

## 青饲大豆主要品质参数的近红外光谱快速分析

姜妍<sup>1,2</sup>, 蒙何<sup>1</sup>, 赵熠荣<sup>1</sup>, 王先虚<sup>1</sup>, 王遂<sup>1</sup>, 薛恩玉<sup>3</sup>, 王绍东<sup>1\*</sup>

1. 东北农业大学/国家大豆工程技术研究中心, 黑龙江 哈尔滨 150030
2. 黑龙江省绿色食品科学研究院, 黑龙江 哈尔滨 150028
3. 黑龙江省乡村产业发展中心, 黑龙江 哈尔滨 150036

**摘要** 饲料是畜牧生产的物质基础, 饲草原料和饲料产品营养价值的检测与评估是饲料生产中的重要环节, 面对饲草资源中粗蛋白含量低和大量依靠进口饲料的局面, 大豆作为优质的高蛋白豆科饲草是畜牧业生产利用的重要资源。不同青饲大豆及其不同刈割期的饲用品质参数可以评价青饲大豆的饲用性能, 但目前主要以化学方法进行检测, 过程繁琐, 试验周期长、易造成人为操作误差, 且青饲大豆主要饲用品质指标的光谱快速检测尚属空白, 亟待开发和利用。鉴于近红外光谱快速分析技术在检测领域及饲料分析中的广泛应用, 利用近红外光谱分析技术在 950~1 650 nm 范围内收集不同大豆品种不同刈割时期的全株样品光谱, 对样品的主要饲用品质参数粗蛋白(CP)含量、中性洗涤纤维(NDF)含量和酸性洗涤纤维(ADF)含量三个指标按国家标准或行业标准的化学方法进行检测, 得到的 150 份样品的品质数据按 3:2 分为校准集和验证集。通过一阶求导(NW<sup>1st</sup>)、二阶求导(NW<sup>2nd</sup>)、标准正态变量变换方法(SNV)、去趋势算法(DE-trending) 4 种不同光谱预处理方法中的一种或多种处理的组合, 结合偏最小二乘(PLS)回归算法建立了青饲大豆三个主要品质参数 CP, NDF 和 ADF 含量的预测模型。通过比较回归模型中的校准集和验证集中决定系数( $R^2$ )和均方根误差(RMSE)得出, NW<sup>1st</sup> + DE-trending + SNV + PLS 处理后所建立的模型效果最好, 青饲大豆 CP 含量模型中的校准集的  $R_c^2$  和验证集的  $R_v^2$  分别为 0.96 和 0.95, NDF 含量模型中的  $R_c^2$  和  $R_v^2$  分别为 0.90 和 0.89, ADF 含量模型中的  $R_c^2$  和  $R_v^2$  分别为 0.94 和 0.93。通过验证集对预测模型的检验分析进一步证实了该模型的准确性和稳定性, 形成了一种便于青饲大豆品质检测的近红外光谱(NIRS)快速分析法。随着青饲大豆品质参数数据量的增加, 将不断完善青饲大豆品质检测模型, 该方法不但扩大了近红外光谱仪对饲草资源品质的检测类别与范围, 而且准确高效, 有利于高蛋白优质饲草资源的开发和有效利用。

**关键词** 近红外光谱; 青饲大豆; 粗蛋白; 偏最小二乘法

**中图分类号:** S126 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)03-0919-05

### 引言

饲料是畜牧生产的物质基础, 随着畜牧业的快速发展, 优良饲草成为畜牧业生产的重要资源, 对饲草原料和饲料产品营养价值的检测与评估是饲料生产中必不可少的重要环节, 饲料生产中的质量控制及其营养成分的快速分析越发重要。饲料品质的管理与科研的主要手段仍沿用传统经典的化学分析方法分析饲料及饲草的常规组分及营养成分, 这些传统的分析方法存在分析周期长、方法复杂、试剂消耗, 污染

环境和有损健康等缺点。因此, 建立快速准确且简便绿色的分析方法十分必要。

近几年, 近红外光谱(NIRS)快速分析技术在分析化学和检测领域得到广泛应用。NIRS 技术适合于多组分的定性和定量分析, 具有制样简单、无污染、操作简单、检测速度快、稳定性好、在线分析等优点<sup>[1-2]</sup>。基于待检样品在近红外光谱区域的光学吸收特性, 通过国标化学方法对样品的化学成分含量与其近红外反射光谱进行回归分析, 建立模型, 从而对同类及其相似类型的未知样品进行品质指标的检测与分析。NIRS 分析过程一般包括扫描样本, 获得近红外光谱数

