

第三部分

近红外光谱分析技术

微信号：qingzhouyuan

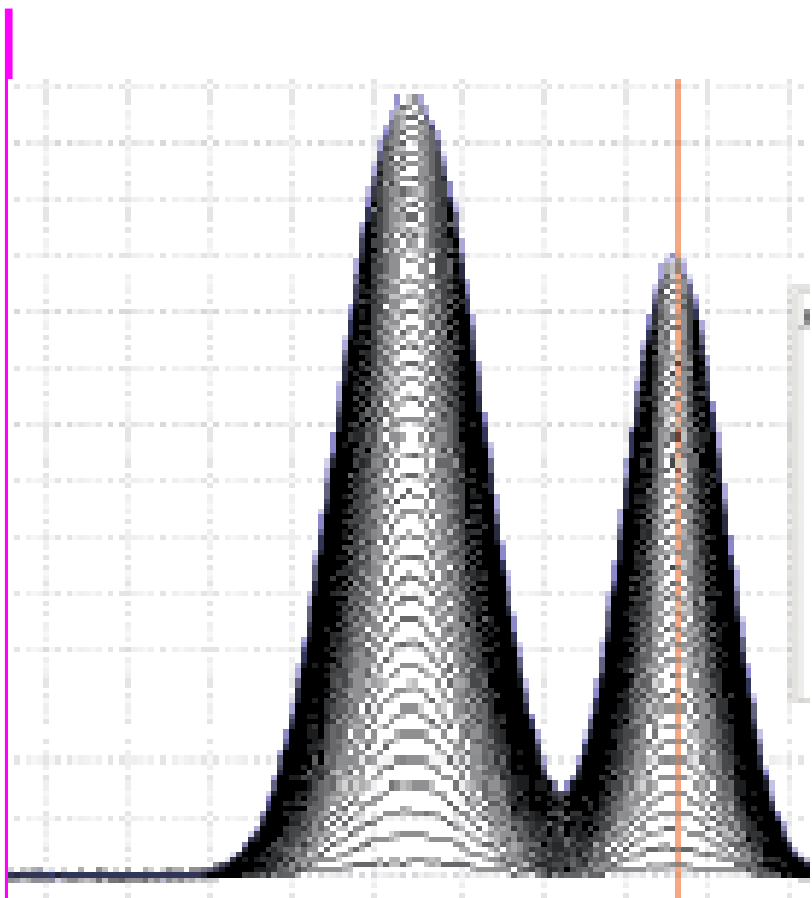
袁洪福

北京化工大学

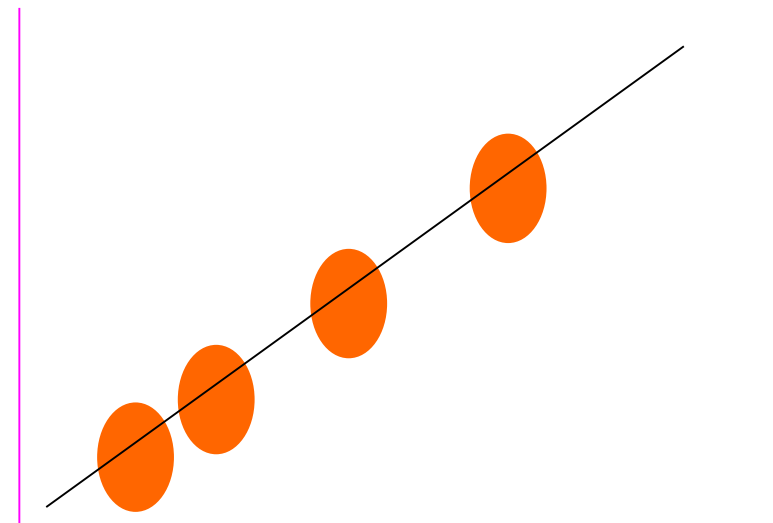
1.定量分析原理

光谱分析的理论基础

$$A = \varepsilon lc$$

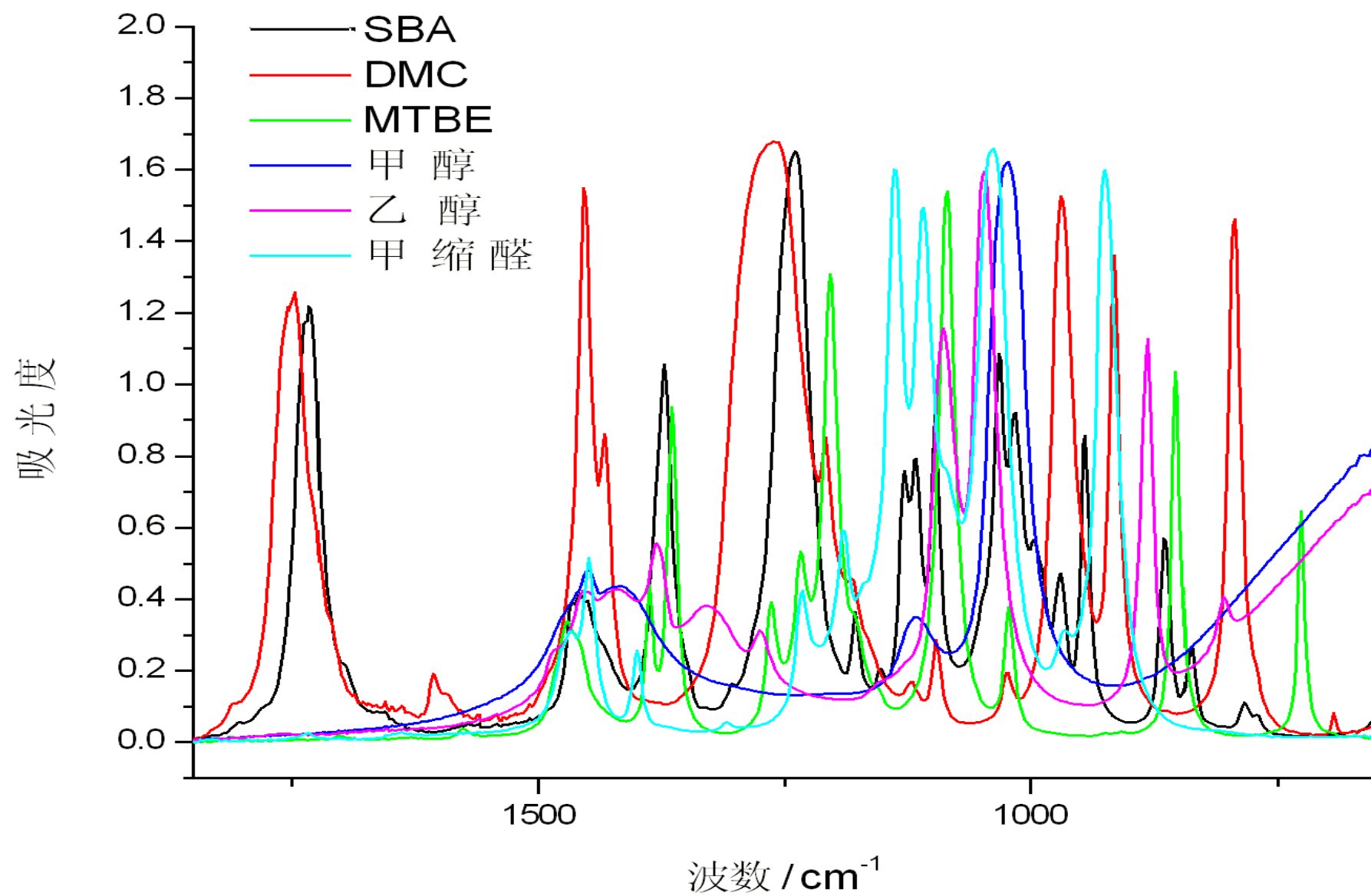


👉 A

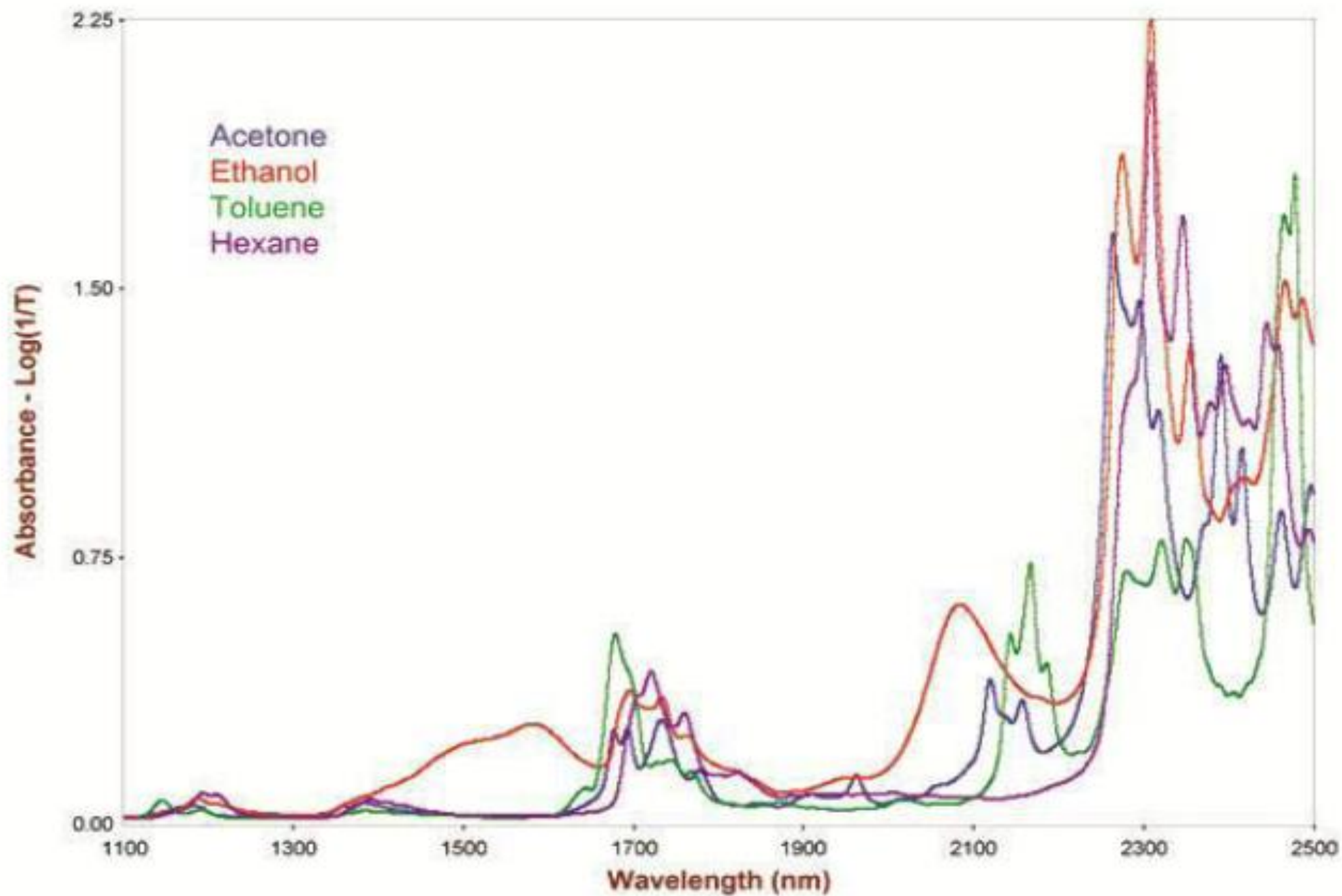


👉 C

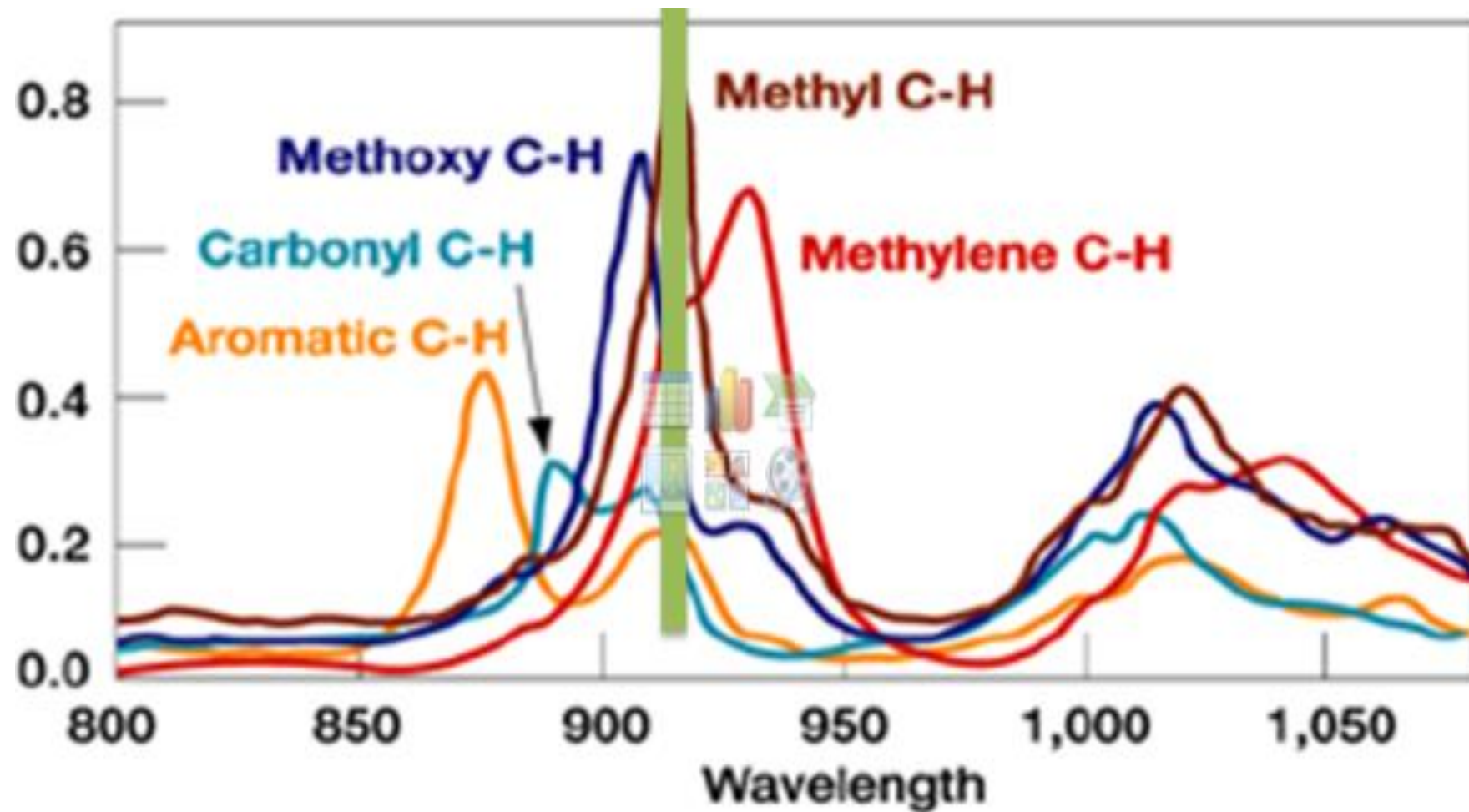
1. 被测组分的信息是分离的
2. 校正标准是纯样品



成分光谱重叠



基团谱带重叠



$$A_{\lambda}^{\text{总}} = \sum_{i=1}^i A_{\lambda}^{\text{组分 } i}$$

混合物的光谱

样品包括 n 种化学组分，假定共存化学组分之间没有相互干扰，则样品光谱数学模型为：

$$\mathbf{a} = c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \cdots + c_n \mathbf{a}_n + \mathbf{e}$$

