

## 近红外光谱的河蟹新鲜度快速检测研究

李鑫星<sup>1</sup>, 姚久彬<sup>1</sup>, 成建红<sup>2</sup>, 孙龙清<sup>1</sup>, 曹霞敏<sup>3</sup>, 张小栓<sup>4\*</sup>

1. 中国农业大学信息与电气工程学院食品质量与安全北京实验室, 北京 100083
2. 中国农业大学烟台研究院, 山东 烟台 264670
3. 苏州大学基础医学与生物科学学院, 江苏 苏州 215200
4. 中国农业大学工学院, 北京 100083

**摘要** 河蟹的新鲜度是大多数消费者在购买时所考虑的最重要的因素, 挥发性盐基氮(TVB-N)是当前国际通用的评价肉类新鲜度的指标, 但其检测工序繁琐、耗费时间长, 无法满足当前市场对河蟹新鲜度评价的迫切需求。因此, 建立一种快速检测河蟹新鲜度的方法是当前急需解决的一大难题。将购于水产市场的河蟹, 采用聚乙烯充氧袋快速运至实验室, 样本数共 126 只。在洁净的工作台上处理后, 将螃蟹分为 42 个实验样品, 每个样品 3 只鲜活螃蟹; 42 个实验样品放在低温 4 °C 的恒温生化培养箱中贮藏, 每天从培养箱中按时取出 6 个螃蟹样品进行光谱数据采集及新鲜度指标 TVB-N 的测定, 历时 7 d。采用近红外光谱(NIRS)对贮藏在不同时间下的河蟹新鲜度进行评价, 使用挥发性盐基氮(TVB-N)作为评价河蟹新鲜度的指标, 首先通过比较经五折交叉验证(5-fold Cross Validation)算法、kennard-stone(KS)算法、光谱一理化值共生距离(SPHY)算法三种样本划分方法处理后所建模型的预测效果确定最优样本划分方法, 最终采用五折交叉验证(5-fold Cross Validation)算法对样本进行划分。其中的 32 个样品被划分为训练集进行模型构建, 其余的 10 个样品被划分为测试集用于模型检验。然后在经过五折交叉验证法对样本进行划分的基础上, 分别采用小波变换(WT)、Savitzky-Golay 平滑、一阶导数法(Db1)、二阶导数法(Db2)这 4 种单一算法以及小波变换(WT)与 Savitzky-Golay 平滑相结合的算法进行预处理, 通过比较预处理后所建模型的预测效果, 确定了小波变换(WT)预处理为最优光谱预处理方法, 从而消除了光谱中的无用信息并提高了信噪比。再次, 在 WT 预处理的基础上, 分别采用主成分分析(PCA)法和连续投影(SPA)算法提取光谱特征波段, 通过建模比较确定主成分分析(PCA)法为最优波长选择方法, 以所选的 16 个特征波长作为模型的输入, 不仅提高了模型的运行速度还可以提高模型的稳定性。最后, 在经过 PCA 特征提取后, 分别采用偏最小二乘回归(PLSR)算法和多元线性回归(MLR)算法构建 TVB-N 定量预测模型, 通过比较两种模型的预测效果, 确定了偏最小二乘回归(PLSR)模型为最优建模方法, 最终确定的最优模型为基于 WT-PCA-PLSR 建立的模型, 模型预测决定系数  $R^2$  为 0.89, 预测均方根误差 RMSEP 为 3.00。综上所述, 所建立的预测模型具有较高的精度, 可以实现对河蟹新鲜度的快速检测, 具有较好的市场应用前景。

**关键词** 近红外光谱; 挥发性盐基氮; 新鲜度; 小波变换; 主成分分析; 偏最小二乘回归

**中图分类号:** O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)01-0189-06

### 引言

河蟹因其美味且富有营养而具有较高的经济价值, 近年来河蟹养殖产业极速扩张<sup>[1]</sup>, 2014 年产量达到 796 535 t<sup>[2]</sup>。

