

湿度对水泥生料近红外光谱检测的影响及补偿方法

肖航, 杨振发, 张雷*, 张法业, 隋青美, 贾磊, 姜明顺

山东大学控制科学与工程学院, 山东 济南 250061

摘要 近红外光谱检测已被应用于水泥生料成分的快速检测, 但现场环境中的湿度等因素会对光谱产生干扰, 从而降低检测精度。为了提高检测精度, 在实验分析湿度对水泥生料近红外光谱检测影响的基础上研究了补偿方法。在水泥厂选取了24份水泥生料样本, 其中18份作为校正集, 6份作为验证集; 水泥生料中的有效成分为 SiO_2 , Al_2O_3 , Fe_2O_3 和 CaCO_3 , 各成分含量的标准值由X射线荧光光谱分析测出。首先, 将校正集的18份样本每份重复装样测5次光谱, 用得到的90个光谱建立模型I; 再每份样品制作5个湿度梯度样本, 其获得过程为, 先将样本放置在电加热平台上, 用玻璃棒将样本摊平, $180\text{ }^\circ\text{C}$ 下加热30 min, 再将样本放置在散热片上进行降温, 待样品恢复室温后取出进行第一次光谱扫描, 得到1个光谱, 将测量后的样本放入搅拌器, 使用装有去离子水的喷雾器对其喷雾两次, 然后搅拌30 s混合均匀, 测量混合后的样本得到下一个光谱, 重复该过程, 得到具有湿度梯度的5个光谱。所有样本均采用烘干法进行湿度测量, 样本湿度变化区间在 $0.6\%\sim 2\%$ 以内。对每个湿度梯度的样本测量1次, 用得到的这90个光谱建立模型II。然后, 将验证集的6份样本每份制作5个湿度梯度, 获取方式与校正集相同, 对每个湿度梯度的样本测量1次, 得到30个光谱。所有光谱均采用多元散射校正预处理, 拟合波段选择 $4\ 000\sim 5\ 000\ \text{cm}^{-1}$, 建模方法采用偏小二乘法。比较同一份样本的5个湿度梯度, 可以看到在 $5\ 200\ \text{cm}^{-1}$ 处光谱差异最大, 在其他位置也有肉眼可见的明显差异, 因此, 湿度变化对全波段光谱有明显的影 响。最后, 将这30个光谱输入模型I与模型II进行验证, 并对比模型I与模型II的预测均方根误差RMSEP。模型II中 SiO_2 , Al_2O_3 , Fe_2O_3 和 CaCO_3 的预测均方根误差RMSEP比模型I分别减小了25%, 31.3%, 33.3%和25%。实验结果表明, 水泥生料样本湿度对近红外光谱模型的预测结果具有一定的影响, 采用具有湿度梯度的样本进行建模可有效降低湿度对预测结果的影响。

关键词 近红外光谱; 水泥生料; 湿度; 偏小二乘法

中图分类号: O657.33 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)03-0867-06

引 言

水泥是广泛应用于房屋等国家基础建设的材料^[1], 在水泥的工业生产中, 水泥各成分的含量决定了水泥的质量, 而水泥生料作为水泥生产过程的中间产物, 其成分含量与最终产物水泥的成分含量有直接的关系。因此, 快速准确的检测水泥生料中各成分含量具有重要意义^[2]。目前水泥工业现场使用的水泥生料检测方法主要为化学滴定法与X荧光检测法^[3], 化学滴定法准确度极高, 但测量时间长, 过程复杂, 且使用大量试剂, 污染环境。X荧光检测法避免了化学试剂的使用, 但检测前的制样过程比较复杂, 检测总时间在1 h

以上, 而过长的检测时间对水泥的成品率有很大的影响。近红外光谱的波段范围很宽($780\sim 2\ 526\ \text{nm}$), 信息量大, 而且光谱扫描速度快, 样本无需预处理, 可以很快的建立新的模型, 在线检测具有很大的优势, 在工业、农业等领域的在线检测中得到了广泛应用^[4-7]。针对水泥工业领域对在线检测的需求, 国内外学者进行了大量的研究。Jicarla Portela Rebouças^[8]使用近红外发射光谱对水泥成品进行了质量检测, 并将检测时间控制在6 min以内。在检测手段及算法上进行了创新性的研究, 达到了较好的检测效果。但是针对湿度对水泥测量结果影响的研究较少, 而湿度补偿是近红外光谱检测在水泥生料检测领域应用中亟待解决的问题之一。