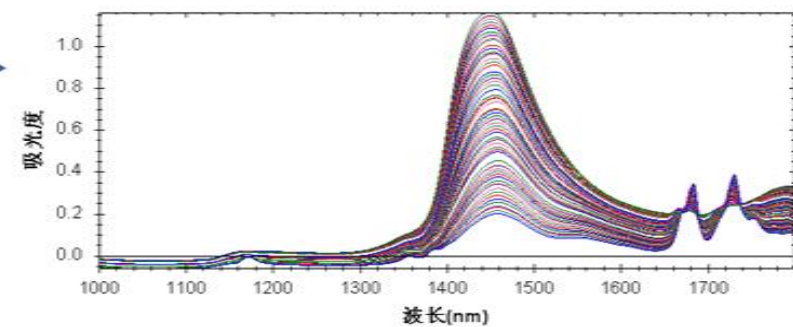
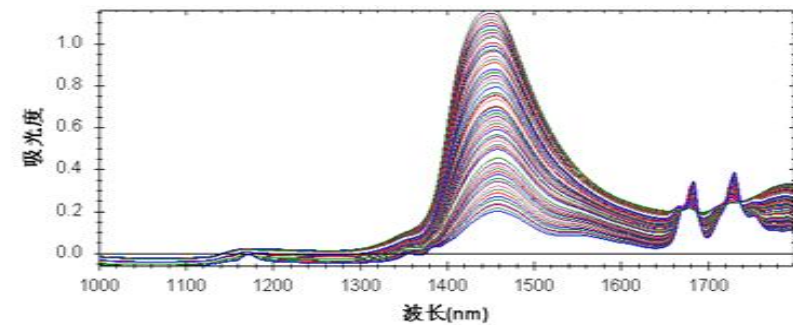
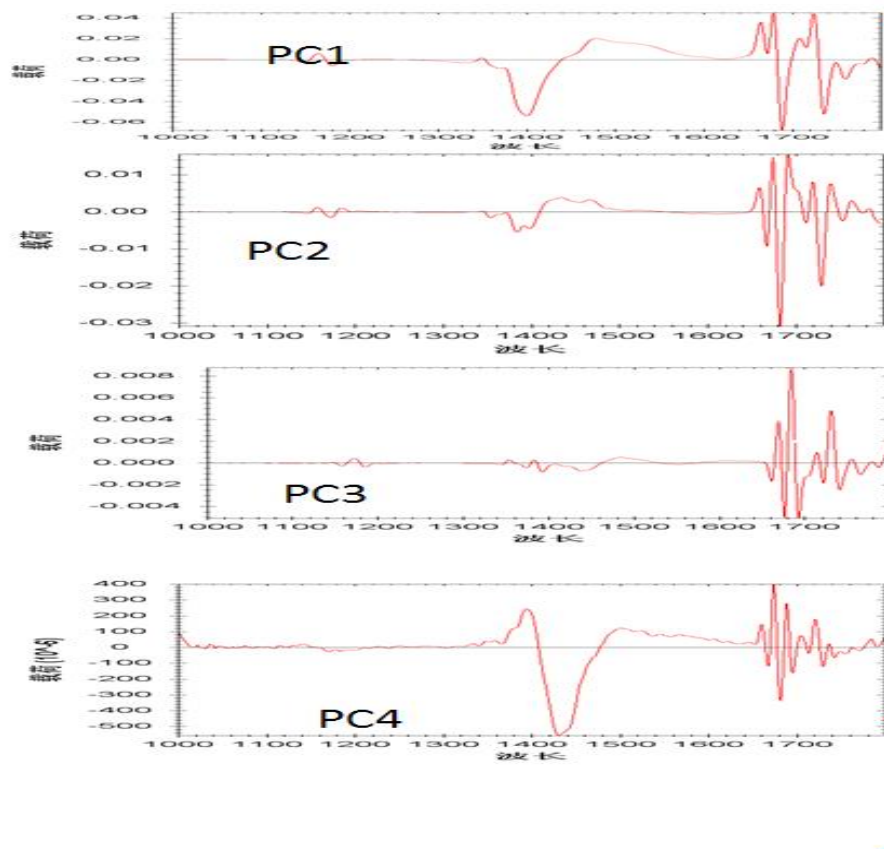
式中， \mathbf{a} 表示混合物的光谱矢量， \mathbf{a}_i 各组分的纯物质光谱矢量， \mathbf{e} 是光谱测量误差矢量。 c 是系数（一般视为相对浓度系数）。

数学分离

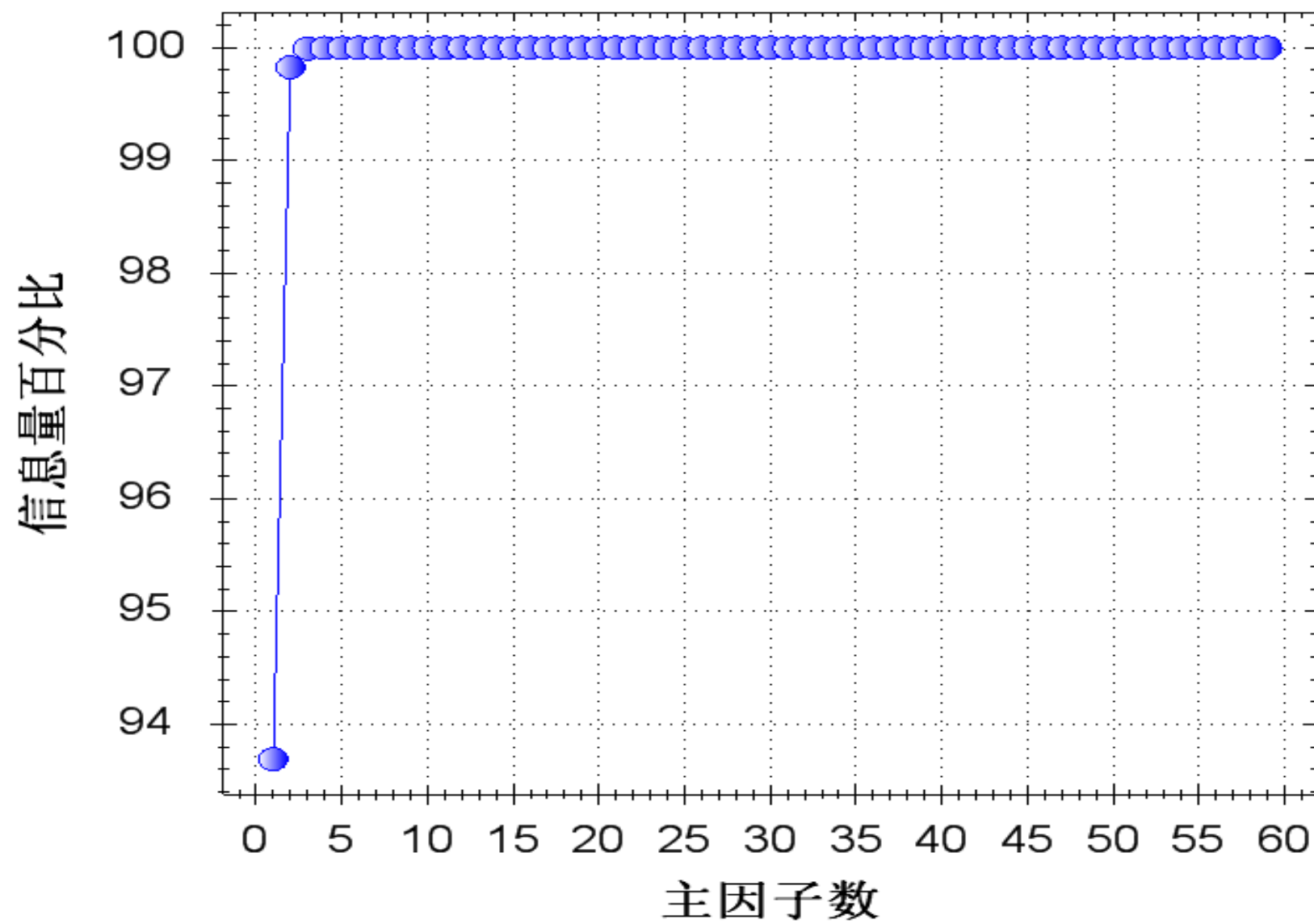
- 建立多元联立方程求解；
- 实际光谱求解过程中存在“共线性”；
- 波长选择或降维处理获得“独立”的光谱变量，进行多元变量求解。

解决共线性问题

主成分分析



用少数主成分光谱表达高维光谱数据



$$\begin{aligned}
 \mathbf{x} &= t_1 \mathbf{p}_1 + t_2 \mathbf{p}_2 + \cdots + t_f \mathbf{p}_f + \sum_{j=1}^n e_j \\
 &= \sum_{i=1}^f \mathbf{t}_i \mathbf{p}_i + \sum_{j=1}^n e_j
 \end{aligned}$$