河蟹的新鲜度是大多数人在购买时所考虑的最重要的因

素, 挥发性盐基氮(TVB-N)是当前国际通用的评价肉类新鲜度的指标<sup>[3]</sup>。然而其检测工序繁琐、耗费时间长, 无法满足当前市场快速、客观评价河蟹新鲜度的迫切需求。因此, 建立一种快速检测河蟹新鲜度的方法是当前急需解决的一大难题。

近红外光谱(NIRS)具有快捷简便, 绿色无污染等优点,

收稿日期: 2018-11-09, 修订日期: 2019-03-24

基金项目: 国家重点研发计划项目(2017YFE0111200)资助

作者简介: 李鑫星, 1983 年生, 中国农业大学信息与电气工程学院食品质量与安全北京实验室副教授 e-mail: lxxcau@cau.edu.cn

\* 通讯联系人 e-mail: zhxshuan@cau.edu.cn

并且已经在食品,石油,医药,农业等领域进行了成功应用<sup>[4-7]</sup>。近年来,国内外学者建立了许多基于近红外光谱的肉质新鲜度评价模型,以快速评价肉类的新鲜度。Abdel-Nour 等<sup>[8]</sup>使用近红外光谱检测鸡蛋的新鲜度,结果显示传输光谱学是评估鸡蛋新鲜度和蛋白 PH 值的一个很好的方法。Huang 等<sup>[9]</sup>结合近红外光谱和化学计量学,实现了鱼新鲜度的快速和非破坏性评估。Lin 等<sup>[10]</sup>将近红外光谱与多变量分析相结合检测蛋的新鲜度,结果表明近红外光谱结合多元校准在蛋新鲜度的分析方面具有显著的潜力。Chuang 等<sup>[11]</sup>采用独立成分分析与近红外光谱学相集成的方法来评价大米新鲜度,结果证明该方法具有可行性。

采用近红外光谱评价河蟹的新鲜度,收集腐败变质过程中蟹肉的光谱信息并测定其挥发性盐基氮含量,采用多种光谱预处理方法与特征波段提取相结合,建立了基于近红外光谱的河蟹新鲜度快速评价方法。

## 1 实验部分

### 1.1 材料

实验用的河蟹是 2017 年 12 月购于清河水产市场当天从河中捕捞的鲜活螃蟹,采用聚乙烯充氧袋快速运至实验室,共 126 只。在洁净的工作台上处理后,将螃蟹分为 42 个实验样品,每个样品 3 只鲜活螃蟹;42 个实验样品放在低温 4 °C 的恒温生化培养箱中贮藏,每天从培养箱中取 6 个螃蟹样品将蟹肉取出并放入洁净的培养皿中充分压实,采用近红外光谱仪进行光谱数据采集,采用凯氏定氮仪进行新鲜度指标 TVB-N 的测定,历时 7 d。

### 1.2 仪器

本实验采用的仪器为德国 BRUKER 公司生产的 MPA 型 Fourier Transformation 近红外光谱仪。



图 1 MPA 型光谱仪

Fig. 1 MPA spectrograph

### 1.3 近红外光谱采集

光谱仪的扫描波数范围为: 3 895~11 988  $\text{cm}^{-1}$  (834~2 567 nm), 扫描次数 32 次, 分辨率 16  $\text{cm}^{-1}$ 。将蟹肉取出并放入洁净的培养皿中完全压实。所得到的蟹肉如图 2 所示。

