

FT-NIR 光谱法同时测定 烟草根茎叶中的氮、磷、氯和钾

王家俊

(红河卷烟总厂产品部 云南 弥勒 652300)

摘要 应用傅里叶变换近红外 (FT-NIR) 光谱法测定 1200 个具代表性的烤烟各个生长期的根茎叶样品的近红外光谱数据, 采用化学计量学中的多元校正方法——偏最小二乘 (PLS) 算法分别对根茎和烟叶建立了测定氮、磷、氯和钾等主要营养元素含量的校正模型。通过对模型进行数理统计检验, 在显著性水平大于 5% 的条件下, 其测定结果与提供基础数据方法测定的结果对比, 两者不存在显著性差异。该分析方法应用于同时测定烟草根茎叶氮、磷、氯和钾的含量, 结果令人满意。

关键词 傅里叶变换近红外光谱法 偏最小二乘法 烟草 氮 磷 氯 钾

一般认为, 氮、磷、氯和钾等都是烤烟中的重要营养元素, 对烤烟的生长、产量和质量均有很大的影响^[1], 追踪烤烟各个生长时期根茎叶营养元素含量的变化, 对于提高烤烟的生长品质有着十分重要的意义。烟叶中这些营养元素含量的测定, 一般采用原子吸收光谱法 (AAS)、电感耦合等离子发射光谱法 (ICP-AES)^[2,3] 和流动注射分析方法 (FIA)^[4,5]。然而, 这些方法的样品前处理费工耗时, 且需用较多的分析设备, 分析费用高, 难以进行大批量样品的快速分析。傅里叶变换近红外 (FT-NIR) 光谱法具有样品前处理简捷、分析速度快、稳定性好、精度高的特点, 一张近红外光谱中含有丰富的理化信息, 应用建立的近红外数学模型可以同时预测未知样品的多种成分或性质, 如近年来我厂已应用 FT-NIR 光谱法同时测定烟草中总氮、总糖和烟碱的含量^[6], 现场控制烤烟收购的内在质量。本文介绍的是应用 FT-NIR 光谱和偏最小二乘 (PLS) 法^[7] 建立预测烤烟各个生长时期根茎叶氮、磷、氯和钾含量校正模型的方法, 旨在为研究烤烟的生长品质提供快速的分析手段。

1 材料与方方法

1.1 样品的采集与制备

烟叶样品采集于红河州烤烟基地, 分别在团棵期、旺长期和成熟期进行采集, 按烤烟主要品种 (包含 K326、云 85、云 87、红大和 V2) 均衡采样, 如在旺长期, 2003 年 7 月所采集的 90 个烟根样品, 这五个品种在其中各占的数量分别为 18 个, 茎也亦然; 采集的烟叶样品 210 个, 上、中、下部位烟叶样品数量均衡, 各占 70 个, 这五个品种在其中又分别各占 14 个。所采集的样品, 烟根为 320 个, 烟茎为 240 个, 烟叶为 650 个 (表 1)。将采集的根、茎和烟叶样品于 105℃ 杀青 15min, 并在 60~70℃ 下烘至样品含水率在 12% 左右, 粉碎, 过 60 目筛, 样品装入密封袋中, 常温下避光下贮存。

表 1 根茎叶样品统计

采样时期		根	茎	烟叶
团棵期	2003年5月	30	-	50
	2003年6月	50	-	60
旺长期	2003年7月	90	90	210 ^①
	2003年8月	80	80	180 ^①
成熟期	2003年9月	70	70	150 ^①
样品总量		320	240	650