收稿日期: 2021-02-21, 修订日期: 2021-05-13

基金项目: 科技部“十三五”七大作物育种专项(2017YFD0101306-05), 黑龙江省“百千万”工程科技重大专项(SC2019ZX16B0039), 东北农业大学“学术骨干”项目(18XG05)资助

作者简介: 姜妍, 女, 1979年生, 东北农业大学国家大豆工程技术研究中心副研究员 e-mail: yanjiang416@163.com

\* 通讯作者 e-mail: wsdhj@neau.edu.cn

据;采用国标分析测试法测定样本的组分或含量数据;建立并优化两者的函数关系,预测模型。NIRS 建模的方式包括主成分分析、多元线性回归、偏最小二乘(partial least squares, PLS)、人工神经网络(artificial neural networks, ANN)等。在众多光谱数据分析中,PLS 是利用频次高、反演精确的一种建模方法,是在多元线性回归的基础上发展起来的一种回归方法<sup>[3-4]</sup>。PLS 利用全部光谱信息进行矩阵分解和回归交互分析。基于 PLS 算法,衍生出多种模型,如多模型 PLS、移动窗口 PLS 建模,及 PLS 与 ANN 联合的 PLS-BP 预测模型,用来检测饲料样品水分、蛋白质、灰分、磷含量。

饲料检测是 NIRS 技术应用最早的领域之一,且发展迅速,应用广泛,马蹄草营养成分的 NIRS 检测<sup>[5]</sup>,便携式光谱仪实时测定饲料中蛋白质、粗纤维和淀粉等组分的研究<sup>[6-7]</sup>,以及建立动物粪便中性洗涤纤维(NDF)的近红外光谱校准方程,NIRS 技术被认为是准确测定牧草化学成分和消化率参数的重要工具<sup>[8-9]</sup>。目前,关于饲料中粗蛋白(CP)含量、粗纤维、粗脂肪等 NIRS 快速测定已有国家标准(GB/T 18868—2002),解决了饲料企业的品质检测和产品开发的需求。现有的 NIRS 技术主要用于玉米、豆粕、鱼粉等饲料原料及浓缩饲料、配方饲料等成品饲料的常规营养组分的性质与含量分析,以及饲料营养价值的快速评定。然而,作为饲草资源的青饲大豆全株品质检测的 NIRS 技术亟待开发和利用。目前,我国饲草存在产量低、品质差,粗蛋白质含量 <15%,优质饲草资源稀缺等问题,大豆作为短日照作物,其光周期反应较为敏感,可以利用大豆这一特性进行南种北移,作为一种新型的青饲豆科资源加以利用,有利于解决饲草蛋白含量低的问题,进一步推进二元种植结构到粮食、经济、饲料作物的三元结构的粮改饲政策。本研究对青饲用不同大豆品种不同时期的整个植株进行粗蛋白(CP)含量、中性洗涤纤维(NDF)含量和酸性洗涤纤维(ADF)含量进行检测,采用不同的光谱预处理方法和变量选择方法来优化模型,利用近红外光谱建立含量分析回归模型,形成一种便于青饲大豆全株品质检测的 NIRS 快速分析法。

## 1 实验部分

### 1.1 样品制备

利用大豆光周期的植株生长特性,选用于来自于东北春夏大豆、黄淮海夏大豆、南方春大豆、南方夏大豆几个生态区域的 50 个大豆品种样本作为青饲大豆全株品质参数检测的品种试验,取样时期为盛花期、鼓粒初期、鼓粒初期 5 d 后去荚处理三个主要青饲大豆全株的利用方式,三次重复,三种方式的地上部分植株经 65 °C 烘干至恒重后粉碎待测。每个样品的 CP, NDF 和 ADF 含量分别按照国家标准(GB/T 6432—2018, GB/T 20806—2006)和行业标准(NY/T 1459—2007)的方法检测,每个品质参数共得到 150 个数据。