通过重新加载将每个样品连续扫描 3 次, 并将获得的 3 条光谱曲线平均用作待分析的样品光谱。光谱仪所采集到的蟹肉原始光谱如图 3 所示。



图 2 蟹肉

Fig. 2 Crab meat

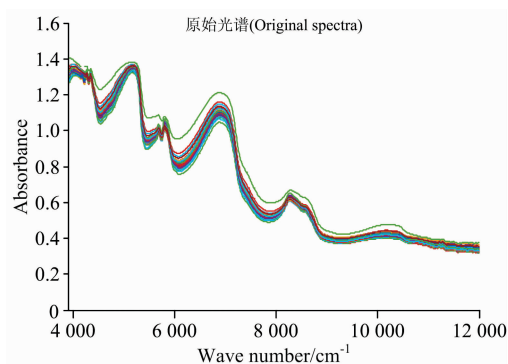


图 3 蟹肉原始光谱曲线

Fig. 3 Crab original NIR spectra

### 1.4 挥发性盐基氮的测定

TVB-N 的测定采用凯氏定氮法。测量方法采用 GB5009.228—2016<sup>[12]</sup>中的方法, 其中每个实验样品做 3 个平行, 取平均值为样品 TVB-N 的含量(单位为  $\text{mg} \cdot 100 \text{g}^{-1}$ )。

### 1.5 模型构建

模型的建立需要经过原始光谱样本集划分、光谱数据预处理、光谱特征波段提取与预测模型构建 4 个步骤, 其中每个步骤都是必不可少的, 每一步的结果都会影响到最终模型的预测效果。

## 2 结果与讨论

### 2.1 样本集的划分

用特定方法划分训练集与预测集是为了评价所用方法对未知输入的拟合效果, 每一次划分的训练数据组合均作为一次独立的建模过程训练模型, 并将对应的测试数据组合作为未知数据评价建模效果。5-fold Cross Validation 法、kennard-stone 法、SPXY 法是常见的三种样本划分方法, 用尽量多的数据组合方式来建模, 这样不仅可以选出效果最好的模型, 还可以对建模方法进行更加全面的评价。分别采用上述 3 种方法将样本光谱划分为训练集与预测集, 其中的 32 个样品被划分为训练集进行模型构建, 其余的 10 个样品被划分为测试集用于模型检验。为了确保所建模型的适用性, 要确保

测试集样品的 TVB-N 含量在训练集样品的含量范围内。

偏最小二乘回归 (partial least squares regression, PLSR) 是自变量与因变量均为两种或以上的线性回归分析, 当各变量内部高度线性相关时, 用 PLSR 非常有效。PLSR 适用于各种参数优化问题以及模型的最优求解问题, 该方法具有典型相关分析、多元线性回归分析和主成分分析的优点。因此, 对每种方法得到的样本划分结果分别采用 PLSR 全波段建模以确定最终的数据划分方式, 结果如表 1 和表 2 所示。

表 1 基于五折交叉验证法的样本划分结果

Table 1 Five results of 5-fold cross validation method for sample selection

样品划分方式	训练集结果		预测集结果	
	$R^2$	RMSEC	$R^2$	RMSEP
组一	0.885 1	3.027 9	0.836 8	6.133 5
组二	0.910 6	2.079 7	0.861 7	5.285 0
组三	0.969 9	1.477 1	0.863 6	3.894 4
组四	0.920 1	2.417 3	0.734 2	5.796 2
组五	0.911 3	2.762 4	0.605 0	4.395 2

注: RMSEC 和 RMSEP 的单位均为  $\text{mg} \cdot 100 \text{g}^{-1}$  (下同)

Note: The units of RMSEC and RMSEP are  $\text{mg} \cdot 100 \text{g}^{-1}$  (the same below)

$R^2$  是模型决定系数, 并且  $R^2$  越趋近 1, 所建模型精确度越高; RMSEC 是训练集均方根误差, RMSEP 是预测集均方根误差, 模型的 RMSEC 和 RMSEP 越小, 所建模型精确