使用 \mathbf{p} 向量替代波长（变量）作为坐标，使用 \mathbf{t} 值替代吸光度（ \mathbf{A} ），实现坐标系的转换。

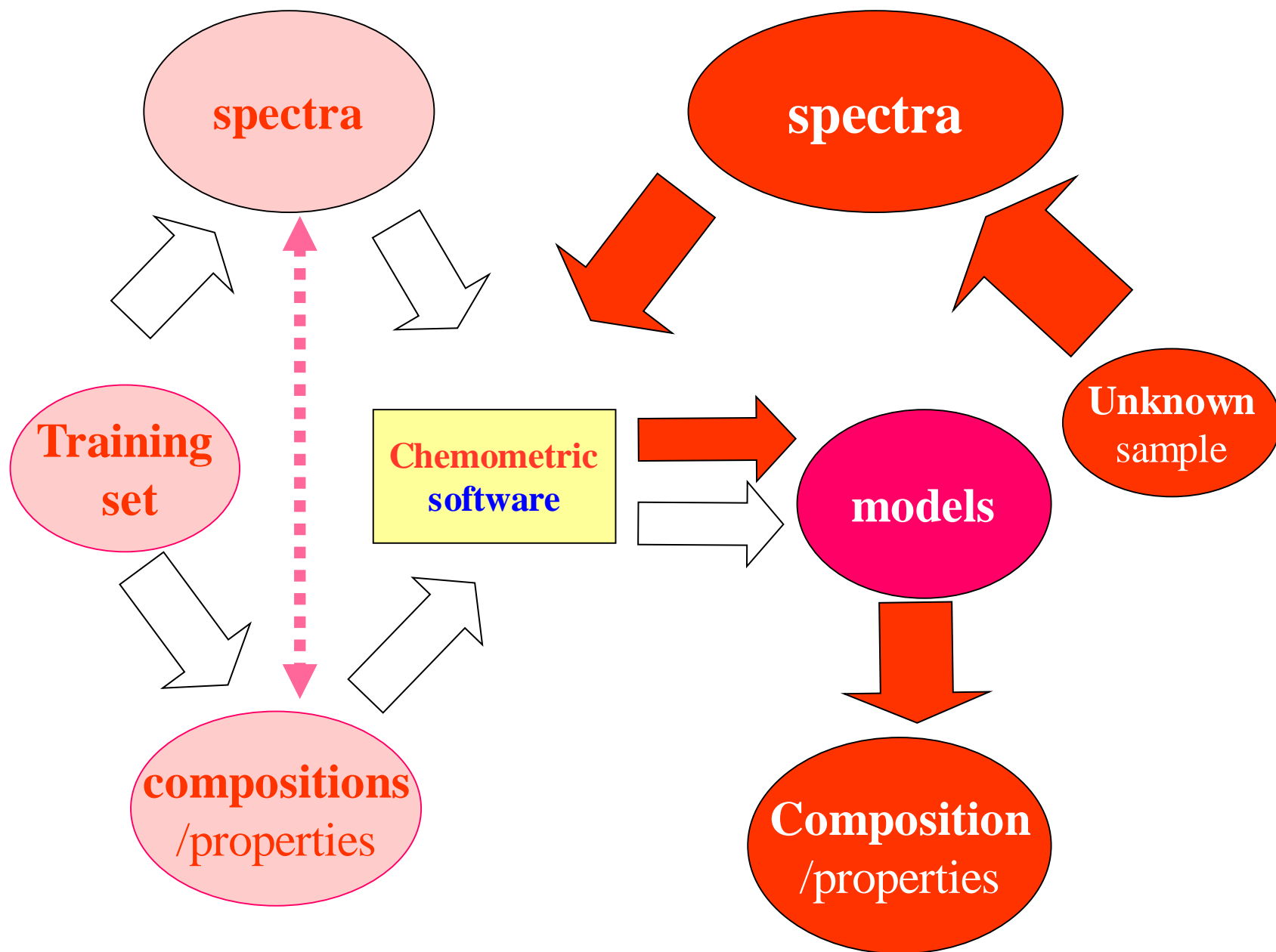
优点：

数据获得降维；

向量之间为正交关系，是独立变量，解决了“共线性”问题。

向量自身在维数上与原光谱数据相同，覆盖了光谱的所有信息通道范围。

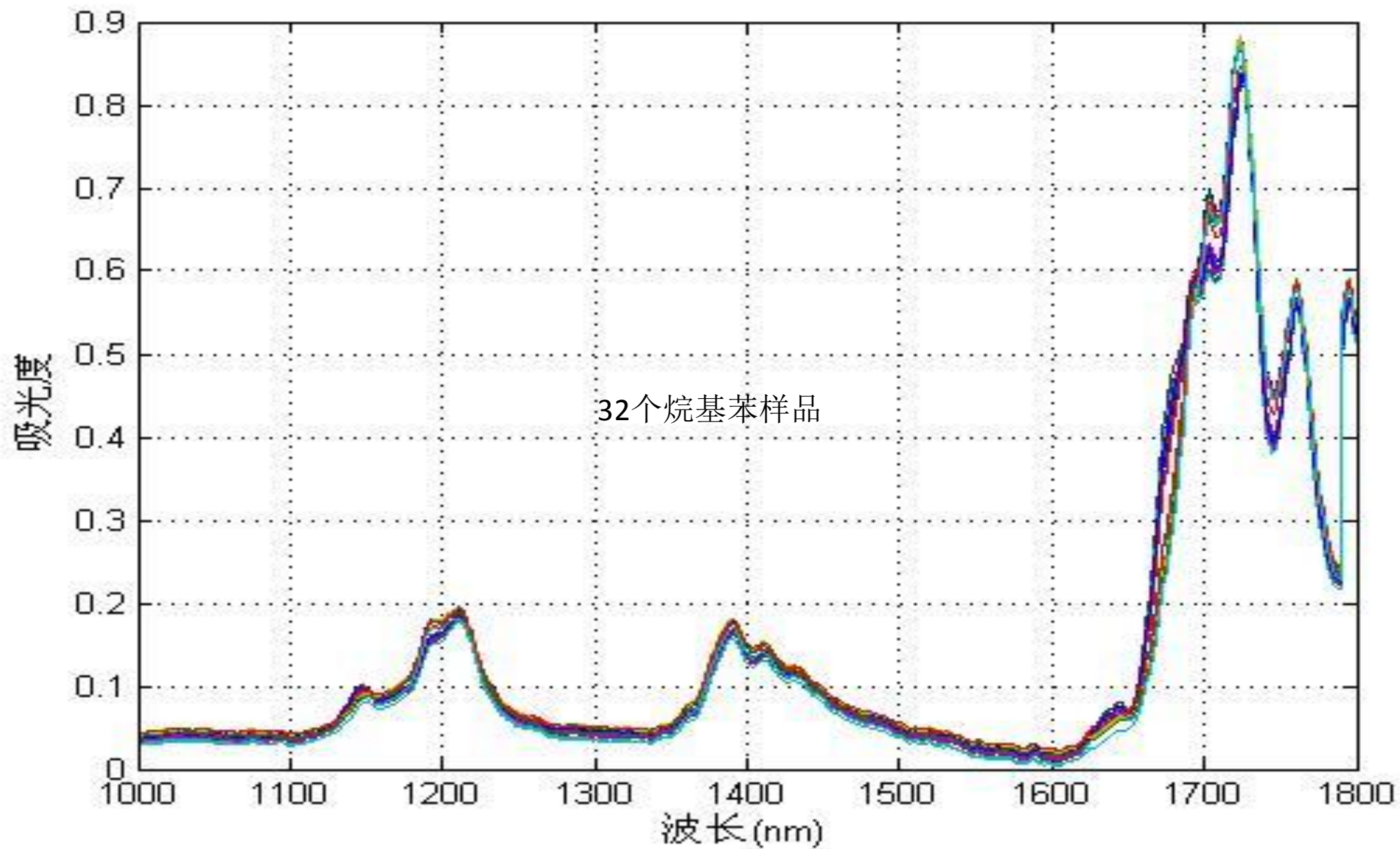
通过选择坐标向量，有效地剥离噪声干扰。

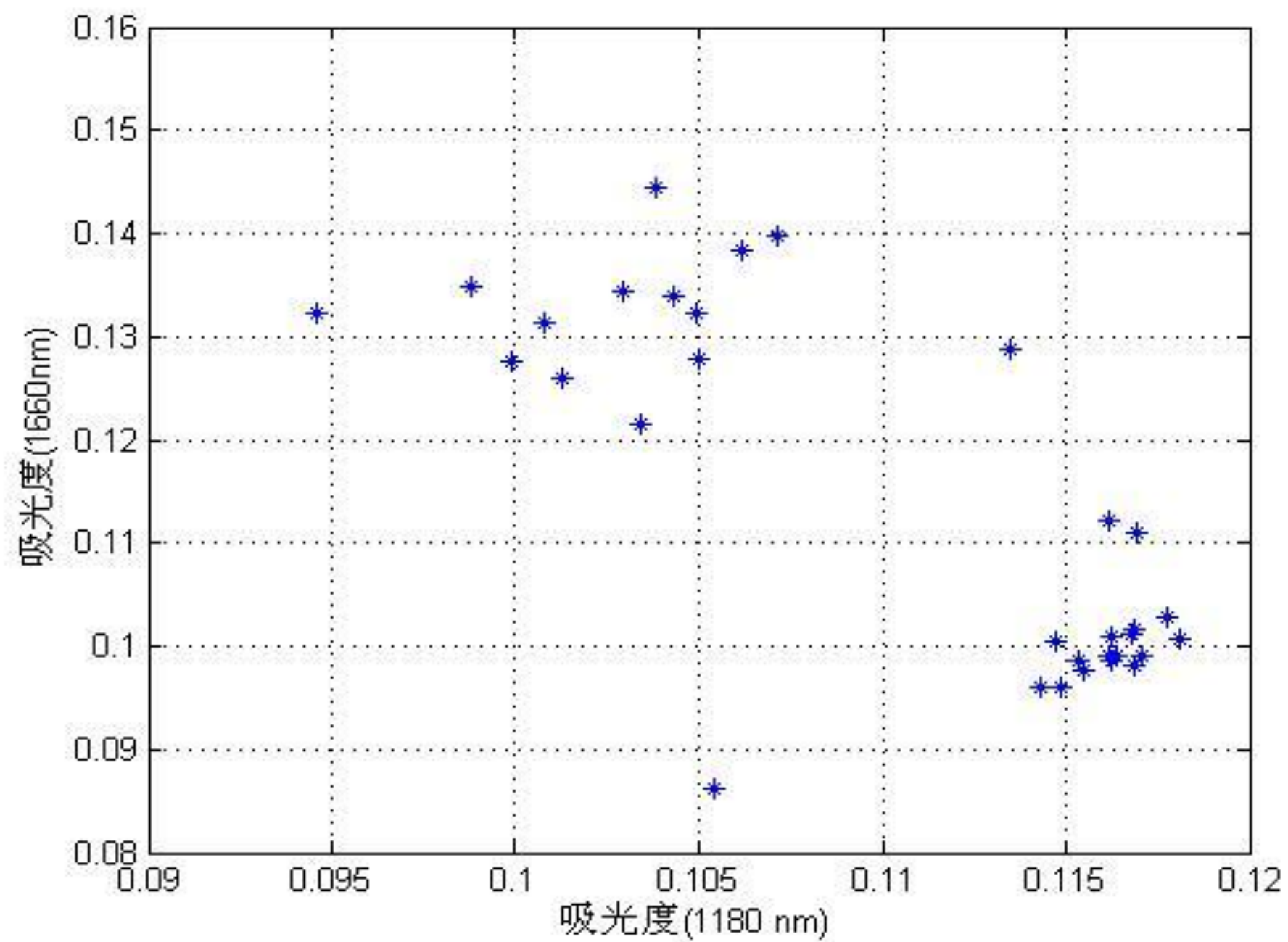


