the Concentration of L-Tryptophan by Fluorescence Correlation Spectroscopy

GU Song[1,2], ZHU Zhuo-wei[1,2], MA Chao-qun[1,2], CHEN Guo-qing[1,2*]
1. College of Science, Jiangnan University, Wuxi　214000, China
2. Jiangsu Provincial Research Center of Light Industrial Optoelectronic Engineering and Technology, Wuxi　214000, China

Abstract　The three-dimensional fluorescence spectrum of the L-tryptophan solution was measured in a fluorescence spectrometer. The result showed that the fluorescence peak of L-tryptophan located at 270 nm/350 nm (excitation wavelength/emission wavelength). As the emission wavelength was fixed at 350 nm, an excitation spectrum can be measured. It can be found from the excitation spectrum that the curve has a high slope and good linearity in the range of 250～260 nm. Thus, the excitation wavelengths of 250, 255 and 260 nm were chosen to excite the sample, and three fluorescence emission spectra were measured. Based on the three spectra, the auto correlation spectra with disturbance variable of wavelength were constructed. In addition, the emission spectra of the ultrapure water under different excitation wavelength were measured and averaged to be reference spectrum. The auto-correlation spectra with disturbance variable of concentration were constructed by correlation calculation between the reference spectrum and the averaged spectrum of the samples. Combined the correlation spectral data with partial least squares regression (PLSR) and radial basis function neural network (RBFNN), the prediction models of the concentration of L-tryptophan were measured, respectively. The prediction results showed that the correlation spectrum constructed with disturbance variable of concentration had a better signal-to-noise ratio and better prediction performance. Furthermore, with the same disturbance variable, the model based on RBFNN was more precise than that based on PLSR. Among all the models, the model based on RBFNN with disturbance variable of concentration had the best prediction performance with correlation coefficient of 99.91% and root-mean-square error of 0.033 $\mu g \cdot mL^{-1}$. This method can provide helps in the food safety supervision for the accurate determination of substances.

Keywords　L-tryptophan; Fluorescence correlation spectrum; Partial least squares regression; Radial basis function neural network

(Received Dec. 20, 2016; accepted Jun. 22, 2017)

* Corresponding author