### 1.2 近红外光谱收集

利用近红外光谱成像仪(DA7250, Perten, Sweden),光谱范围为 900~1 700 nm,实际工作的光谱波长范围为 950~1 650 nm,波长精度 <0.3 nm,光谱分辨率为 7 nm,二极管

间距(像素间距)3.1 nm·像素<sup>-1</sup>,检测器为 InGaAs,电温控制冷处理,256 像素,内置线性参考光源。可驱动载有样本的旋转杯,用来控制系统运行的计算机以及近红外成像系统采集软件。为保证获取光谱的一致性,检测温度为 15~25 °C,样品厚度一致为 5 mm,湿度范围为 30%~70%,每个样品至少扫描三次以获得光谱数据,所有光谱曲线的波形呈现相同的趋势(图 1),在全波段范围内进行青饲大豆全株品质指标的模型建立。

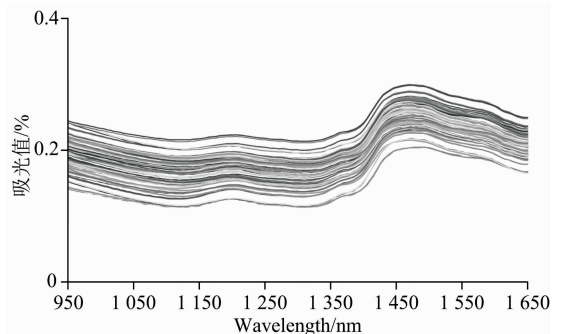


图 1 原始光谱曲线

Fig. 1 Original spectral curve

### 1.3 光谱数据处理

建立 CP, NDF 和 ADF 含量的定量识别模型时,将 150 个数据集按 3:2 比例划分为校准集和验证集。为提高光谱分辨率,去除样品表面颗粒不均匀带来的杂散光等,光谱预处理主要采用一阶导数直接差分法求导法(NW<sup>1st</sup>)、二阶导数直接差分法求导法(NW<sup>2nd</sup>)、标准正态变量变换法(standard normal variate transformation, SNV)和去趋势算法(De-trending)的一种或几种方法组合进行比较分析。NW<sup>1st</sup>和 NW<sup>2nd</sup>对光谱进行基线校正和光谱分辨预处理;De-trending 算法是为了消除漫反射光谱的基线漂移,将原始光谱的吸光度和波长拟合成一条趋势线,再从原光谱中减掉趋势线;SNV 主要是用来消除颗粒大小不均匀和表面非特异性散射。光谱定量建模分析为多元校正,在物质浓度与光谱之间建立一种定量关联关系。PLS 集合主成分分析、典型相关分析和多元线性回归分析 3 种分析方法的优点,利用响应矩阵提取出反映数据变异的最大信息,具有预测功能,建模性能更优越<sup>[10]</sup>。鉴于 PLS 的建模优势和所用光谱仪参数的匹配度,利用 PLS 法,首先对光谱数据进行运算,求取主因子或隐含变量,再根据主因子或隐含变量的累积贡献率得分值,以及化学指标数据变量,建立模型。通过不同预处理后的全光谱结合 PLS 建立的模型效果图,利用校准集和验证集的决定系数(R<sup>2</sup>)和均方根误差(RMSE)比较分析出适合的光谱预处理建模方法,用于青饲大豆全株主要品质指标的快速检测。

## 2 结果与讨论

### 2.1 样品化学值分析

从表 1 中各指标的最大值、最小值和标准差值可以看出,作为青贮饲草资源的不同大豆品种在不同取样时期的

CP, NDF 和 ADF 含量的变化幅度比较大, CP 含量为 11.21%~22.79%, NDF 含量为 41.02%~63.48%, ADF 含量为 25.00%~42.51%, 数据覆盖面较大, 样品化学值具有一定的代表性。

表 1 样品中主要饲用品质指标的含量统计 (%)

Table 1 Content statistics of main forage quality indexes in samples (%)

指标	最大值	最小值	平均值	标准差
CP	22.79	11.21	16.21	2.04
NDF	63.48	41.02	51.25	5.02
ADF	42.51	25.00	33.88	4.51

表 2 不同预处理方法的青饲大豆全株主要品质指标的模型结果

Table 2 Model results of main quality indexes of whole plants of forage soybean with different pretreatments