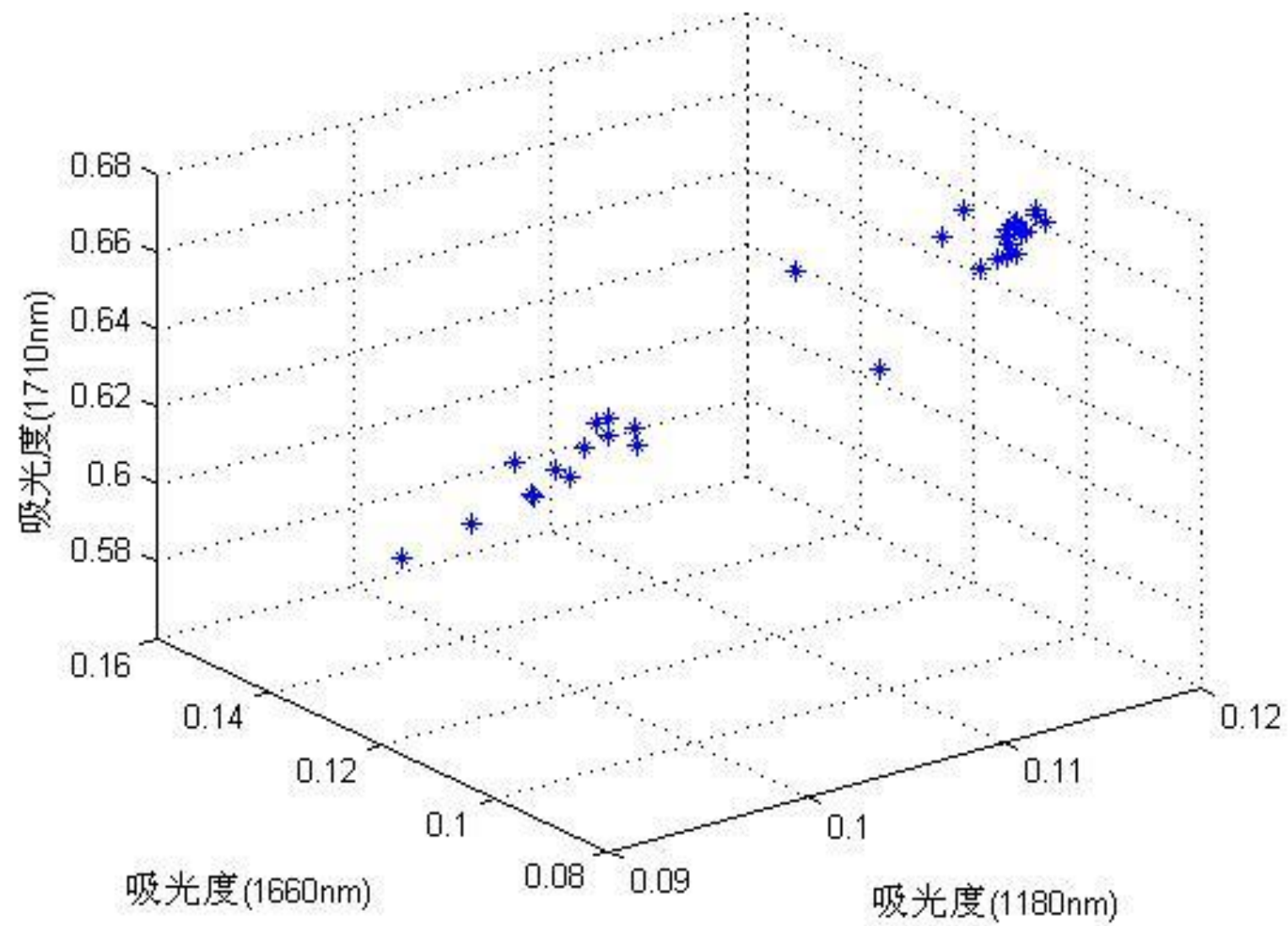
常用多元定量校正方法

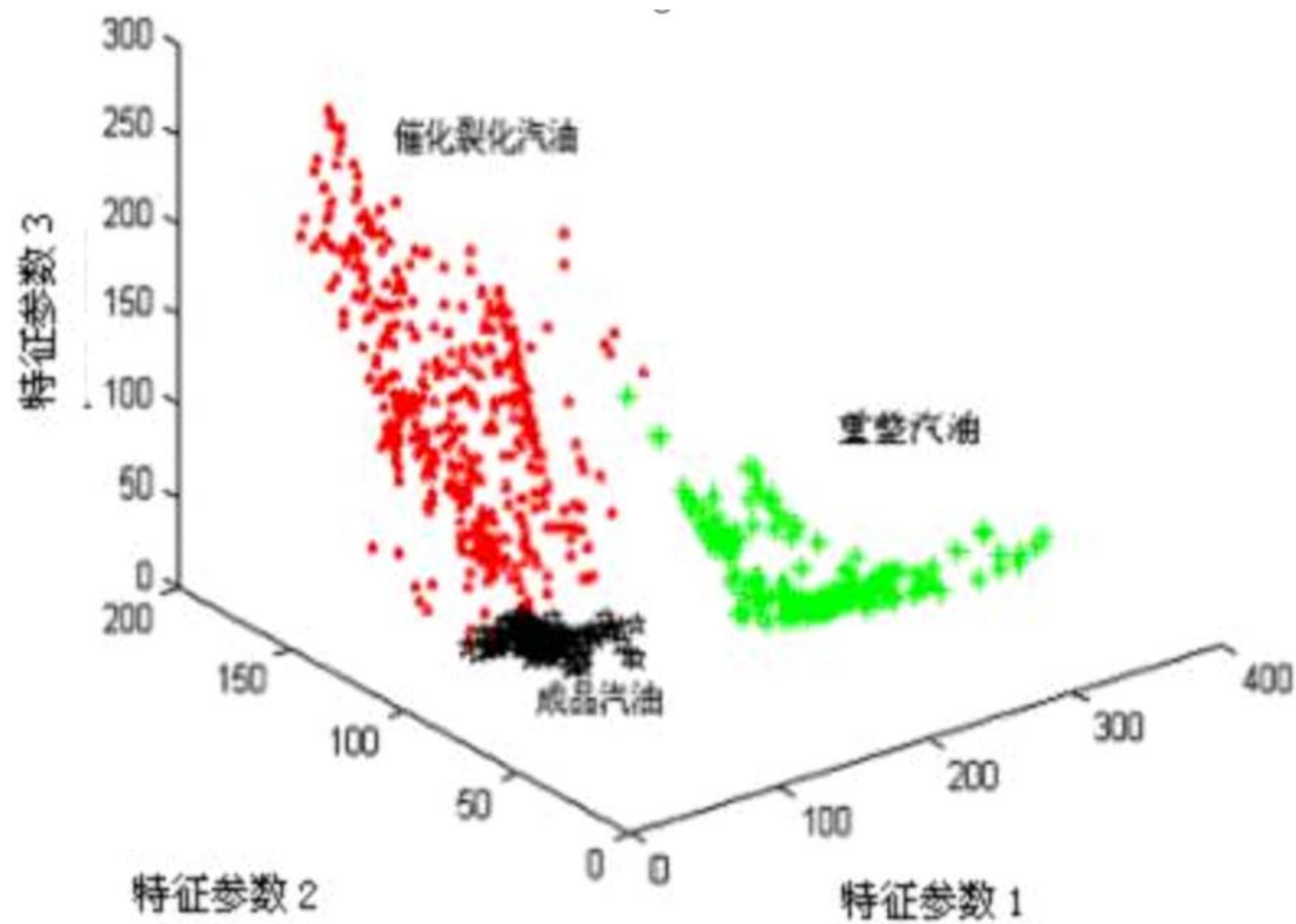
- 多元线性回归（**MLR**）
- 偏最小二乘法（**PLS**）
- 拓扑（**TP**）

定性分析原理





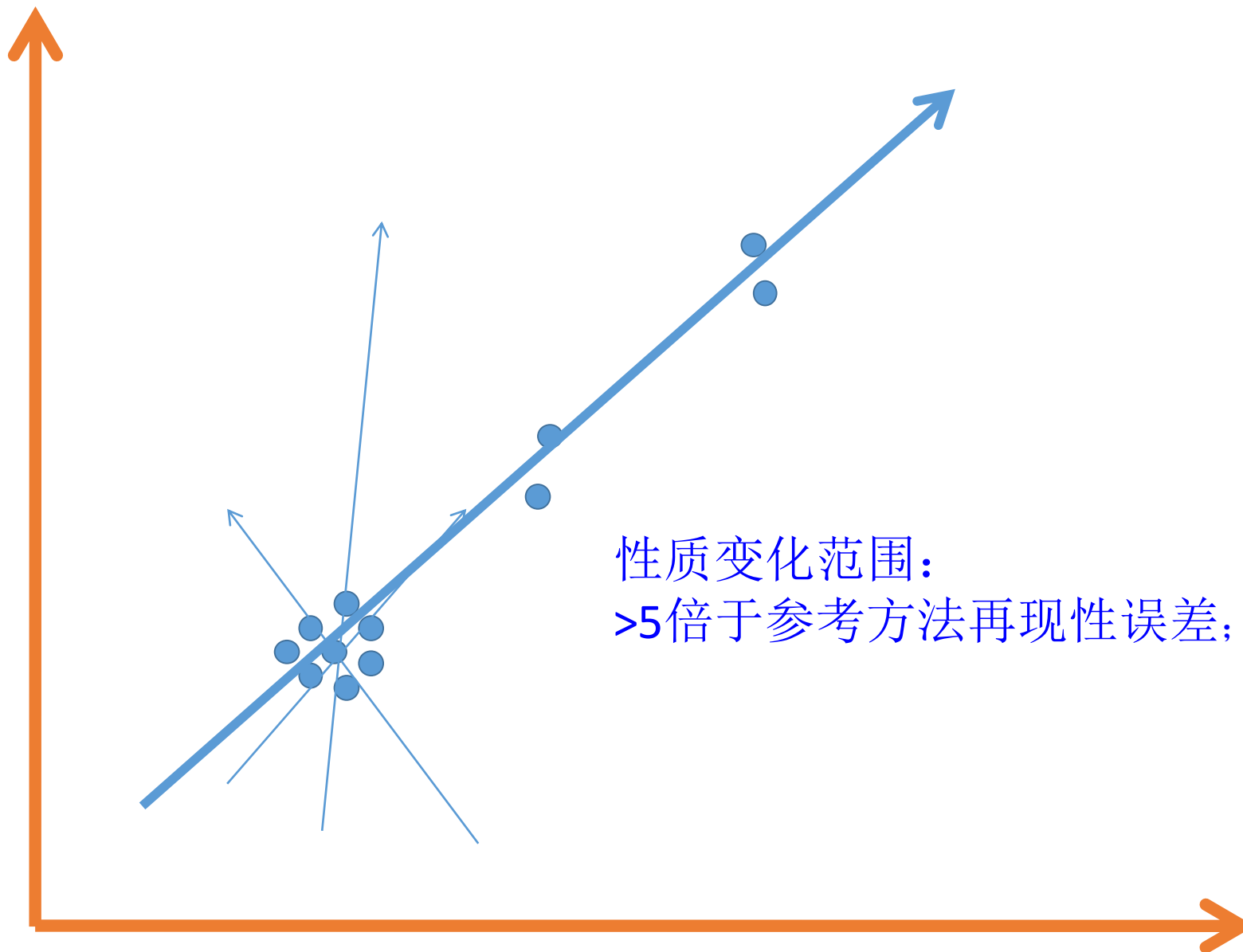




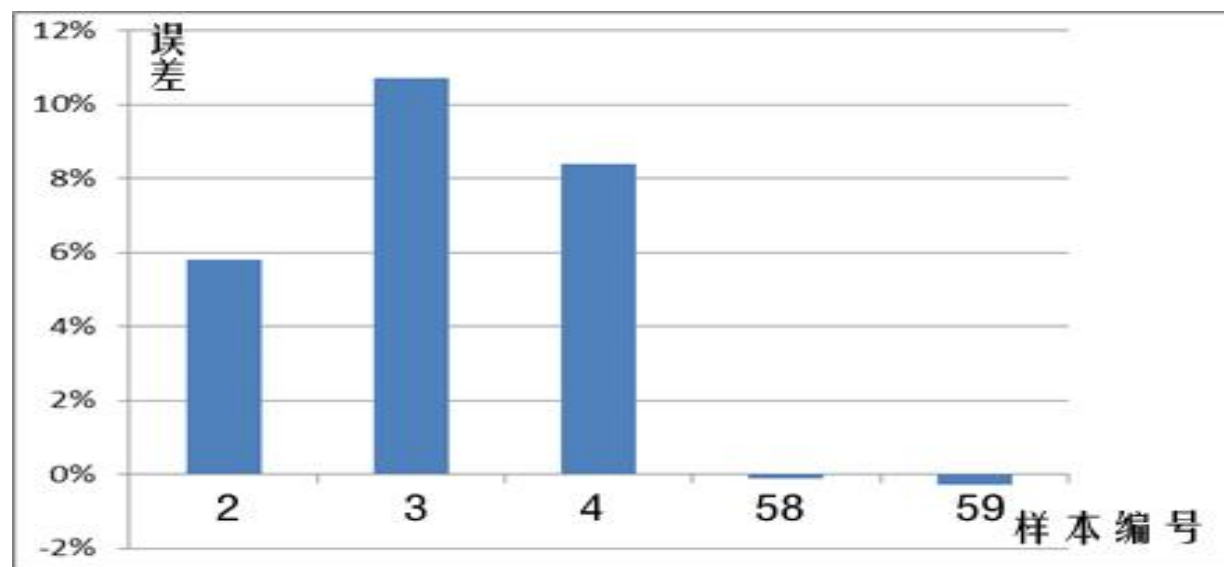
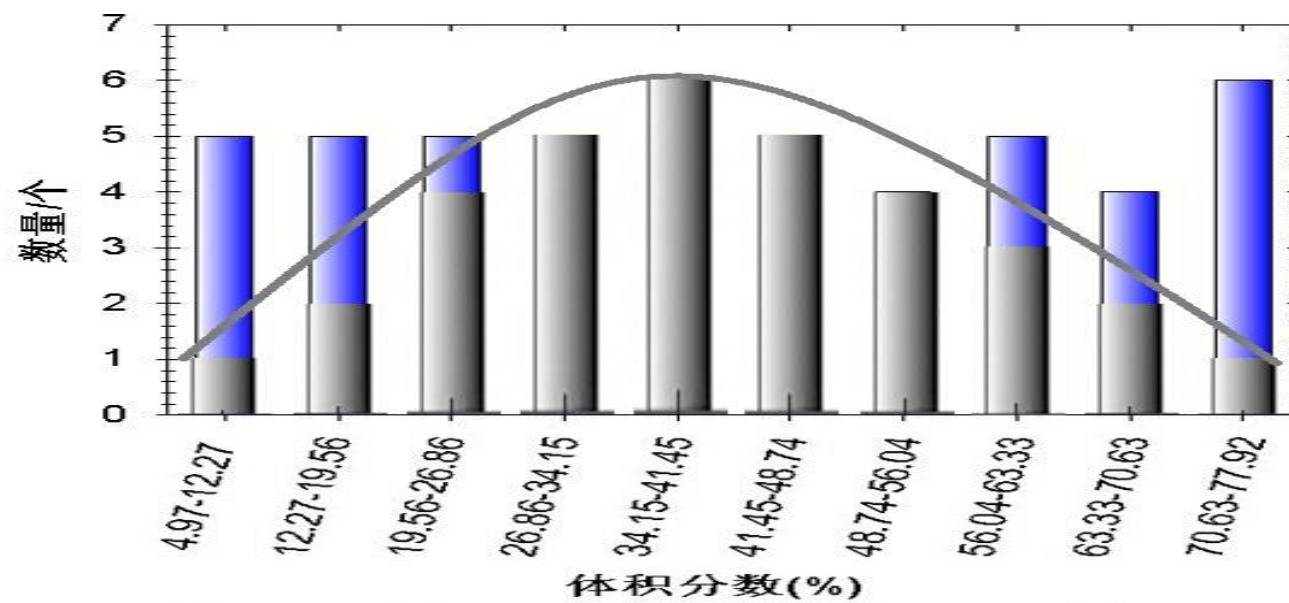
常用分类（classification）鉴别（identification）方法