由于水对光谱有较强的吸收作用, 因此光学检测方法的

收稿日期: 2019-05-05, 修订日期: 2019-09-10

基金项目: 国家自然科学基金项目(11504201)资助

作者简介: 肖航, 1991年生, 山东大学控制科学与工程学院博士研究生 e-mail: nirxiaohang@163.com

* 通讯联系人 e-mail: drleizhang@sdu.edu.cn

测量精度通常会受到样品湿度的影响,如 Pan^[9]在使用近红外光谱检测一种抗高血压药物的有效成分依贝沙坦时发现湿度因素会对预测结果产生一定的影响,Tibor^[10]在使用近红外光谱进行定量分析时发现,低浓度咖啡因的预测准确性受相对湿度的影响,而扑热息痛和布洛芬的预测结果对湿度影响具有很好的鲁棒性。上述研究表明,对于不同物质,湿度的影响效果是不同的。在水泥生料工业现场的在线检测过程中,随季节与天气的变化,空气湿度会影响水泥生料湿度的检测结果。湿度对检测结果会产生多大的影响,可否在模型上进行补偿成为了一个亟待需要的问题。目前常用的温湿度补偿方法分为两种。第一种方法是在建模时加入同一样本不同温湿度的干扰样本,从而提高模型的抗干扰能力。第二种方法是将样本的温度或湿度作为一个新的变量加入自变量或因变量矩阵,通过检测当前的温湿度值对结果进行补偿。以温度为例,有研究表明,相对于将样本的温度作为一个新的变量加入自变量或因变量矩阵这一方法而言,将同一样本在不同温度下测量得到的光谱加入校正集进行建模的效果更好^[11]。因此在实际应用中更多使用第一种补偿方法。

首先对比了同一样本不同湿度下的光谱,阐述了湿度对近红外光谱的直观影响,然后基于近红外光谱检测,采用偏最小二乘拟合,分别使用不带有湿度梯度的样本和带有湿度梯度的样本建立模型 I 与模型 II,对比 2 组模型预测不同湿度下样本的预测误差,讨论并总结了水泥生料样本湿度对近红外光谱模型预测结果的影响,为水泥生料在线检测的湿度补偿方法提供了一定的参考。

1 实验部分

1.1 系统

实验系统由计算机与近红外光谱仪(ABB, MB3600 型)组成,近红外光谱仪包含光源、反射镜片、检测窗口(蓝宝石材质)、检测器、数据处理模块,系统结构原理如图 1 所示。近红外光谱仪中光源发出的近红外光通过镜片反射,到达检测窗口处的蓝宝石窗片,近红外光穿过蓝宝石窗片和蓝宝石材质的样品杯底部被样本吸收并反射,反射回的光再由镜片反射到检测器,检测器将检测到的光谱信号传送至数据处理模块,数据处理模块与计算机通信进行指令和数据的传输。

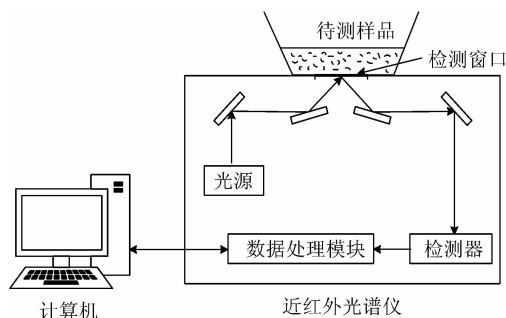


图 1 系统结构原理图

Fig. 1 Diagram of experimental system

1.2 样本制备

实验选用的 24 份水泥生料样本来自曲阜中联水泥厂(中国,山东省),水泥生料中的有效成分为 SiO_2 , Al_2O_3 , Fe_2O_3 和 CaCO_3 ,各成分含量的标准值由 X 荧光光谱分析仪(赛默飞, ARL ADVANT 'X)测出,含量情况如表 1 所示,序号 c1—c18 为校正集, v1—v6 为验证集。