预处理	CP				NDF				RMSEC			
	RMSEC	$R_c^2$	RMSEP	$R_p^2$	RMSEC	$R_c^2$	RMSEP	$R_p^2$	RMSEC	$R_c^2$	RMSEP	$R_p^2$
NW <sup>1st</sup>	0.67	0.93	0.73	0.92	1.97	0.87	2.13	0.85	1.23	0.93	1.34	0.92
NW <sup>1st</sup> +DE	0.54	0.96	0.61	0.95	1.72	0.90	1.89	0.87	1.20	0.93	1.30	0.92
NW <sup>1st</sup> +DE+SNV	0.49	0.96	0.55	0.95	1.74	0.90	1.87	0.89	1.10	0.94	1.20	0.93
NW <sup>2nd</sup> +DE+SNV	0.49	0.96	0.56	0.96	1.71	0.91	1.94	0.88	1.23	0.93	1.32	0.92

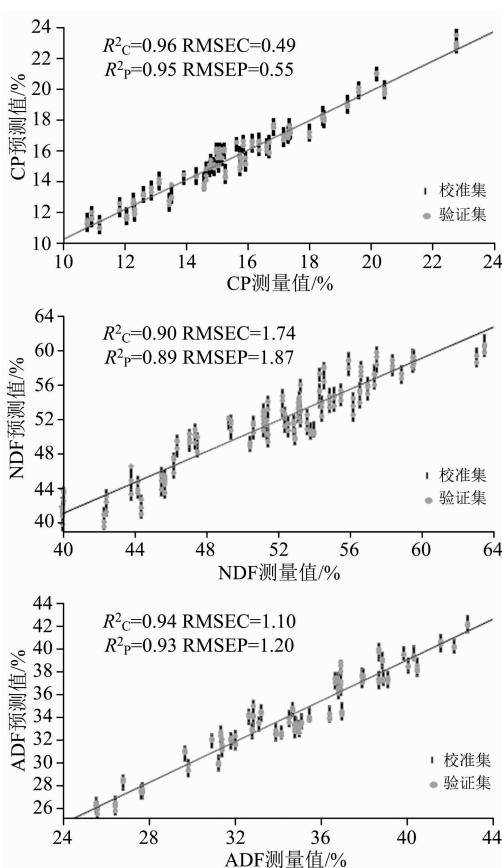


图 2 NW<sup>1st</sup>+DE-trending+SNV 处理的青饲大豆品质 PLS 模型

Fig. 2 PLS model of forage soybean quality by NW<sup>1st</sup>+DE-trending+SNV treatments

## 2.2 全波段模型效果评价

为了提高预测模型精确度, 以往研究多关注不同光谱处理条件的组合、不同定量建模分析以及数学方法的应用上, 但实际经验证实, 光谱仪的检测精度、光谱收集过程的条件控制, 以及化学值测量的准确度才是建立校准模型的基础和关键。为准确定量分析青饲大豆植株的饲用品质, 将数据分为校准集( $R_c^2$ , RMSEC)和验证集( $R_p^2$ , RMSEP)分别建立了 NW<sup>1st</sup>, NW<sup>1st</sup>+DE-trending, NW<sup>1st</sup>+DE-trending+SNV, NW<sup>2nd</sup>+DE-trending+SNV 四种光谱预处理的 PLS 模型。通过比较校准集和验证集的  $R^2$  和 RMSE,  $R^2$  越接近 1, 说明回归效果显著, RMSE 越接近 0, 说明模型具有较好的稳定性和预测能力。