- 系统树
- PCA
- SIMCA
- PLS-DA

样品收集



a



光谱测量

光程差

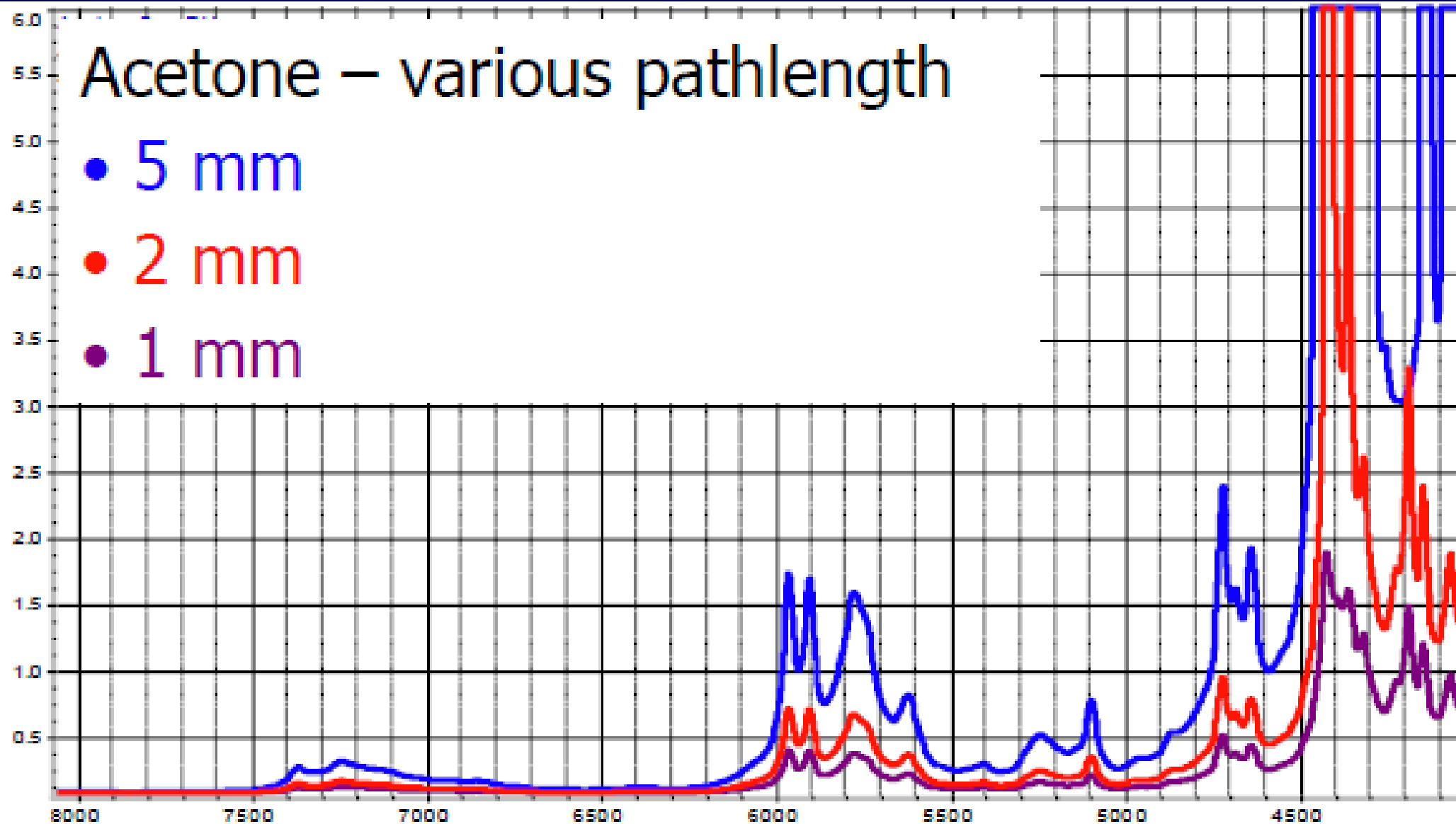
Acetone – various pathlength

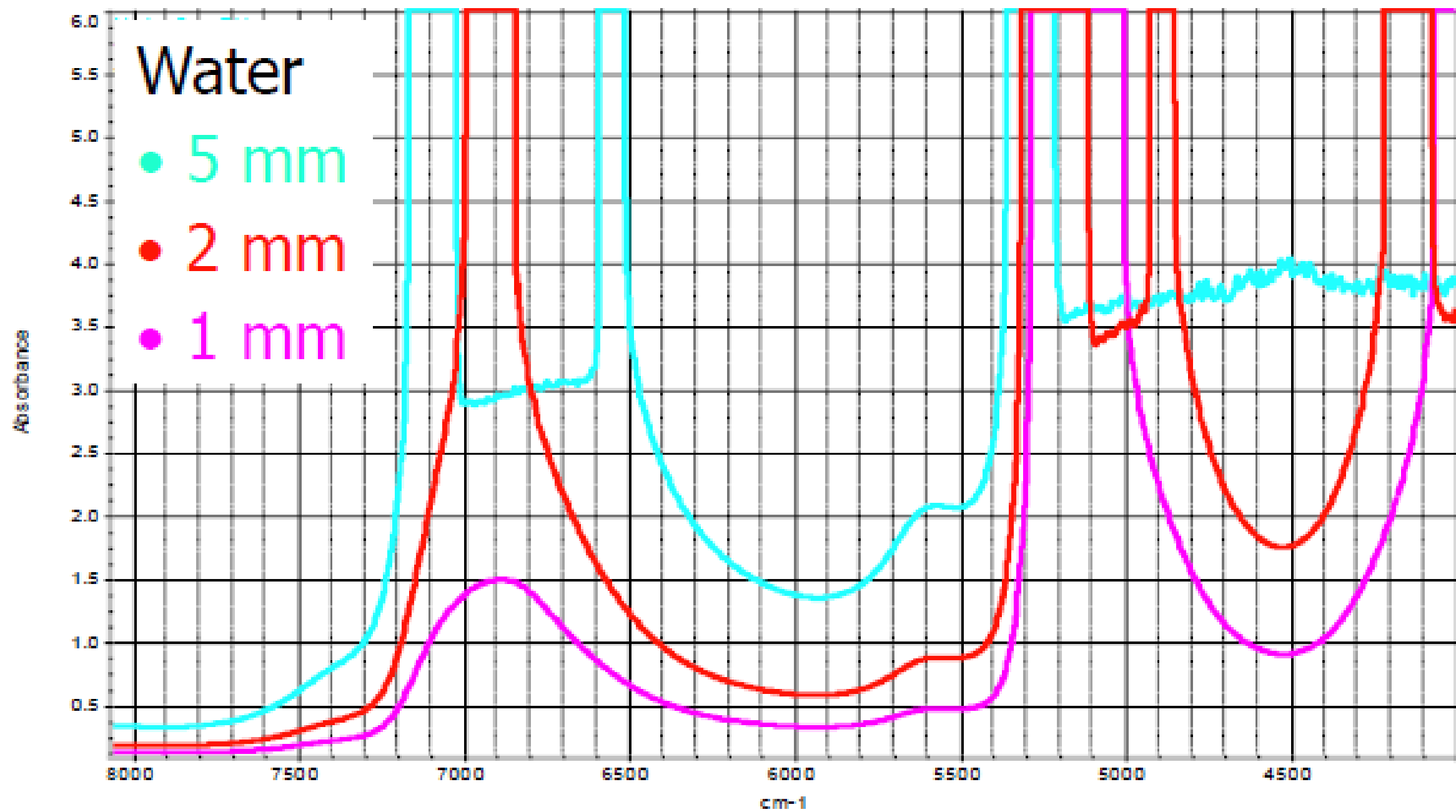
• 5 mm

• 2 mm

• 1 mm

Absorbance





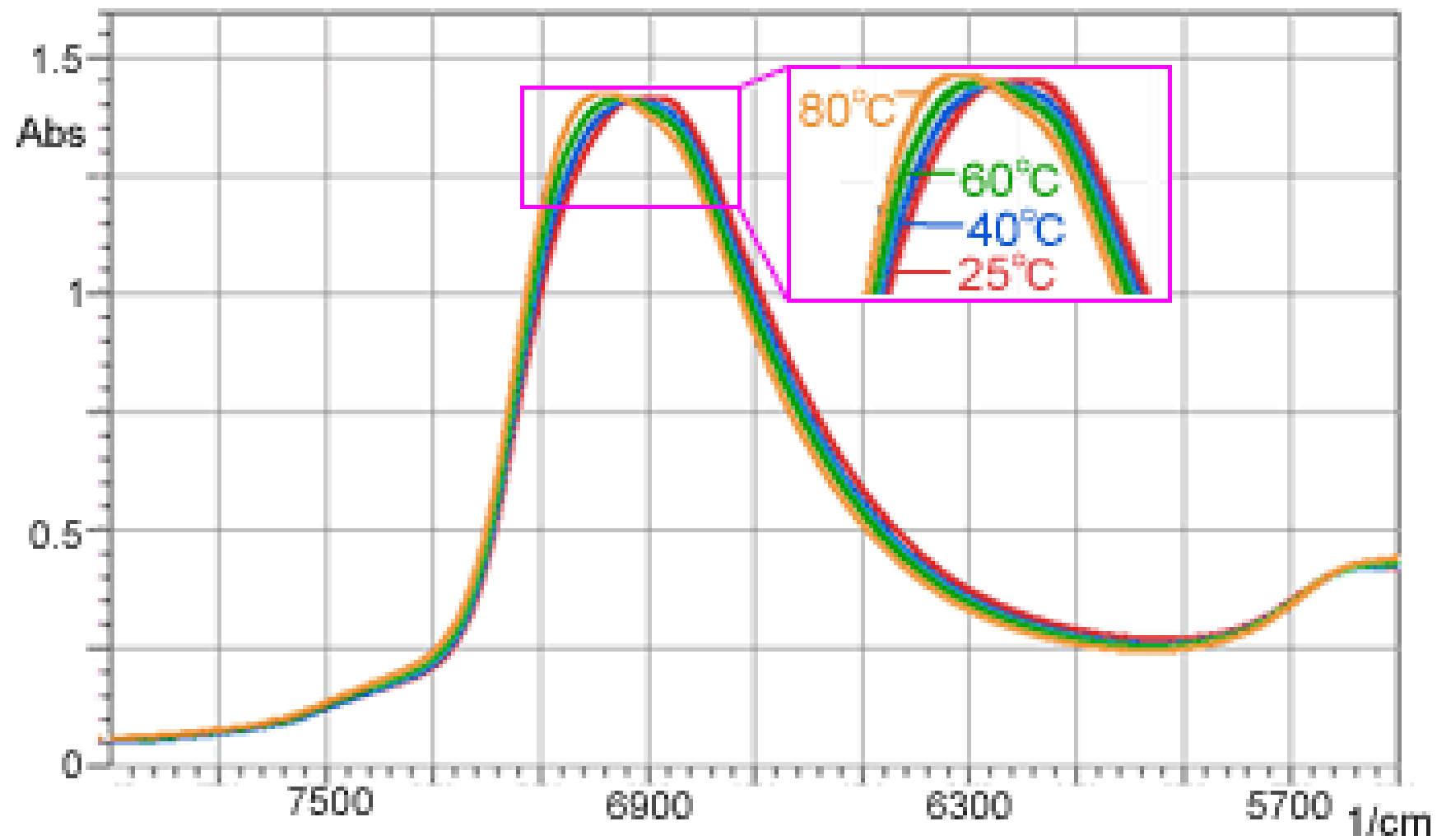


Figure 3 Changes in Peak Wavenumber of Water due to Temperature

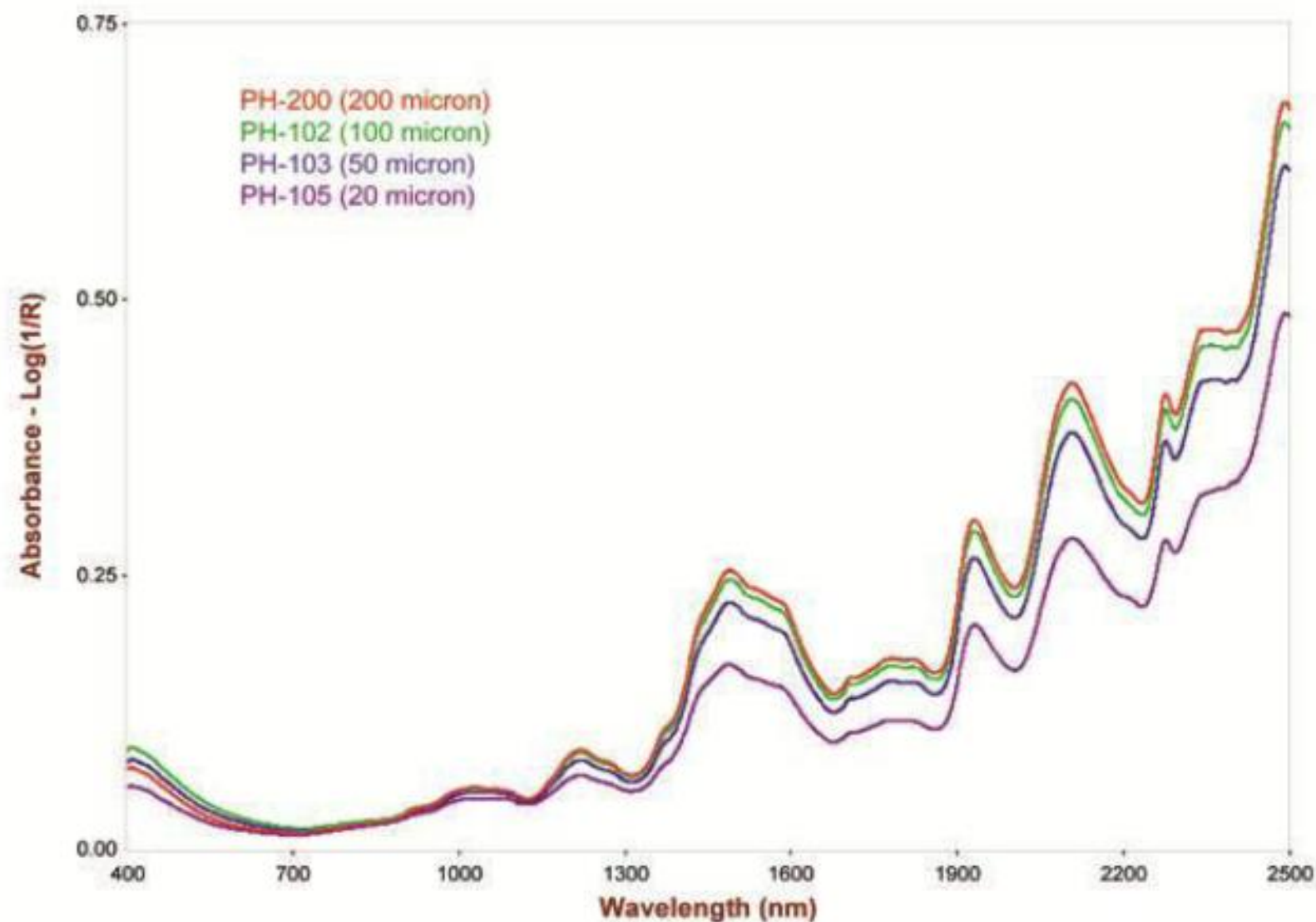


Figure 3: Visible (400–700 nm) and NIR (700–2500 nm) diffuse reflectance ($\log(1/R)$) spectra of differing grades of microcrystalline cellulose (Avicel™) which vary in particle size. Baseline increases as particle size increases due to changes in penetration depth and absorption.

基础数据测定

重复性和再现性

建模与预测



模型参数

+ x

样品集

模型组

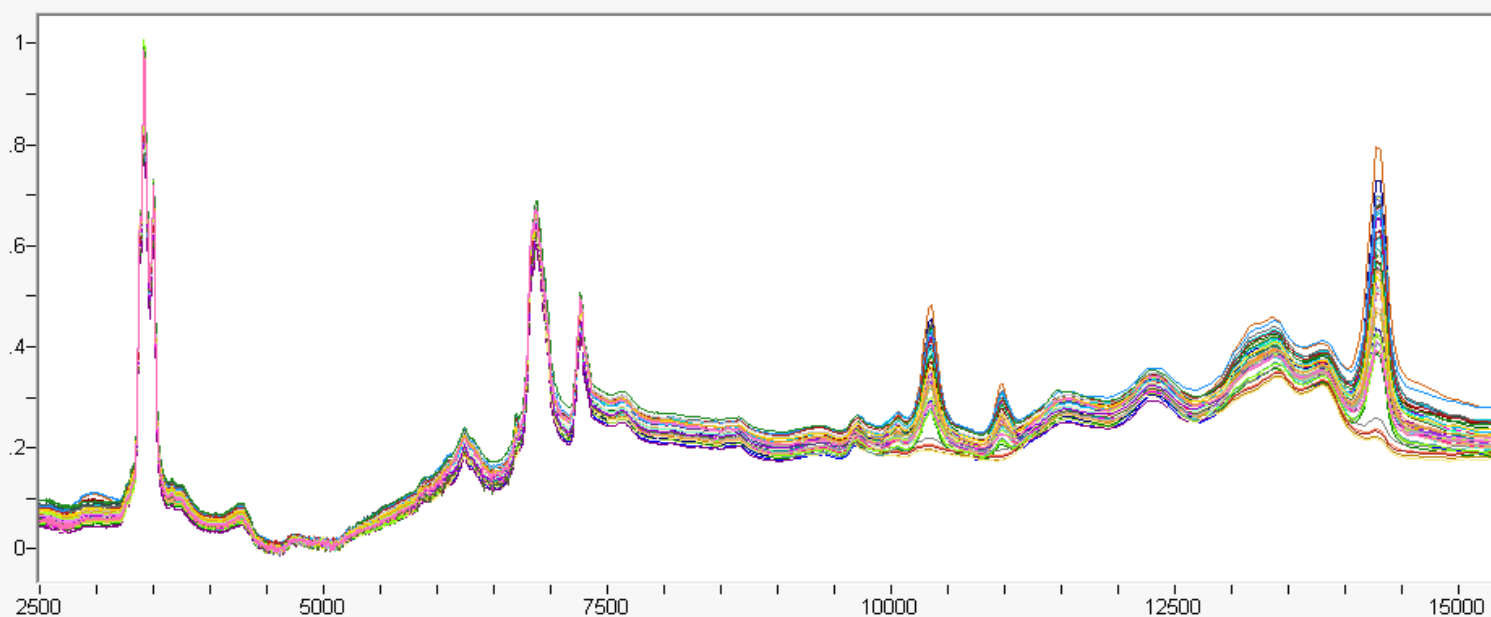
预测

样本号...

—模型组

模型参...

	谱图名称	起始波长	结束波长	文件维数	分类	总沥青含量	胶粉	SBS	C9
1	1#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.924	0.00606	0.06930	0.00103
2	3#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.927	0.00581	0.06650	0.00090
3	4#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.932	0.00538	0.06150	0.00089
4	5#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.943	0.00454	0.05190	0.00079
5	6#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.894	0.00693	0.09830	0.00119
6	7#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.876	0.00809	0.11500	0.00139
7	8#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.798	0.00693	0.19400	0.00119
8	9#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.816	0.00638	0.17700	0.00100
9	10#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.831	0.00582	0.16200	0.00097
10	11#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.843	0.00544	0.15000	0.00090
11	12#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.977	0.01460	0.00561	0.00249
12	13#-mean.txt	2,499.4	15,331.3	1,804.0	无	0.982	0.01400	0.00457	0.00109