表 1 样本主要成分含量表

Table 1 Main component contents of samples

	$\text{SiO}_2/\%$	$\text{Al}_2\text{O}_3/\%$	$\text{Fe}_2\text{O}_3/\%$	$\text{CaCO}_3/\%$
c1	13.03	3.12	2.05	42.15
c2	13.09	3.15	2.01	42.19
c3	13.15	3.22	2.09	41.87
c4	13.27	3.34	2.07	41.78
c5	13.11	3.24	2.01	42.02
c6	13.32	3.30	2.05	41.76
c7	13.45	3.27	2.06	41.87
c8	13.17	3.23	2.04	41.89
c9	13.79	3.44	2.12	41.44
c10	13.38	3.27	2.03	41.94
c11	13.31	3.16	2.01	41.99
c12	13.24	3.25	2.05	42.01
c13	13.52	3.39	2.08	41.66
c14	13.22	3.19	2.00	42.09
c15	12.54	2.82	1.95	42.54
c16	13.21	3.21	2.00	42.00
c17	13.14	3.14	2.09	42.03
c18	12.75	2.96	2.02	42.57
v1	13.02	3.09	2.04	41.92
v2	13.19	3.22	2.06	42.12
v3	13.12	3.23	2.05	42.02
v4	13.39	3.32	2.01	41.85
v5	13.32	3.13	2.05	41.82
v6	12.92	2.97	1.97	42.32

1.3 方法

使用无湿度梯度的样本建立模型 I,再使用有湿度梯度的样本建立模型 II,观察这两个模型对于带有湿度变化的样本的预测能力,整个实验的流程图如图 2 所示。实施方案为:

(1)将 24 份水泥生料样本分为 18 份的校正集样本与 6 份的验证集样本。校正集中 18 份样本重复装样测量 5 次,得到 90 个光谱,将这 90 个光谱进行建模,得到模型 I。

(2)将 18 份校正集样本每份制作 5 个湿度梯度,每个湿度梯度检测 1 个光谱,得到 90 个光谱,将这 90 个光谱进行建模,得到模型 II。湿度梯度样本的获得过程为,先将样本放置在电加热平台上,用玻璃棒将样本摊平,180 °C 下加热 30 min,再将样本放置在散热片上进行降温,待散热片恢复室温后将样本取出进行第一次光谱扫描,得到 1 个光谱;将测量后的样本放入搅拌器,使用装有去离子水的喷雾器对其喷雾两次,然后搅拌 30 s 混合均匀,测量混合后的样本得到下一个光谱,重复该过程,得到具有湿度梯度的 5 个光谱。

(3)将验证集的 6 份样本也按步骤(2)中方法每份制作 5

个湿度梯度，每个湿度梯度检测 1 个光谱，得到 30 个光谱作为验证集。将验证集分别输入模型 I 与模型 II，对比两个模型得到的预测结果。

实验在实验室环境[温度(24±1)℃，湿度 35%±1%]进行，尽可能保证所有样本的检测条件一致；每次测试将 20 g

样本装入蓝宝石材质的样本杯中，用质量为 200 g 的圆柱形压样器压平后检测。扫描分辨率为 4 cm⁻¹，扫描次数为 32 次，光谱范围 10 000~4 000 cm⁻¹。所有样本在光谱测量后取少量进行烘干法湿度测量，湿度梯度样本的湿度范围在 0.6%~2% 以内，该湿度为水分质量除以样本总质量。