从不同预处理方法对青饲大豆全株主要品质指标的模型分析结果可以得出(表 2), NW<sup>1st</sup>+DE-trending+SNV 处理的模型中 CP 含量的  $R_c^2$  和  $R_p^2$  分别为 0.96 和 0.95, 且 RMSEC 和 RMSEP 值较其他三个处理低, 分别为 0.49 和 0.55; NW<sup>2nd</sup>+DE-trending+SNV 处理中 CP 含量的  $R_c^2$  和  $R_p^2$  均为 0.96, 但 RMSEP 值为 0.56, 高出 0.01, 说明 NW<sup>1st</sup>+DE-trending+SNV 的 PLS 模型用于检测 CP 含量的效果好于他处理, 从 NW<sup>1st</sup>+DE-trending+SNV 处理的青饲大豆全株 CP 含量 PLS 模型中可以看出(图 2), 校正集和验证集的回归直线基本重合。NDF 含量模型中(表 2), NW<sup>1st</sup>+DE-trending+SNV 的  $R_c^2$  和  $R_p^2$  分别为 0.90 和 0.89, RMSEC 和 RMSEP 分别为 1.74 和 1.87, 其 RMSEP 值相比其他处理最低, 且校正集和验证集的回归直线基本重合; ADF 含量模型中, NW<sup>1st</sup>+DE-trending+SNV 处理的  $R_c^2$  和  $R_p^2$  最高, 分别为 0.94 和 0.93, 其 RMSEC 和 RMSEP 最低, 分别为 1.10 和 1.20, 相比于 NW<sup>2nd</sup>+DE-trending+SNV 处理低 0.12~

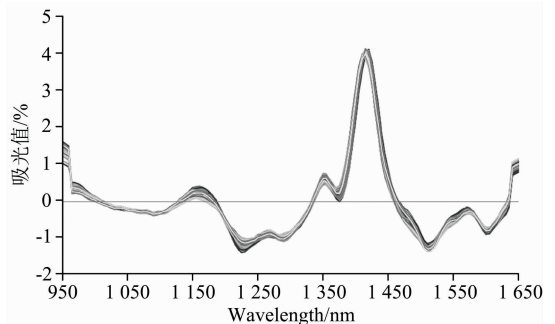


图 3 NW<sup>1st</sup>+DE-trending+SNV 处理的全光谱曲线图  
Fig. 3 The full spectrum curve after NW<sup>1st</sup>+DE-trending+SNV treatments

0.13。综合比较 CP, NDF 和 ADF 三个含量指标,  $NW^{1st} + DE-trending + SNV$  光谱预处理的 PLS 模型效果较好, 更稳定更准确(图 2)。经过  $NW^{1st} + DE-trending + SNV$  这三种光谱预处理组合的方法, 原始光谱曲线变得平滑, 曲线的噪声、基线漂移、共线性现象消除效果较好(图 3)。采用  $NW^{1st} + DE-trending + SNV$  的 PLS 模型检测青饲大豆全株的主要饲用品质指标。

### 3 结 论

利用近红外光谱分析仪获得稳定有效光谱, 通过国标法

或行标法对不同品种不同处理的青饲大豆植株中三个主要饲用品质参数 CP, NDF 和 ADF 含量进行定量检测, 采用  $NW^{1st}$ ,  $NW^{2nd}$ , DE-trending 和 SNV 四种光谱预处理方法的不同组合建立品质预测模型, 经比较分析校准集和验证集的  $R^2$  和 RMSE, 得出  $NW^{1st} + DE-trending + SNV + PLS$  方法建立的模型优化效果好, 并验证了 PLS 模型的准确性及稳定性, 随着青饲大豆资源的利用和推广, 其品质检测数据量会随之增加, 将进一步优化模型和补充参数, 使青饲大豆品质的近红外检测方法更加完善。

致谢: 非常感谢波通公司在模型解析中给予的指导。

### References

- [1] Ferreira D S, Pallone J A L, Poppi R J. Food Research International, 2013, 51: 53.
- [2] Leite D C, Corrêa A A P, Júnior L C C, et al. Journal of Food Composition and Analysis, 2020, 91: 103536.
- [3] Guo W L, Du Y P, Zhou Y C, et al. World Journal of Microbiology and Biotechnology, 2012, 28(3): 993.
- [4] LIU Yan-de, XU Hai, SUN Xu-dong, et al(刘燕德, 徐海, 孙旭东, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2020, 40(3): 922.
- [5] Harris P A, Nelson S, Carslake H B, et al. Journal of Equine Veterinary Science, 2018, 71: 13.
- [6] Modroño S, Soldado A, Martínez-Fernández A, et al. Talanta, 2017, 162: 597.
- [7] Xu R, Hu W, Zhou Y, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2020, 224: 117400.
- [8] Brogna N, Palmonari A, Canestrari G, et al. Dairy Science, 2017, 101: 1234.
- [9] Righi F, Simoni M, Visentin G, et al. Livestock Science, 2017, 206: 105.
- [10] REN Zhi-shang, PENG Hui-hui, HE Zhuang-zhuang, et al(任志尚, 彭慧慧, 贺壮壮, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2020, 51(S2): 466.