模型参数

- 模型组:未命名
 - 模型:未命名1
 - 校正方法:偏最
 - 样品集:59/0/0
 - 预处理:2种方法
 - 波长点:1804/18
 - 模型生成

K-S分集

随机分集

应用

上一步<<

下一步>

样品集 模型组 预测

样品集分集

样品集

序号	谱图名称	状态	总沥青含量
1	1#-mea...	校正	0.924
2	3#-mea...	校正	0.927
3	4#-mea...	校正	0.932
4	5#-mea...	校正	0.943
5	6#-mea...	校正	0.894
6	7#-mea...	校正	0.876
7	8#-mea...	校正	0.798
8	9#-mea...	校正	0.816
9	10#-me...	校正	0.831
10	11#-me...	校正	0.843
11	12#-me...	校正	0.977
12	13#-me...	校正	0.982
13	14#-me...	校正	0.984
14	15#-me...	校正	0.987
15	16#-me...	校正	0.988
16	17#-me...	校正	0.866
17	18#-me...	校正	0.867
18	19#-me...	校正	0.87
19	20#-me...	校正	0.874
20	21#-me...	校正	0.879
21	22#-me...	校正	0.886
22	23#-me...	校正	0.893
23	24#-me...	校正	0.901
24	25#-me...	校正	0.913
25	26#-me...	校正	0.922
26	27#-me...	校正	0.822
27	28#-me...	校正	0.826
28	29#-me...	校正	0.828
29	30#-me...	校正	0.825

参数设置

比例
校正样本所占比80 %

分类方法
分类方法: ☒ 马氏距离
☐ 欧氏距离

确定 取消

校正样本性质值对比图

E及验证样本性质值统计图



模型参数

- 模型组:未命名
 - 模型:未命名1
 - 校正方法:偏最
 - 样品集:59/0/0
 - 预处理:2种方法
 - 波长点:1804/18
 - 模型生成

K-S分集

随机分集

样品集

模型组

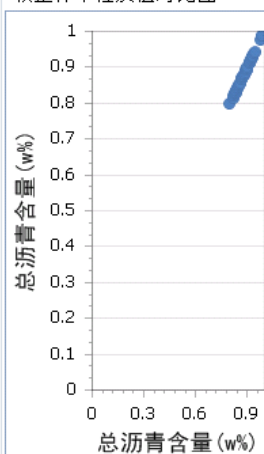
预测

样品集分集

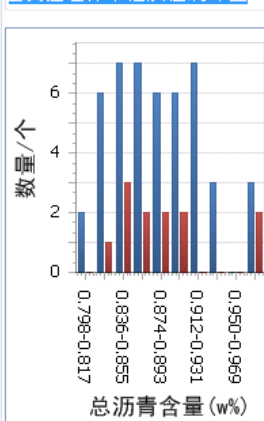
样品集

序号	谱图名称	状态	总沥青含量
1	1#-mea...	校正	0.924
2	3#-mea...	校正	0.927
3	4#-mea...	校正	0.932
4	5#-mea...	校正	0.943
5	6#-mea...	校正	0.894
6	7#-mea...	校正	0.876
7	8#-mea...	校正	0.798
8	9#-mea...	校正	0.816
9	10#-me...	校正	0.831
10	11#-me...	校正	0.843
11	12#-me...	校正	0.977
12	13#-me...	验证	0.982
13	14#-me...	验证	0.984
14	15#-me...	校正	0.987
15	16#-me...	校正	0.988
16	17#-me...	校正	0.866
17	18#-me...	验证	0.867
18	19#-me...	校正	0.87
19	20#-me...	校正	0.874
20	21#-me...	校正	0.879
21	22#-me...	验证	0.886
22	23#-me...	校正	0.893
23	24#-me...	验证	0.901
24	25#-me...	校正	0.913
25	26#-me...	校正	0.922
26	27#-me...	校正	0.822
27	28#-me...	校正	0.826
28	29#-me...	校正	0.828
29	30#-me...	校正	0.825

校正样本性质值对比图



E及验证样本性质值统计图



模型参数

样本号... 模型参...