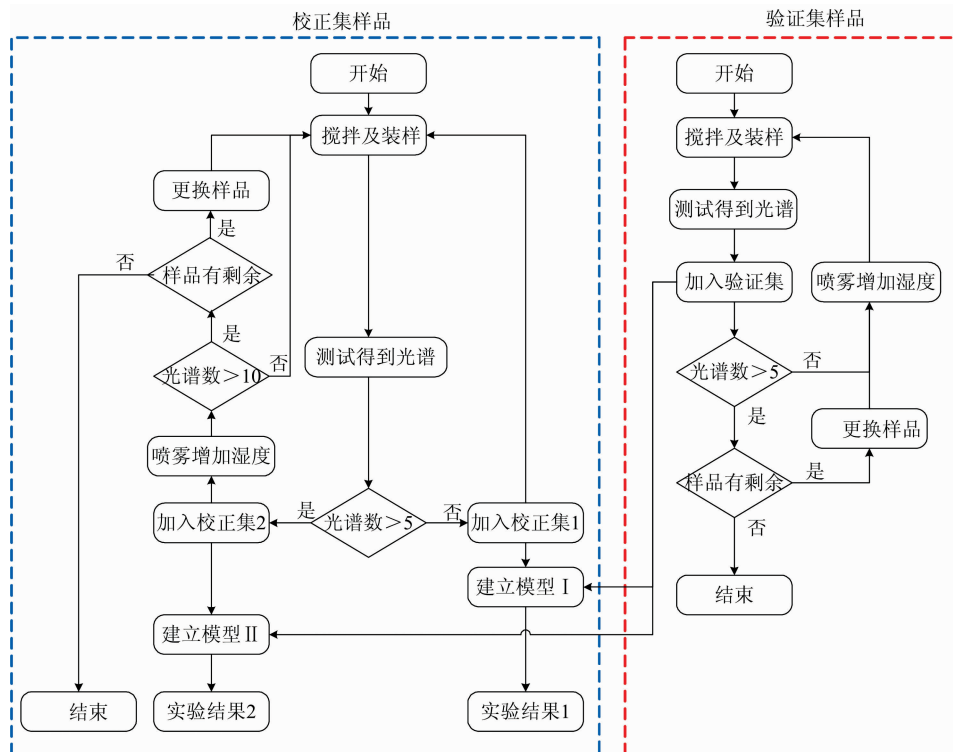


图 2 实验过程流程图

Fig. 2 Flow chart of experimental process

1.4 数据预处理

近红外漫反射方法进行物质成分测量时，由于样本颗粒度大小不均匀或装填密度不一致等引起的光的散射效应常常会导致被测样本光谱有较大的差异。通过多元散射校正(MSC)可以有效降低散射效应对水泥生料样本光谱的影响。

1.5 波段选择

在经过多元散射校正预处理后，将同一份样本的 5 个湿度梯度光谱绘制在同一坐标下，如图 3 所示。图中 a 和 b 两条虚线所标注的吸收峰为水分子中氢氧键的吸收峰，a 和 b 两吸收峰随水分变化峰值变化明显。因此，波段选择应当在包括待测物质的吸收峰的同时尽可能避免将水峰选入拟合波段，由于 SiO₂，Al₂O₃，Fe₂O₃ 和 CaCO₃ 的吸收峰分布在 4 000~5 000 cm⁻¹ [12]，因此选择 4 000~5 000 cm⁻¹ 波段，即图 3 中阴影部分作为拟合波段。

1.6 拟合方法

偏最小二乘法的原理在于建立自变量与因变量的之间的线性回归，模型如式(1)所示

$$Y = XB + E \quad (1)$$

式(1)中，Y 是具有 m 个变量、n 个样本点的响应矩阵，X 是具有 p 个变量、n 个样本点的输入矩阵，B 是回归系数矩阵，E 为噪音校正模型，与 Y 具有相同的维数。将经过预处理后