度也就越高。本文采用  $R^2$  与 RMSEP 作为判别模型优劣的指标。由表 1 可知, 采用第三组的样本划分方式所建立的模型决定系数最高, 均方根误差最小, 因此五折交叉验证最终选择第三组的方式进行样本划分。由表 2 可知, 通过五折交叉验证法划分样本所建立的偏最小二乘回归模型训练集与预测集决定系数都在 0.85 以上, KS 算法所建立的模型 RMSEP 最小, 而通过 SPXY 算法所建立的模型预测集  $R^2$  最小且 RMSEP 最大, 模型精度最低。与 KS 算法相比, 五折交叉验证法训练集决定系数、预测集决定系数更大, 训练集均方根误差更小, 预测集均方根误差稍大, 综合考虑模型决定系数与均方根误差, 最终选择五折交叉验证法进行样本划分。表 3 是最终训练集与预测集样本 TVB-N 含量, 单位  $\text{mg} \cdot 100 \text{g}^{-1}$  表示每 100 g 蟹肉中所含有 TVB-N 的含量。

表 2 基于 3 种不同方法的样本划分结果

Table 2 Statistics description of sample selection for the aquaculture sample based on three different algorithms

样品划分方式	训练集结果		预测集结果	
	$R^2$	RMSEC	$R^2$	RMSEP
5-fold CrossValidation	0.969 9	1.477 1	0.863 6	3.894 4
KS	0.826 2	3.818 2	0.709 9	3.281 8
SPXY	0.997 5	0.459 6	0.473 7	4.904 3

注: RMSEC 和 RMSEP 的单位均为  $\text{mg} \cdot 100 \text{g}^{-1}$

Note: The units of RMSEC and RMSEP are  $\text{mg} \cdot 100 \text{g}^{-1}$

表 3 样品 TVB-N 含量

Table 3 TVB-N content of samples in training set and prediction set

训练集样本数	训练集/ $(\text{mg} \cdot 100 \text{g}^{-1})$				预测集样本数	预测集/ $(\text{mg} \cdot 100 \text{g}^{-1})$			
	最小值	最大值	平均值	标准差		最小值	最大值	平均值	标准差
32	3.97	46.53	14.58	8.52	10	13.69	34.78	21.46	6.31

## 2.2 光谱数据预处理

原始光谱曲线不仅包含有用信息, 还包含噪声信号, 同时还可能存在诸如基线平移和漂移之类的问题, 所以对原始光谱进行 S-G 平滑、小波变换处理以及导数处理是很有必要的。在比较了经无预处理、SG 平滑、小波变换 (wavelet transform, WT)、一阶导数法 (first-derivative, Db1)、二阶导数法 (second-derivative, Db2) 以及不同方法组合等预处理方法以后, 采用 PLSR 法建立了河蟹新鲜度检测模型, 结果如表 4 所示。

由表 4 可知, 不同的光谱预处理方法对模型的建立有明显的影 响, WT, SG 平滑及 SG 平滑与 WT 相结合的处理效果较好, Db1 和 Db2 的处理结果较差, 结果甚至远劣于无预处理方法, 因而不适用于该模型的光谱预处理。小波变换处理与 SG 平滑、小波变换结合 SG 平滑处理相比, 模型决定系数更大, 均方根误差更小, 建模效果明显更好。而小波变换处理与无预处理建模相比, 预测集决定系数 0.850 7 略低于 0.863 6, 差值为 0.01, 但预测集均方根误差 3.379 3 明显小于 3.894 4, 差值为 0.52。决定系数本质是比较两列数据的

表 4 基于不同预处理的 PLSR 预测效果比较

Table 4 PLSR models based on different pretreatment methods

预处理方法	训练集结果		预测集结果	
	$R^2$	RMSEC	$R^2$	RMSEP
无	0.969 9	1.477 1	0.863 6	3.894 4
SG 平滑	0.944 9	1.998 4	0.834 7	3.464 1
WT	0.944 6	2.004 3	0.850 7	3.379 3
Db1	0.828 4	3.527 8	0.503 7	4.803 8
Db2	0.842 6	3.378 8	0.552 8	4.667 9
SG 平滑+WT	0.944 6	2.004 6	0.833 3	3.457 7
WT+SG 平滑	0.942 7	2.037 7	0.830 9	3.543 9