模型组:未命名
模型:未命名1
校正方法:偏最小二乘法
样品集:47/12/0
预处理:2种方法
波长点:1804/1804
模型生成

方法类型

方法列表

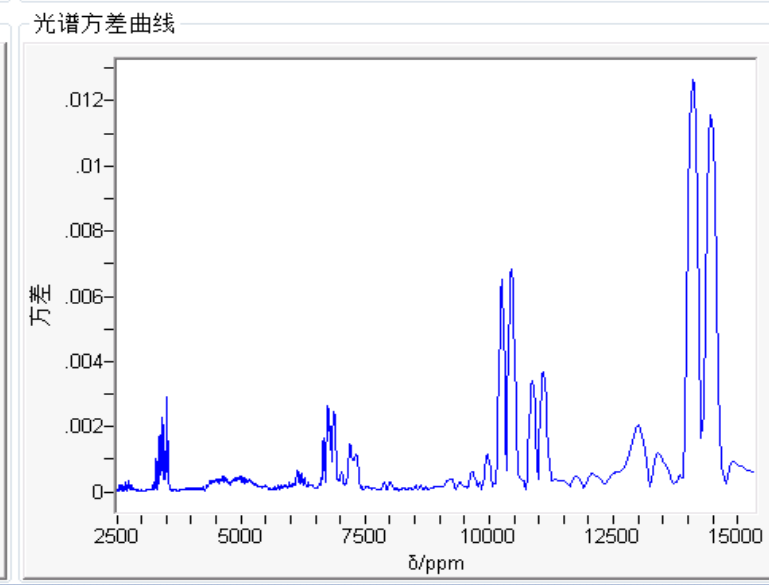
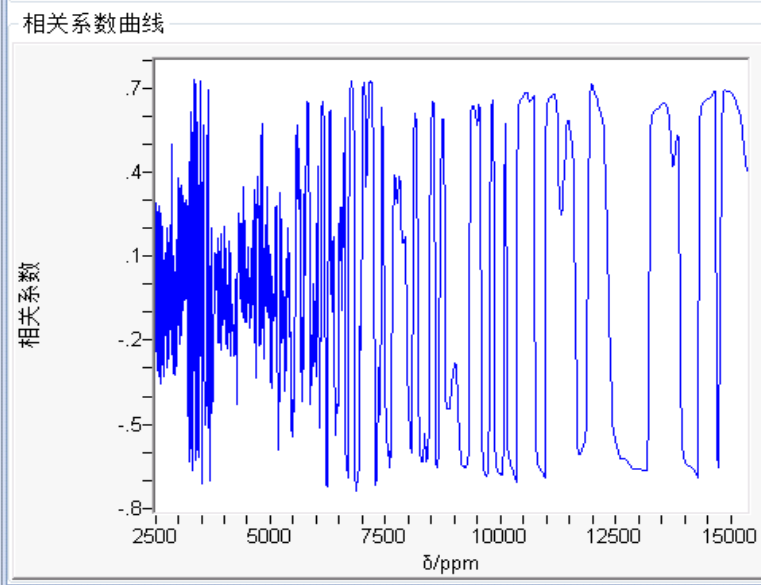
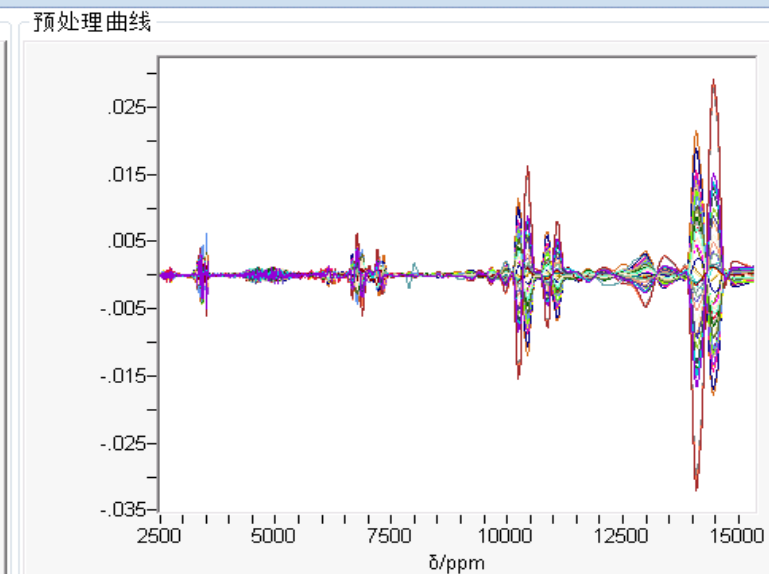
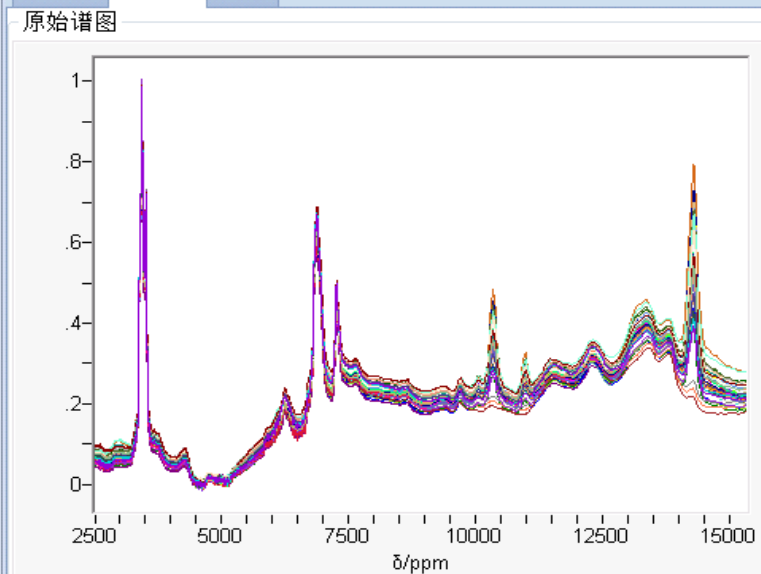
已选方法
均值中心化
Savitzky-Golay导数

清空 删除

应用

上一步<< 下一步>

样品集 模型组 预测



- 模型组:未命名
 - 模型:未命名1
 - 校正方法:偏最
 - 样品集:34/9/16
 - 预处理:2种方法
 - 波长点:1804/18
 - 模型生成

交互检验

计算

推荐主因子: 4

主因子调整

主因子: 8

保存模型

更新

离群样本

- [校正样本]6#-mea
- [校正样本]8#-mea
- [校正样本]13#-me
- [校正样本]25#-me
- [校正样本]26#-me
- [校正样本]46#-me

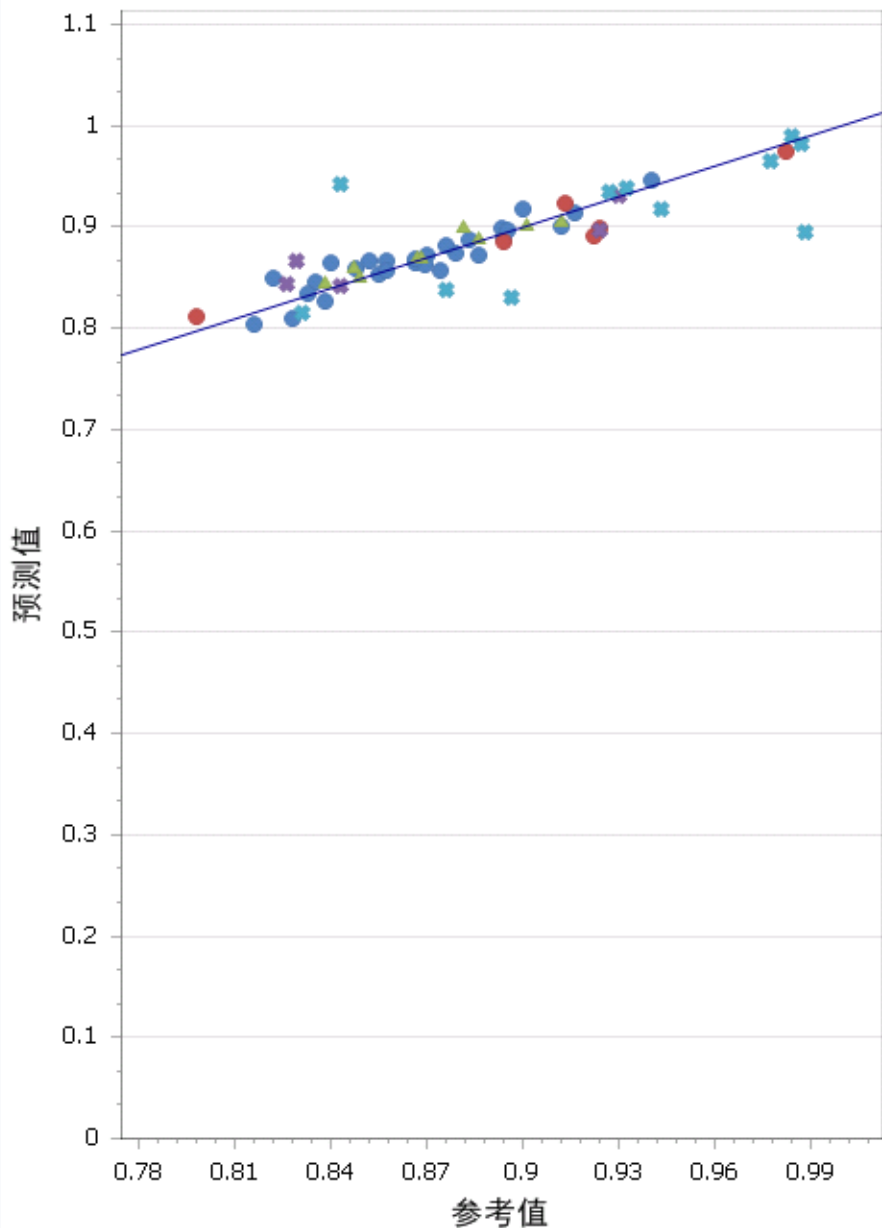
离群图

剔除离群样

视图窗口1

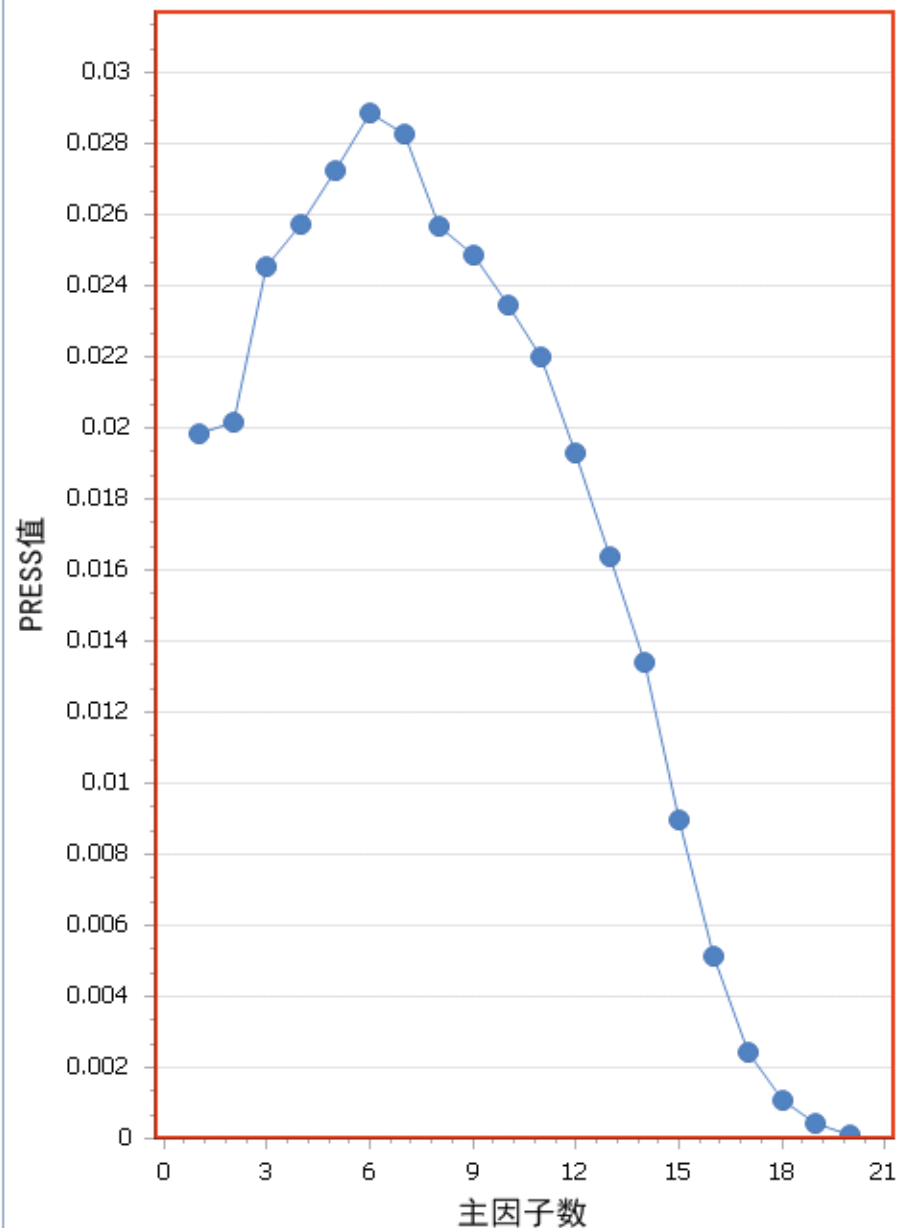
预测值-参考值

全部



视图窗口2

PRESS





样品集 模型组 预测

描述信息:

标题: 未命名1/

模型组: 未命名/

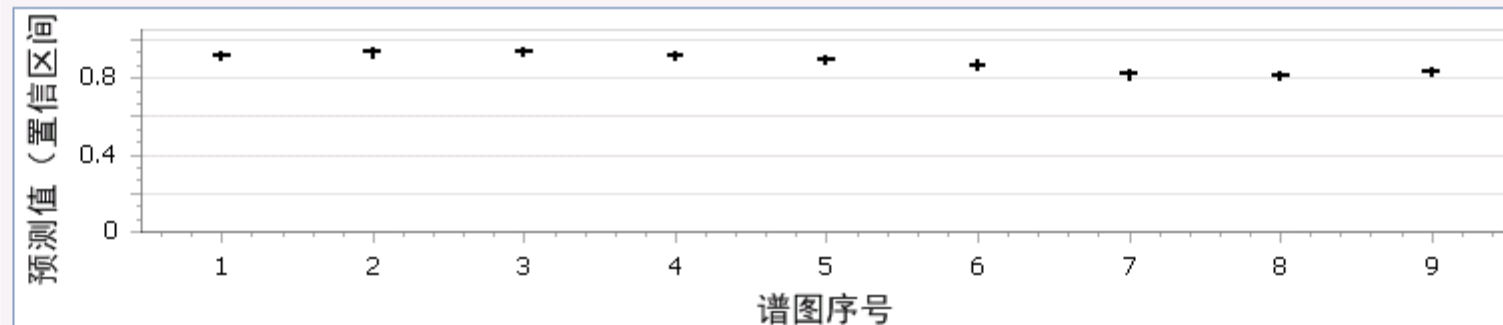
编辑器: /

样品集: 防水材料/

描述:

视图

未命名1(总沥青含量)



应用

序号	谱图名称	总沥青含量	总沥青含量	总沥青含量
序号	谱图名称	outlier	预测值	置信区间
1	1#-mean.txt	否	0.919	0.022
2	3#-mean.txt	否	0.933	0.022
3	4#-mean.txt	否	0.935	0.022
4	5#-mean.txt	否	0.916	0.022
5	6#-mean.txt	否	0.895	0.023
6	7#-mean.txt	是	0.868	0.025
7	8#-mean.txt	否	0.818	0.023
8	9#-mean.txt	否	0.809	0.022
9	10#-mean.txt	否	0.834	0.023



模型参数

✕

添加数据
删除数据

模型组:未命名

模型:未命名1

校正方法选择

样品集:59/0/0

预处理:2种方法

波长选择:256/256

模型生成

建模方法

校正类型: ☐ 定量校正☒ 定性校正

校正方法:

偏最小二乘判别分析
偏最小二乘判别分析(PL)
主成分马氏距离法
SVM
KLBP
聚类树
SIMCA方法

样品集

模型组

预测

含水分析

校正方法选择

性质选择



模型参数

样本号...