注：①上、中、下部位烟叶均衡。

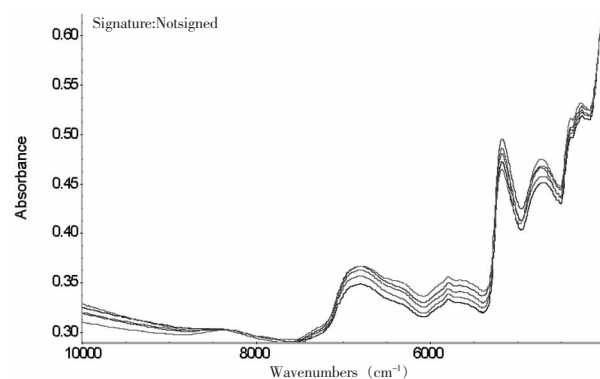


图 1 根、茎的近红外漫反射光谱图

1.2 主要实验设备及应用软件

Nicolet Antaris™ FT-NIR 光谱仪、RESULT™ 集成软件和 TQ Analyst 6.2 光谱数据处理软件（美国 Thermo Nicolet 公司）；Skalar SAN^{PLUS} 流动分析仪（荷兰 Skalar 公司）；AA800 原子吸收光谱仪（美国 PE 公司）；ICP-AES 等离子发射光谱仪（美国 Leeman 公司）。

1.3 采集样品的光谱数据

采集前，先调用 RESULT™ 集成软件编写采集流程，再设置仪器工作参数。本试验设置的仪器工作参数为：光谱扫描范围（Data Range）：10000~4000cm⁻¹；分辨率（Resolution）：8 cm⁻¹；扫描次数（Number of scan）：72 次（所需时间约 35 秒）。而后于恒温（25℃）恒湿（相对湿度 65%）的条件下，开机预热 Antaris™ FT-NIR 光谱仪 2h。将制好的粉末样品置于 5cm 石英杯中，装样量控制在石英杯容量的一半左右，并用压样器轻压至平。样品装载完成后，即可上机采集光谱数据，采集到的根、茎和烟叶的近红外漫反射光谱分别为图 1、图 2。

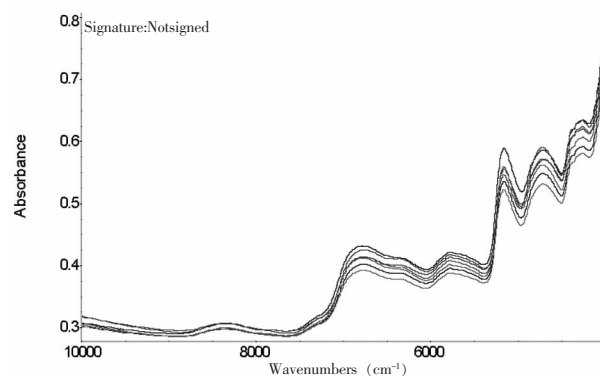


图 2 烟叶的近红外漫反射光谱图

1.4 测定样品的氮、氯、磷和钾

采用 Skalar SAN^{PLUS} 流动分析仪测定样品中氮和氯的含量^[4,5]，采用 AA800 原子吸收光谱仪和 ICP-AES 等离子发射光谱仪分别测定样品中钾和磷的含量^[2,3]。

1.5 建立模型

应用 TQ Analyst 6.2 光谱数据处理软件中的偏最小二乘（PLS）算法，将采集的 560 个根、茎样品（其中根 320 个、茎 240 个）和 650 个烟叶样品的近红外光谱与其相应的氮、氯、磷和钾等的基础测定数据分别进行拟合建模。所获得的校正模型的数理指标见表 2。相关系数（Correlation coefficient）、主成分数（Factor）、估计误差（RMSEC），以及测定值（Actual）与模型自身预测值（Calculated）的散点图和残差分布图分别见图 3~图 9。

2 结果与讨论

2.1 建模训练集样品的选择

表 2 根茎烟叶氮、氯、磷和钾等校正模型的有关数理指标

类别	模型名称	模型中的样本数量	相关系数	最佳主因子	预测误差 (RMSECV)	预测范围 (%)	备注
根茎	氮	514	0.9749	11	0.0860	0.80 ~ 2.30	图 3
	氯	512	0.9960	21	0.0450	0.14 ~ 1.35	图 4
	钾	533	0.9931	21	0.1063	0.50 ~ 2.80	图 5
烟叶	氮	619	0.9960	14	0.1062	0.90 ~ 4.80	图 6
	氯	580	0.9709	18	0.0968	0.70 ~ 1.85	图 7
	磷 ^①	432	0.9230	7	186	1000 ~ 2700	图 8
	钾	621	0.9920	15	0.2162	1.10 ~ 5.80	图 9