协方差与对应方差积的比重, 由于数据本身存在一定的误差, 所以在  $R^2$  相对接近的情况下, 采用 RMSEP 作为评价模型优劣的指标。因此, 本研究采用 WT 作为预处理的方法。

## 2.3 最优波段的确定

由表 4 可知, 经小波变换处理采用全波段光谱建立的模型效果较优, 模型精度可以满足在线检测的要求, 但在在线检

测不仅要求精度要高,而且还要求模型检测速度快,因此需要对模型进行优化。特征波段提取不但能够提高模型的运行速度,而且能够提高模型的稳定性。目前常用的特征波段提取方法有连续投影变换(successive projection algorithm, SPA)、主成分分析(principal component analysis, PCA)、遗传算法(genetic algorithm, GA)、无信息变量消除算法(uninformative variable elimination, UVE)等。PCA 通过筛选原数据集各组成向量(即原数据集矩阵的每个字段,一个字段代表一个特征,也可称为变量)协方差矩阵的累积方差,组合出将原数据集矩阵变换到主元空间的变换矩阵,原数据集矩阵与该变换矩阵的乘积即是原数据集各变量在主元空间的线性组合<sup>[13]</sup>,即利用原数据在主元空间上的映射来表示原数据集矩阵,由于在主元空间上可用更少的变量表示,从而实现了原数据集的降维。SPA 选择含有最少多余信息的波长变量组合以最小化信息重复叠加。本研究所包含的光谱波段较多,利用主成分分析法和连续投影算法可以对光谱特征波段进行有效降维,有效地提高模型运行的速度。

采用 PCA 和 SPA 分别提取经过相同预处理的光谱特征波段,然后对提取出的最优波段采用 PLSR 进行建模,结果如表 5 所示。

表 5 基于不同特征提取的 PLSR 效果比较

Table 5 PLSR models based on different feature extraction methods

特征提取方法	预处理方法	波段数	训练集结果		预测集结果	
			$R^2$	RMSEC	$R^2$	RMSEP
无	WT	1 050	0.944 6	2.004 3	0.850 7	3.379 3
PCA	WT	16	0.844 9	3.353 5	0.886 1	2.996 7
SPA	WT	14	0.768 4	4.098 0	0.827 5	3.099 9

由表 5 可知,与全波段建模相比较,经过 PCA 特征提取之后建立的模型,预测集决定系数更大,均方根误差更小,模型精度更高;在 SPA 特征提取之后,该模型具有更小的  $R^2$  和 RMSEP。全波段建模包含了更多无用的光谱信息,特征波段提取有效去除了干扰信息,使模型性能得到优化。本文选用 PCA 作为光谱特征提取的方法。

#### 预测模型的建立

目前常用的建模方法主要有偏最小二乘回归(partial least squares regression, PLSR)、主成分回归(principal component regression, PCR)、多元线性回归(multivariable linear regression, MLR)、人工神经网络(artificial neural network, ANN)等。MLR 可以通过线性拟合多个自变量和因变量之间的关系来确定模型的最优参数。本研究中涉及多次最优求解问题,因此本文采用 PLSR 与 MLR 进行建模。本文利用偏最小二乘回归和多元线性回归分别对经相同预处理、特征波段提取后的光谱进行建模,获得建立河蟹 TVB-N 定量模型

的最优算法。最终的建模结果如表 6 所示。

表 6 基于 PLSR 和 MLR 的预测模型效果比较

Table 6 Comparison of PLSR and MLR model performance

建模方法	预处理	特征提取	训练集结果		预测集结果	
			$R^2$	RMSEC	$R^2$	RMSEP
PLSR	WT	PCA	0.844 9	3.353 5	0.886 1	2.996 7
MLR	WT	PCA	0.850 4	4.810 0	0.838 9	3.364 5

从表 6 可以看出,通过 PLSR 获得的模型的预测决定系数  $R^2$  大于通过 MLR 建模获得的模型的预测决定系数  $R^2$ ,并且 RMSEP 小于 MLR 建模获得的模型的 RMSEP,PLSR 优于 MLR,因此本文最终选择 PLSR 作为建模方法。实验构建的最佳河蟹挥发性盐基氮模型如图 4 所示。