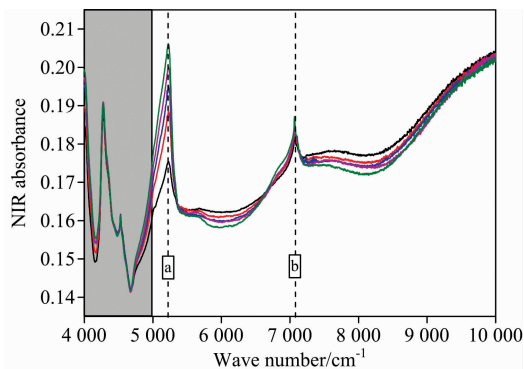


图 3 同一样本不同湿度下的光谱图

Fig. 3 Spectra of the same sample at different humidities

的光谱数据作为模型的自变量输入 X，待分析元素的浓度作为因变量来建立标定模型 [13]。

2 结果与讨论

模型的好坏主要通过模型相关系数(R²)、交叉验证均方根误差(RMSECV)和预测均方均误差(RMSEP)来判断 [14]，

R^2 越大, RMSECV 与 RMSEP 越小, 模型效果越好。2 组实验的分析结果如图 4 以及表 2 所示。

图 4 展示了 2 组模型的预测结果, 红色方块为无湿度梯度样本建立模型(模型 I)的预测结果, 蓝色三角为有湿度梯度样本建立模型(模型 II)的预测结果。图中横纵坐标分别为 X 荧光测量值与近红外光谱模型的预测值, 点越靠近图中的

直线代表模型的预测能力越强。从图中可以看出 SiO_2 , Al_2O_3 , Fe_2O_3 和 CaCO_3 在模型 I 与模型 II 中的预测结果均较为理想, 但对比图中模型 I 与模型 II, 可以看出在同一横坐标下模型 II 的预测结果(蓝色三角)比模型 I (红色方块)更加靠近直线, 因此模型 II 的预测效果优于模型 I。