备注：① $\times 10^{-6}$ 。

图 3 根茎-氮的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 4 根茎-氯的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 5 根茎-钾的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 6 烟叶-氮的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 7 烟叶-氯的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 8 烟叶-磷的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 9 烟叶-钾的测定值 (Actual) 与模型预测值 (Calculated) 的散点图和残差分布图

图 10 根、茎和烟叶的分类分布图

在选择建模的训练集样品时,应充分考虑样品含量分布、组成和结构的代表性和样品基体的一致性,代表性越广,所建立的模型的适应能力也就越强。但若组成、结构特征和样品基体存在较大的差异,就必须考虑分类建模,以获得较为准确的预测结果^[8]。据此,对各个品种进行均衡采样,其目的就是使各个品种的样品在模型中分布均匀,以保证所建模型在将来预测这五个品种样品的未知化学成分时有良好的适应性。从采集到的根、茎和烟叶的近红外光谱图(分别为图 1、图 2)来看,似乎都很相似,但通过采用化学模式识别——判别分析方法^[9],对根、茎和烟叶的近红外光谱进行分类处理时,结果见图 10,由此发现,根、茎光谱相似,但与烟叶却存在明显的差异。所以,在进行建模时,有必要将根、茎样品合并为一类,作为一个训练集,烟叶样品作为另一个训练集分别进行建模。

2.2 光谱数据的采集

为了获得较高质量的光谱数据,应在严格的实验条件下采集光谱数据,仪器的工作参数设置应保持一致,并定期对仪器进行校验,一般情况下,每次分析时,均可调用自动校验程序让仪器自检。此外,还要做到:①样品杯中样品的密实程度要一致;②从装载样品到上机扫描,整个过程应在最短的时间内完成;③每扫描 2 个样品应扫描一次背景,以获得信噪比高的光谱数据;④在时间允许的情况下,尽量增加扫描次数,以提高光谱的信噪比,这样要比后期对光谱数据进行预处理的效果理想。

2.3 光谱数据的预处理

对光谱数据进行必要的平滑和求导等数学预处理,其目的是过滤噪声、提高信噪比,消除基线飘移的干扰,优化光谱信息^[10]。从本实验的建模效果来看,Norris 导数滤波(Norris derivative filter)优于 Savitzky-Golay 平滑滤波;二阶求导优于一阶求导;二阶导数光谱与 Norris 导数滤波法相结合,可获得理想的建模效果。值得注意的是,采用 Norris 导数滤波时,应根据建模对象的差异慎重选择段长(Segment length)和段间

表 3 配对样本的统计分析

类别	成分	分析方法	平均值	样本容量	标准差
根茎	氮	YC/T 161 - 2002	1.4891	33	0.41871
		FT - NIR	1.4964	33	0.41122
	氯	YC/T 162 - 2002	0.7009	33	0.29017
		FT - NIR	0.7061	33	0.29153
	钾	AAS	1.6473	33	0.72160
		FT - NIR	1.6455	33	0.70280
烟叶	氮	YC/T 161 - 2002	2.3803	32	0.88650
		FT - NIR	2.3637	32	0.91920
	氯	YC/T 162 - 2002	1.2603	32	0.27540
		FT - NIR	1.2591	32	0.26090
	磷	ICP - AES	1775.2	30	518.76
		FT - NIR	1764.0	30	483.98
	钾	AAS	3.0327	33	1.4634
		FT - NIR	3.0294	33	1.4476

表 4 配对样本的 t-检验

类别	配对名称	配对差的平均值	配对差的标准差	t 值	显著性水平
根茎	氮	-0.0073	0.0630	-0.6636	0.5117
	氯	-0.0052	0.0271	-1.0938	0.2822
	钾	0.0018	0.0632	0.1653	0.8697
烟叶	氮	0.0166	0.0726	1.2908	0.2063
	氯	0.0012	0.0659	0.1055	0.9166
	磷	11.21	169.87	0.3615	0.7204
	钾	0.0033	0.1511	0.1267	0.9000