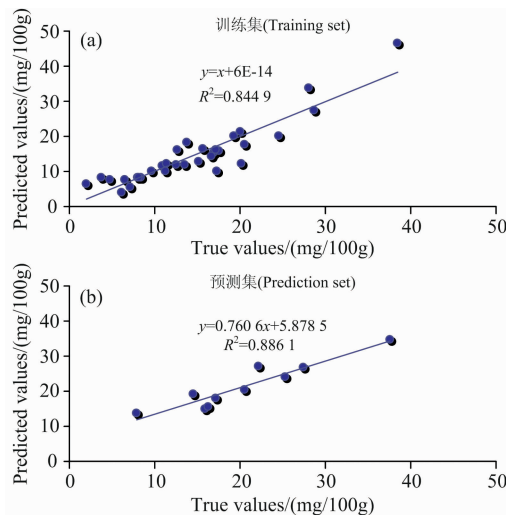


图 4 TVB-N 预测值与真实值散点图

Fig. 4 Scatter plots of TVB-N predicted value and true value

### 3 结论

利用 NIRS 构建了河蟹贮存过程中 TVB-N 预测模型。首先采用 5-fold CrossValidation 法对样本进行划分,其中的 32 个样品被划分为训练集进行模型构建,其余的 10 个样品被划分为测试集用于模型检验。然后利用 WT 对原始光谱进行预处理,消除了无用的信息并提高信噪比。再次,采用 PCA 法提取光谱最优波段,利用该最优波段作为模型的输入,不仅提高模型的运行速度,还可以提高模型的稳定性。最后采用 PLSR 法构建了 TVB-N 定量预测模型, $R^2$  和 RMSEP 分别达到 0.89 和 3.00。结果表明,预测模型具有较高的精度,该方法可以实现对河蟹新鲜度的快速检测,具有较好的市场应用前景。

## References

- [ 1 ] Gu S Q, Wang X C, Tao N P, et al. *Food Research International*, 2013, 54(1): 81.
- [ 2 ] China Fisheries Yearbook(中国渔业年鉴). Beijing: China Agriculture Press(北京: 中国农业出版社), 2015. 42.
- [ 3 ] ZHU Rong-guang, YAO Xue-dong, DUAN Hong-wei, et al(朱荣光, 姚雪东, 段宏伟, 等). *Spectroscopy and Spectral Analysis (光谱学与光谱分析)*, 2016, 36(3): 806.
- [ 4 ] Sivertsen A H, Kimiya T, Heia K. *Journal of Food Engineering*, 2011, 103(3): 317.
- [ 5 ] Cai Jianrong, Chen Quansheng, Wan Xinmin, et al. *Food Chemistry*, 2011, 126(3): 1354.
- [ 6 ] XU Fu-bin, HUANG Xing-yi, DING Ran, et al(徐富斌, 黄星奕, 丁然, 等). *Journal of Food Safety Quality(食品安全质量检测学报)*, 2012, 103(6): 644.
- [ 7 ] Kimiya T, Sivertsen A H, Heia K. *Journal of Food Engineering*, 2013, 116(3): 758.
- [ 8 ] Abdel-Nour N, Ngadi M, Prasher S, et al. *Food and Bioprocess Technology*, 2011, 4(5): 731.
- [ 9 ] Huang X, Ding R, Han F K, et al. *Analytical Methods*, 2014, 6(24): 9675.
- [ 10 ] Lin H, Zhao J, Sun L, et al. *Innovative Food Science and Emerging Technologies*, 2011, 12(2): 182.
- [ 11 ] Chuang Y K, Hu Y P, Yang I C, et al. *Journal of Cereal Science*, 2014, 60(1): 238.
- [ 12 ] GB 5009. 228—2016. National Standard of the People's Republic of China(中华人民共和国国家标准). National Standards for Food Safety Determination of Total Volatile Basic Nitrogen in Food(食品安全国家标准 食品中挥发性盐基氮的测定), 2016.
- [ 13 ] Yang H, Hassan S G, Wang L, et al. *Computers and Electronics in Agriculture*, 2017, 141: 96.