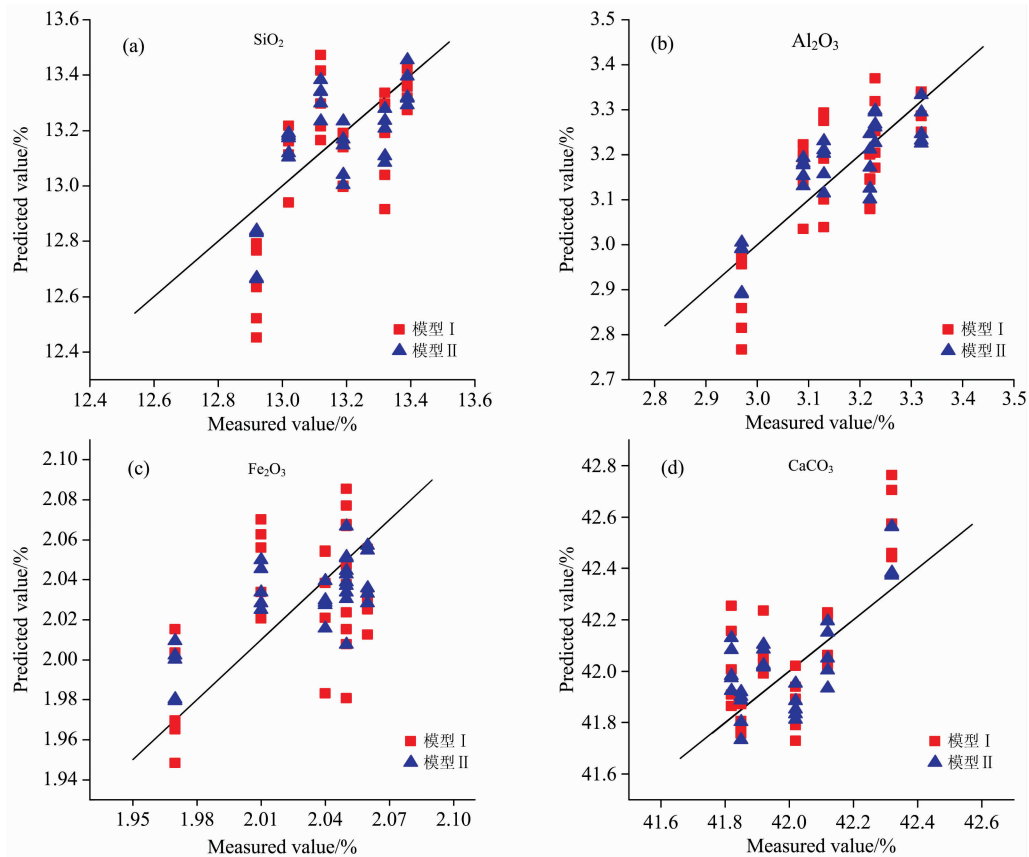


图 4 模型 I 与模型 II 的预测结果对比

Fig. 4 Comparison of predicted results between model I and model II

表 2 两组模型结果指标对比

Table 2 Comparison of evaluation indexes between model I and model II

模型编号	R^2	RMSECV/%	RMSEP/%	
SiO_2	I	0.66	0.16	0.20
	II	0.58	0.18	0.15
Al_2O_3	I	0.77	0.072	0.096
	II	0.69	0.84	0.066
Fe_2O_3	I	0.33	0.034	0.033
	II	0.34	0.034	0.022
CaCO_3	I	0.73	0.14	0.20
	II	0.65	0.17	0.15

表 2 直观的展示了 2 组模型的 R^2 , RMSECV 和 RMSEP 三项指标, 以 4 种成分相互作用比较来看, Fe_2O_3 的模型相关系数 R^2 为 0.33 与 0.34, 明显低于其他 3 种成分, 该结果与已有的近红外发射光谱检测水泥生料文献[8]一致。由于生

产现场水泥生料中 Fe_2O_3 的含量很低, 波动很小, 因此该结果对水泥生料质量检测影响较小, 仍然可以作为生产指导使用。以每种成分的 2 组模型作比较来看, 加入湿度因素后建模指标 R^2 与 RMSECV 会略微变差, 这是由于模型中加入了湿度这一新的干扰因素造成的, 但由于建模过程中考虑了湿度因素的影响, 因此对于验证集光谱中的湿度干扰有较好的鲁棒性, 预测均方根误差 RMSEP 会有明显提升。在使用无湿度梯度样本建立的模型, 即模型 I 中, SiO_2 , Al_2O_3 , Fe_2O_3 和 CaCO_3 的 RMSEP 分别为 0.2%, 0.096%, 0.033% 和 0.2%, 在使用带有湿度梯度样本建立的模型, 即模型 II 中, SiO_2 , Al_2O_3 , Fe_2O_3 和 CaCO_3 的 RMSEP 分别为 0.15%, 0.066%, 0.022% 和 0.15%, 模型 II 比模型 I 减小了 25%, 31.3%, 33.3% 和 25%。

结果表明同一样本的湿度变化会对模型的预测结果有一定的影响, 而在建模时加入带有湿度梯度的样本能有效降低预测误差。

3 结 论

为了研究水泥生料样本湿度对近红外光谱模型预测结果的影响,采用偏最小二乘拟合,使用无湿度梯度的样本建立模型 I,再用有湿度梯度的样本建立模型 II,观察这两个模型对于带有湿度变化的样本的预测能力。对比 2 组模型的预测结果,带有湿度梯度样本建立的模型,4 种成分 SiO_2 ,

Al_2O_3 , Fe_2O_3 和 CaCO_3 的预测均方根误差 RMSEP 分别减小了 25%, 31.3%, 33.3% 和 25%。实验结果表明,由于水分对光具有吸收作用,湿度变化会使光谱产生明显的变化,因此,水泥生料样本湿度对近红外光谱模型预测结果具有一定的影响。采用具有湿度梯度的样本进行建模可有效降低湿度对预测结果的影响,提高检测精度,为优化水泥生产的检测过程提供了一种新的方法。