距 (Gap Between Segments)。另外, 虽然光谱数据处理软件会建议选择最佳数据波段, 但对不具有近红外活性的成分如钾元素来说, 只有选择相对较宽的波段才能得到理想的 (通过间接的相关关系) 建模效果。

2.4 模型的优化与检验

当校正模型的雏形建立后, 必须进一步优化模型, 优化模型是一个重复运算模型的过程, 其目的是剔除异常值^[11](Outlier), 直到获得最佳数理指标的校正模型。异常值是指远离模型整体的量测数据, 对模型的回归分析影响较大, 这类数据一般具有较大的杠杆值 (Leverage) 及残差值, 较为方便的是, 通过 TQ Analyst 软件提供的诊断功能, 采用交互效验 (Cross-validation) 的建模方法, 便可剔除这类异常数据。

通过交互效验建立的校正模型, 从模型本身的数理指标 (表 2 和图 3~图 9), 如残差分布、相关系数、估计误差和预测误差等便可衡量模型预测能力的优劣^[12], 此外, 可采用独立校验集验证所建模型的预测能力, 即另外分别随机抽取根、茎和烟叶样品 (样本容量不低于 30 个) 作为独立校验集, 应用建好的模型对其氮、氯、磷和钾含量进行预测, 其预测值与测定结果进行配对 t-检验, 统计分析结果 (表 3 和表 4) 表明, 在 5% 显著性水平下, 不存在显著性差异。表 5 为采用 FT-NIR 光谱法和上述仪器分析方法对另外抽样的根、茎和烟叶样品的氮、氯、磷和钾含量的预测值和测定结果对比。由此可知, 以上所建立的模型具有较好的预测效果。

采用 Antaris™ FT-NIR 光谱仪和校正模型对同一根、茎和烟叶样品分别进行 8 次测定, 结果 (表 6) 显

表 5 烟叶和根茎样品 FT-NIR 光谱法与 FIA、AAS 和 ICP-AES 方法测定结果的比较 (n=9)

类别	样品 编号	氮(%)		氯(%)		钾(%)		磷($\times 10^{-6}$)	
		YC/T 161-2002	FT-NIR	YC/T 162-2002	FT-NIR	AAS	FT-NIR	ICP-AES	FT-NIR
烟叶	1#	3.60	3.50	1.79	1.86	4.65	4.86	1230.0	1100.2
	2#	1.79	1.72	1.13	1.06	1.32	1.23	951.0	1124.6
	3#	2.34	2.27	1.49	1.40	2.19	2.12	1466.0	1367.9
	4#	1.93	1.90	1.24	1.19	1.98	1.94	1620.0	1470.7
	5#	3.31	3.29	1.36	1.46	3.73	3.75	1378.0	1498.0
	6#	1.14	1.14	0.84	0.89	1.47	1.37	1541.0	1579.5
	7#	1.81	1.84	1.23	1.27	2.80	2.66	1472.0	1643.8
	8#	2.59	2.65	1.36	1.36	3.06	3.06	2742.0	2603.2
	9#	2.84	2.92	1.53	1.49	3.30	3.43	2548.0	2679.9
根茎	1#	0.91	0.94	0.34	0.30	0.81	0.85	-	-
	2#	1.25	1.26	0.95	0.93	1.78	1.82	-	-
	3#	1.34	1.40	0.43	0.43	0.62	0.61	-	-
	4#	1.68	1.68	0.97	0.96	1.33	1.44	-	-
	5#	1.80	1.73	0.66	0.65	1.99	2.07	-	-
	6#	1.89	1.88	1.19	1.17	2.20	2.10	-	-
	7#	1.93	1.95	0.23	0.22	2.33	2.30	-	-
	8#	1.56	1.54	0.51	0.52	0.99	1.03	-	-
	9#	2.16	2.09	0.90	0.92	2.74	2.77	-	-