## Rapid Detection of Crab Freshness Based on Near Infrared Spectroscopy

LI Xin-xing<sup>1</sup>, YAO Jiu-bin<sup>1</sup>, CHENG Jian-hong<sup>2</sup>, SUN Long-qing<sup>1</sup>, CAO Xia-min<sup>3</sup>, ZHANG Xiao-shuan<sup>4\*</sup>

1. Beijing Laboratory of Food Quality and Safety, College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. Yantai Institute of China Agricultural University, Yantai 264670, China
3. School of Biology and Basic Medical Sciences, Soochow University, Suzhou 215200, China
4. College of Engineering, China Agricultural University, Beijing 100083, China

**Abstract** The freshness of river crab is the most important factor that most consumers consider when buying. Total volatile base nitrogen (TVB-N) is a commonly used international index for evaluating meat freshness, However, its detection process is cumbersome and time-consuming, which can not meet the urgent needs of the current market for rapid and objective evaluation of river crab freshness. Therefore, it is an urgent problem to establish a rapid method for detecting freshness of river crabs. A total of 126 crabs purchased from aquatic market were rapidly transported to the laboratory by polyethylene oxygenation bag. After treatment on a clean bench, the crabs were divided into 42 experimental samples, with 3 fresh crabs in each sample; 42 experimental samples were stored in a constant temperature biochemical incubator at low temperature of 4 °C. 6 crab samples were taken from the incubator on time every day for spectral data collection and freshness index determination for 7 days. In this paper, near infrared spectroscopy (NIRS) was used to evaluate the freshness of river crabs stored at different time, and total volatile base nitrogen (TVB-N) was used as an index to evaluate the freshness of crabs. Firstly, by comparing influence on the model prediction effect of 5-fold Cross Validation, Kennard-stone algorithm and sample set partitioning based on joint X-Y distance algorithm, finally, the 5-fold CrossValidation algorithm was used to divide the samples. 32 samples were used as training sets for model building, and the remaining 10 samples were used as test sets for model testing. Then, on the basis of dividing the samples by five fold cross validation algorithm, wavelet transform (WT), Savitzky-Golay smoothing, first derivative (Db1), second derivative (Db2) and wavelet transform (WT) combined with Savitzky-Golay smoothing were used to pretreat. Wave transform (WT) pretreatment was the best spectral pretreatment method, which eliminated the useless information in the spectrum and improved the signal-to-noise ratio. Once more, on the basis of the WT pretreatment, principal component analysis (PCA) and successive projection algorithm (SPA) were used to extract spectral feature bands, and the principal component analysis (PCA) was used as the optimal wavelength selection method by comparing the model prediction effect. With the selected 16 feature bands as the input of the model, which not only improved the running speed of the model, but also improve the stability of the model. Finally, after PCA feature extraction, by using partial least squares regression (PLSR) and multiple linear regres-

sion (MLR) built the TVB-N quantitative prediction model, by comparing the two kinds of model prediction effect to determined the partial least squares regression (PLSR) model for the optimal modeling method, this paper finally determine the optimal model based on WT-PCA-PLSR model, model prediction determination coefficient  $R^2$  was 0.89, and the root mean square error of prediction RMSEP was 3.00. In conclusion, the prediction model established in this study has a high accuracy, and this method can realize the rapid detection of the freshness of river crabs, and has a good market application prospect.

**Keywords** Near infrared spectroscopy; Total volatile base nitrogen; Freshness; Wavelet transform; Principal component analysis; Partial least squares regression

(Received Nov. 9, 2018; accepted Mar. 24, 2019)

\* Corresponding author