References

- [1] Ali M B, Saidur R, Hossain M S. *Renewable & Sustainable Energy Reviews*, 2011, 15(5): 2252.
- [2] Yu H, Lian G, Wan X, et al. *Intelligent Control System for Cement Raw Meal Quality Based on Online Analysis*. IEEE International Conference on Cyber Technology in Automation, 2015.
- [3] Tyopine A A, Wangum A J, Idoko E A. *American Journal of Analytical Chemistry*, 2015, 6(5): 492.
- [4] Jan U Porep, Dietmar R Kammerer, Reinhold Carle. *Trends in Food Science & Technology*, 2015, 46(2): 211.
- [5] Posom J, Saechua W, Sirisomboon P. *Renewable Energy*, 2017, 103: 653.
- [6] Machado J C, Faria M A, Ferreira M P L V O, et al. *Talanta*, 2018, 180: 69.
- [7] Ventura M, Silva J R, Andrade L H C, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2018, 188: 32.
- [8] Reboúças J P, Rohwedder J, Pasquini C. *Analytica Chimica Acta*, 2018, 1024: 136.
- [9] Pan D, Crull G, Yin S, et al. *Journal of Pharmaceutical and Biomedical Analysis*, 2014, 89: 268.
- [10] Casian T, Iurian S, Gavan A, et al. *Talanta*, 2018, 188: 404.
- [11] Wülfert F, Kok W T, Noord O E D, et al. *Chemometrics & Intelligent Laboratory Systems*, 2000, 51(2): 189.
- [12] YAN Yan-lu(严衍禄). *Basic and Application of Near Infrared Spectroscopy(近红外光谱分析基础与应用)*. Beijing: China Light Industry Press(中国轻工业出版社), 2005.
- [13] GUO Zhi-wei, SUN Lan-xiang, ZHANG Peng(郭志卫, 孙兰香, 张 鹏). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2019, 39(1): 278.
- [14] Sampaio P S, Soares A, Castanho A, et al. *Food Chemistry*, 2018, 242: 196.

Effect of Humidity on Determination of Main Components in Cement Raw Meal Using Near Infrared Spectroscopy and Compensation Method

XIAO Hang, YANG Zhen-fa, ZHANG Lei*, ZHANG Fa-ye, SUI Qing-mei, JIA Lei, JIANG Ming-shun
College of Control Science and Engineering, Shandong University, Ji'nan 250061, China

Abstract As a new detection method, near infrared spectroscopy (NIR) has been applied to the rapid compositional analysis of cement raw meal. However, environmental factors such as humidity may have a potential impact on its detection. In order to improve the accuracy of the detection, we studied compensation method based on the impact of humidity upon near infrared spectroscopy of cement raw meal. Twenty four cement raw meal samples were obtained from cement factories. Eighteen of them were used as the calibration set; six of them were used as the validation set. The effective ingredients in cement raw meal were SiO_2 , Al_2O_3 , Fe_2O_3 , and CaCO_3 . The standard value of the contents of each ingredient was measured by X-ray fluorescence analysis. Firstly, eighteen samples of calibration set were repeatedly measured five times, and ninety spectra were obtained, which were used to establish model I. Then five humidity gradients were made for eighteen samples of calibration set. The process to generate the humidity gradient sample was as follows: first, the samples were placed on an electric heating platform, flattened with glass rod and heated at 180 °C for 30 minutes, then the samples were placed on the radiation fin to cool. When the sample restored to room temperature, they were taken out and a spectrum was obtained. The samples were placed in the agitator, sprayed with deionized water for two times, and then stirred for 30 seconds to be mixed evenly. After that, the mixed sample was measured to get the next spectrum. Five spectra with certain humidity gradients were obtained by repeating the process. The humidity of all samples was measured by drying method. The range of humidity change was within 0.6%~2%. Each sample with certain humidity was measured once, and these ninety spectra were used to establish the model II. Then, five humidity

gradients were made for the validation set in the same way as the calibration set. Thirty spectra were obtained by detecting each humidity gradient sample in the validation set. All spectra were pre-processed by multivariate scattering correction, and the fitting band was $4\,000\sim 5\,000\text{ cm}^{-1}$. Partial least squares method was used for modeling. Comparing the five humidity gradients of the same sample, we could see that the spectra have the greatest differences at $5\,200\text{ cm}^{-1}$, and there were also obvious differences at other locations, so the humidity change has a significant impact on the whole spectrum. Finally, the root mean square error of prediction (RMSEP) of the 30 spectra in model I and model II were compared. The RMSEP of SiO_2 , AlO_3 , Fe_2O_3 and CaCO_3 in model II was reduced by 25%, 31.3%, 33.3% and 25% compared with model I. The experimental results show that the humidity of cement raw meal sample has a certain influence on the prediction results of near infrared spectroscopy model. Modeling with humidity gradient samples can effectively reduce the influence of humidity on the prediction results.

Keywords Near infrared spectroscopy; Cement raw meal; Humidity; Partial least squares

(Received May 5, 2019; accepted Sep. 10, 2019)

* Corresponding author