表 6 根茎叶校正模型测量精度的实验结果

测定 次数	根茎			烟叶			
	氮(%)	氯(%)	钾(%)	氮(%)	氯(%)	磷 ppm	钾(%)
1	1.89	1.01	2.01	2.72	1.52	1516	3.21
2	1.87	0.98	2.00	2.75	1.5	1523	3.23
3	1.88	1.03	1.98	2.69	1.49	1496	3.19
4	1.91	1.02	1.99	2.74	1.50	1517	3.22
5	1.91	1.00	2.05	2.73	1.48	1520	3.19
6	1.88	1.02	2.03	2.72	1.47	1529	3.18
7	1.92	0.99	2.01	2.67	1.51	1495	3.24
8	1.93	0.97	1.99	2.68	1.48	1522	3.17
平均值	1.90	1.00	2.01	2.71	1.49	1514.75	3.20
标准差	0.0217	0.0212	0.0231	0.0292	0.0169	11.7234	0.0250
RSD%	1.14	2.12	1.15	1.07	1.13	0.77	0.78

示, Antaris FT-NIR 光谱仪和校正模型具有良好的测量精度。

2.5 模型的应用与维护

应用 Antaris™ FT-NIR 分析系统的集成软件 (Result-Integration) 将建好的校正模型编入分析流程, 然后通过分析系统的操作软件 (Result-Operation) 调出相应的分析流程即可进行未知样品的氮、氯、磷和钾的同

时测定。通常情况下，1h可分析60个样品。

模型投入应用后，必须对模型的适应性适时进行校验，采用独立校验集验证是一种比较可靠、灵活的评价方法，可以随时对模型的适应性进行评价。通常，在模型使用1~2年后，可采集一批具有代表性的样品（至少30个）作为独立校验集，应用FT-NIR光谱法和FIA、AAS、ICP-AES等方法分别测定其相应化学成分的含量，然后进行统计评价，其目的是删除原校正模型中不适应的光谱数据，增加欠缺的光谱数据，优化模型，使模型始终保持良好的适应状态。

3 结论

总之，采用FT-NIR光谱法同时测定烟草根、茎和烟叶中的氮、氯、磷和钾含量，样品前处理简单、无损，分析速度快，精度高，具有广阔的应用前景。

参考文献

- [1] D Layten Davis, Mark T Nielsen 编. 国家烟草专卖局科技教育司, 中国烟草科技信息中心组织翻译. 烟草——生产, 化学和技术 (第1版) [M]. 北京: 化学工业出版社, 2003, 75~82.
- [2] Meyer G A, Keliher P N. An overview of analysis by inductively coupled plasma-atomic emission spectrometry [M]. VCH Publishers Inc. New York, NY. 1992, 473~505.
- [3] 贺与平, 崔娅, 王淑华等. ICP-AES法同时测定烟草中16种元素 [J]. 理化检验 (化学分册), 2001, 37 (11): 510~511, 516.
- [4] YC/T 161-2002 烟草及烟草制品 总氮的测定 连续流动法 [P].
- [5] YC/T 162-2002 烟草及烟草制品 氯的测定 连续流动法 [P].
- [6] 王家俊. FT-NIR光谱分析技术测定烟草中总氮、总糖和烟碱 [J]. 光谱实验室, 2003, 20 (2): 181~185.
- [7] 梁逸曾, 俞汝勤. 化学计量学 [M]. 北京: 高等教育出版社, 2003, 125~127
- [8] 冯新庐, 史永刚. 近红外光谱及其在石油产品分析中的应用 [M]. 北京: 中国石化出版社, 2002, 178~187.
- [9] 梁逸曾, 俞汝勤. 化学计量学 [M]. 北京: 高等教育出版社, 2003, 196~198
- [10] 梁逸曾, 俞汝勤. 化学计量学 [M]. 北京: 高等教育出版社, 2003, 85~96.
- [11] 梁逸曾, 俞汝勤. 分析化学手册 (第十分册, 第2版): 化学计量学 [M]. 北京: 化学工业出版社, 2000, 211~213.
- [12] 陆婉珍, 袁洪福, 徐广通等. 现代近红外光谱分析技术 [M]. 北京: 中国石化出版社, 2000, 170.