## Rapid Analysis of Main Quality Parameters in Forage Soybean by Near-Infrared Spectroscopy

JIANG Yan<sup>1,2</sup>, MENG He<sup>1</sup>, ZHAO Yi-rong<sup>1</sup>, WANG Xian-xu<sup>1</sup>, WANG Sui<sup>1</sup>, XUE En-yu<sup>3</sup>, WANG Shao-dong<sup>1\*</sup>

1. Northeast Agricultural University/National Soybean Engineering Technology Research Center, Harbin 150030, China

2. Heilongjiang Academy of Green Food Science, Harbin 150028, China

3. Heilongjiang Rural Industry Development Center, Harbin 150036, China

**Abstract** Forage is the material basis of animal husbandry production. The detection and evaluation of the nutritional value of forage raw materials and feed products are an important link in feed production. Facing the situation of low crude protein content in forage resources and relying on many imported feeds, soybean, as a high-quality, high protein legume forage, is an important resource for animal husbandry production and utilization. The feeding quality parameters of different forage soybean and different cutting periods can evaluate the feeding performance of forage soybean. However, the chemical method is mainly used for detection, which is cumbersome, long test cycle and is easy to cause operation error. Moreover, the rapid detection of main feeding quality indexes of forage soybean is still blank, which needs to be developed and utilized urgently. Because of the wide application of near-infrared spectroscopy in detection and feed analysis, the whole plant samples of different soybean varieties in different cutting periods were collected by near-infrared spectroscopy in the range of 950~1 650 nm. The content of crude protein (CP), neutral detergent fiber (NDF) and acid detergent fiber (ADF) was detected according to the national standard or industry standard chemical method. The 150 samples data were divided into calibration and verification set according to 3 : 2. The prediction models of three main quality parameters CP, NDF and ADF content of forage soybean, were established by combining one or more of four different spectral pretreatment methods, including first-order derivative ( $NW^{1st}$ ), second-order derivative ( $NW^{2nd}$ ), standard normal variable transformation (SNV) and detrending (DE-trending), and partial least squares

(PLS) regression algorithm. By comparing the coefficient of determination ( $R^2$ ) and root mean square error (RMSE) of calibration set and validation set in regression models, the results showed that the model established by  $NW^{1st} + DE\text{-trending} + SNV + PLS$  had the best effect. The  $R_C^2$  and  $R_P^2$  of the calibration set and validation set in forage soybean CP content model were 0.96 and 0.95 respectively, the  $R_C^2$  and  $R_P^2$  of NDF content model were 0.90 and 0.89 respectively, and the  $R_C^2$  and  $R_P^2$  of ADF content model were 0.94 and 0.93 respectively. The accuracy and stability of the model were further confirmed by the test and analysis of the validation, and rapid analysis of near-infrared spectroscopy (NIRS) method for the qualitative detection of forage soybean was formed. With the increase of forage soybean quality parameter data, the quality detection model of forage soybean will continuously improve. This method expands the detection category and range of forage resource quality by near-infrared spectrometer and is accurate and efficient, which is conducive to the development and effective utilization of high-quality high protein forage resources.

**Keywords** Near-infrared spectroscopy; Forage soybean; Crude protein; Partial least squares method

(Received Feb. 21, 2021; accepted May 13, 2021)

\* Corresponding author

---

## 本刊声明

近期以来，一些不法分子假冒《光谱学与光谱分析》期刊社名义，以虚假网站等形式欺骗广大作者、读者。这些虚假网站公然假冒《光谱学与光谱分析》期刊名义进行大肆的征稿并骗取作者的审稿费和版面费。经部分作者及读者举报，现有关部门已就此介入调查。本刊将通过法律途径向假冒者追究相应的责任，维护本刊权利。

本刊官方网站已正式开通，网址为

<http://www.gpxygpx.com/>

在此郑重声明，本网址为《光谱学与光谱分析》期刊唯一开通运行的官方网站。本刊从未授权任何单位或个人以任何形式（包括网上网下）代理本刊征稿、审稿等项业务。

希望广大读者和作者切实维护好自身的合法权益，防止受骗上当。

《光谱学与光谱分析》期刊社

2019年3月15日