模型参...

- 模型组:未命名
 - 模型:未命名1
 - 校正方法:聚类树
 - 样品集:59/0/0
 - 预处理:2种方法
 - 波长点:256/256
 - 模型生成

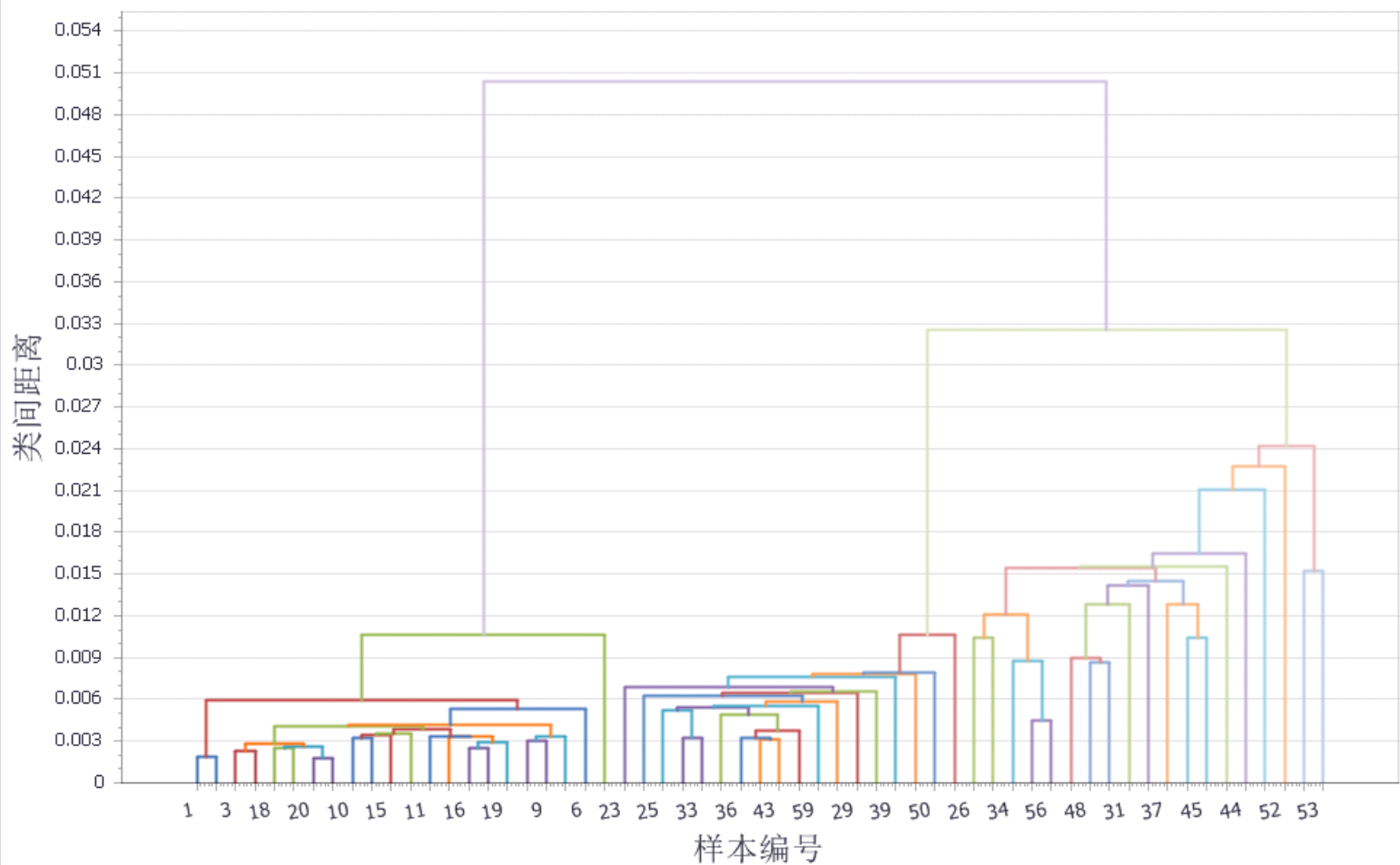
聚类树绘

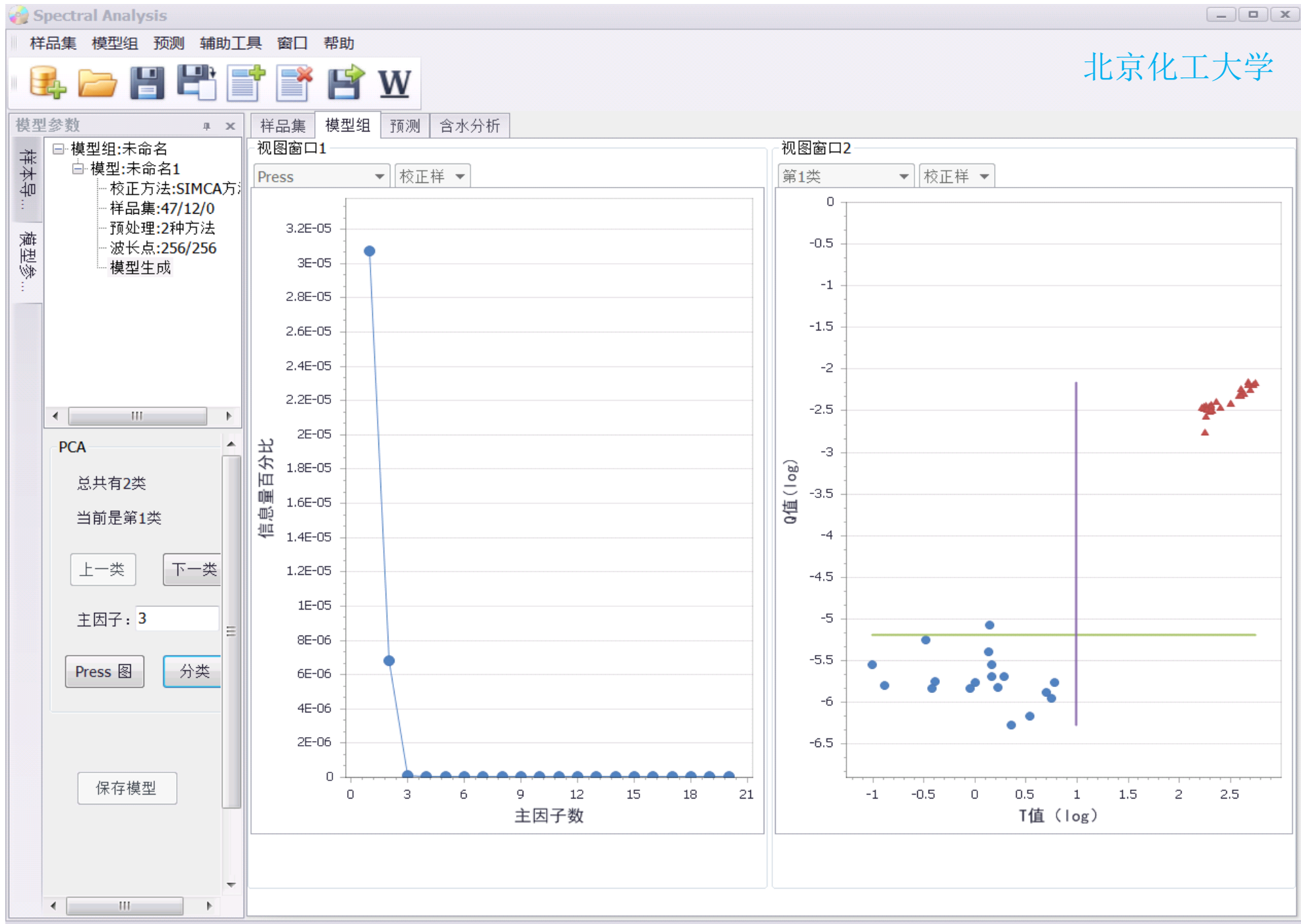
样品集 模型组 预测 含水分析

视图窗口1

聚类树

聚类树

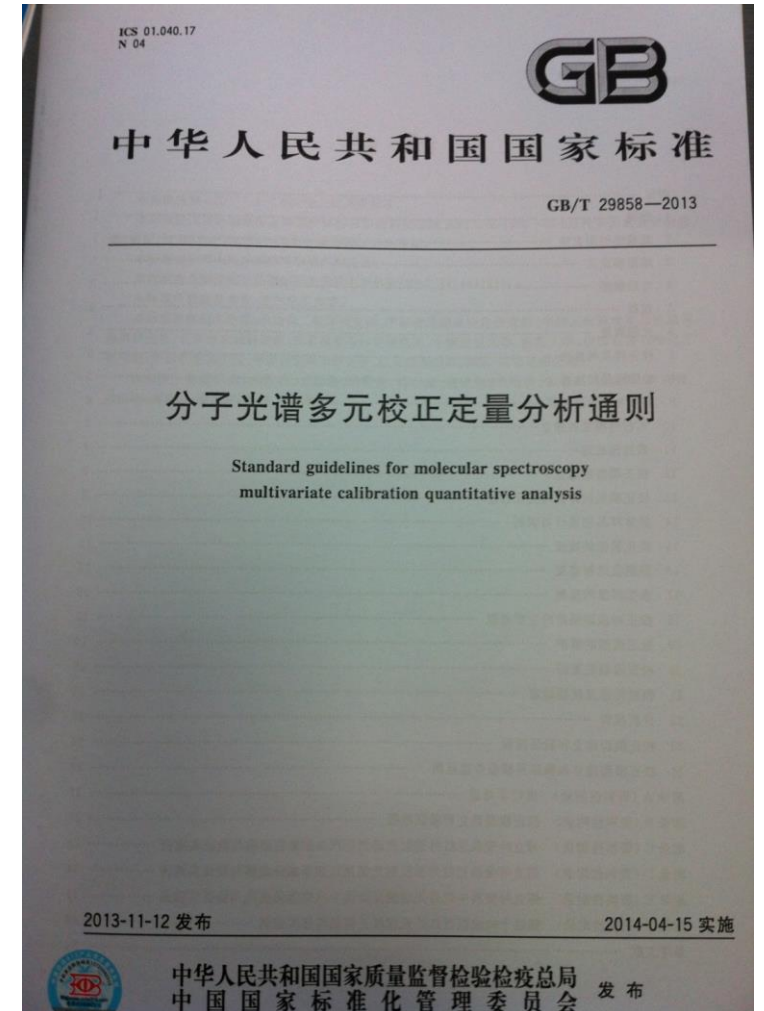




近红外光谱标准

分子光谱多元分析框架GB标准进展情况

- “分子光谱多元分析定量校正通则”（GB/T 28959-2013）；
- “近红外定性分析通则”已经获批GB立项，正在编制之中。
- “近红外光谱仪性能测定方法”已通过TC481标委会批准申报GB。



近红外光谱分析技术特点

- ➡ 可直接测量液体，固体样品；
- ➡ 可直接测量透明和不透明样品；
- ➡ 可直接测量均匀和不均匀体系；
- ➡ 可直接测量静止和可移动样品；
- ➡ 可直接测量低温和高温样品；
- ➡ 可直接测量带压的样品；
- ➡ 通过光纤远距离测量；
- ➡ 大多数样品不做任何样品前处理；
- ➡ 无创伤分析,环保绿色分析；
- ➡ 实验室,野外,现场,离线,在线,原位；
- ➡

- 定性和定量分析；
- 同时测定多种性质；
- 无需制样、无损、速度快；
- 操作方便和智能化；
- 质量监督（真假识别，质量分析）；
- 过程分析与优化控制。

局限性

- 近红外光谱方法适用于常量分析，不适用于微量和痕量成分分析；
- 近红外光谱分析需要建模，工作量大和成本高；
- 缺乏技术标准是制约进一步推广应用的问题；
- 模型传递尚需建立和完善。