

化学计量学在食品风味领域的应用进展

张 牵¹, 韩颖颖¹, 孟繁宇¹, 王亚东¹, 王 蓓^{1,*}, 江 滔^{2,*}

(1.北京工商大学食品与健康学院, 北京 100048; 2.法国勃艮第大学里昂神经科学研究中心, 法国 布龙 69500)

摘 要: 风味物质影响食品的感官属性及消费者选择, 风味物质分析对食品质量改善以及新产品开发都至关重要。然而, 风味物质数据庞杂、统计分析运用不恰当极大程度限制了该领域发展, 因此, 正确、合理地使用新型化学计量学方法, 如人工智能算法获取有效信息, 在该领域至关重要。近年来化学计量学方法广泛应用于食品研究中, 除降维、分类与回归方法外, 多种神经网络方法也在食品领域研究中兴起, 但对其合理应用的归纳总结较少。因此, 本文对可用于研究食品风味的统计分析方法进行归纳总结, 包括主成分分析和线性判别分析, 线性回归方法如偏最小二乘回归系列, 及非线性方法如模糊逻辑、人工神经网络等多种神经网络方法, 对其原理进行讲解并进行应用举例, 旨在为化学计量学在后续食品风味领域的研究提供有效的方法和思路。

关键词: 化学计量学; 食品风味科学; 感官分析; 统计分析

Progress in the Application of Chemometrics in the Field of Food Flavor

ZHANG Qian¹, HAN Haoying¹, MENG Fanyu¹, WANG Yadong¹, WANG Bei^{1,*}, JIANG Tao^{2,*}

(1. School of Food and Health, Beijing Technology and Business University, Beijing 100048, China;

2. Lyon Neuroscience Research Centre, University of Burgundy, Bron 69500, France)

Abstract: Flavor substances affect the sensory properties of food and consumer choices, and flavor substance analysis is crucial for improving food quality and developing new products. However, the vast amount of flavor substance data and inappropriate statistical analysis greatly limit the development of this field. Therefore, it is crucial to use new chemometrics methods, such as artificial intelligence algorithms, correctly and reasonably to obtain effective information in this field. In recent years, chemometrics methods have been widely applied in food research. In addition to dimensionality reduction, classification and regression methods, various neural network methods have also emerged in the field of food research. However, a summary of their reasonable application is lacking. Therefore, this article summarizes the statistical analysis methods available to study food flavor, including principal component analysis, linear discriminant analysis, linear regression methods such as partial least squares regression, and nonlinear methods such as fuzzy logic and artificial neural networks, explains their principles and provides application examples. This article aims to provide effective methods and ideas for further research on chemometrics in the field of food flavor.

Keywords: chemometrics; food flavor science; sensory analysis; statistical analysis

DOI:10.7506/spkx1002-6630-20240328-214

中图分类号: TS201

文献标志码: A

文章编号: 1002-6630 (2024) 21-0307-09

引文格式:

张牵, 韩颖颖, 孟繁宇, 等. 化学计量学在食品风味领域的应用进展[J]. 食品科学, 2024, 45(21): 307-315. DOI:10.7506/spkx1002-6630-20240328-214. <http://www.spkx.net.cn>

ZHANG Qian, HAN Haoying, MENG Fanyu, et al. Progress in the application of chemometrics in the field of food flavor[J]. Food Science, 2024, 45(21): 307-315. (in Chinese with English abstract) DOI:10.7506/spkx1002-6630-20240328-214. <http://www.spkx.net.cn>

随着计算机技术的进步, 化学计量学在实验研究中的作用越来越受到关注。化学计量学是指使用统计学和

人工智能方法, 根据大量的分析数据对样品进行表征和分类。作为一个跨学科领域, 化学计量学使用数学统

收稿日期: 2024-03-28

基金项目: 国家自然科学基金面上项目 (32072345)

第一作者简介: 张牵 (2000—) (ORCID: 0009-0005-2866-2561), 女, 硕士研究生, 研究方向为发酵乳制品风味。

E-mail: 1317187258@qq.com

*通信作者简介: 王蓓 (1981—) (ORCID: 0000-0001-5407-7810), 女, 教授, 博士, 研究方向为乳制品分子感官组学及智能感官大数据分析。E-mail: wangbei@th.btbu.edu.cn

江滔 (1961—) (ORCID: 0000-0001-7167-9987), 女, 研究员, 博士, 研究方向为感觉系统的心理物理、心理生理和神经生理及感官品评。E-mail: tao.jiang@univ-lyon1.fr

计方法设计或选择最佳的测量程序和实验，并通过分析数据提供最大的有效信息^[1]，因此化学计量学的使用是食品领域研究中非常重要的一环。在食品研究中，最常见的应用是识别样本的类别、了解产品质量属性和研究样品分析中未涵盖的属性^[2]。

食品的风味包含了风味物质所导致的味觉、嗅觉以及化学物质刺激产生的综合感知^[3]，风味感知依赖于不同感官形式的多种组合，如物理感觉（形状、质构）和化学感觉（气味、滋味）^[4]，是不同食品之间相互区别的重要特征指标，是影响消费者购买意向的重要感官属性^[5]。根据不同的评价目的，食品风味物质常用来了解产品的差异性、感官属性的分类、数据的建模预测等，需要通过合适的统计方法将风味数据与其他类型数据结合分析。风味物质分析主要涉及消费者品尝的食品滋味和气味以及相关的风味化合物的测定，旨在检测食品中可能直接或间接地影响食品风味的挥发性和非挥发性化合物^[3]。例如在研究食品风味与感官属性关系时，通过感官描述，可在产品的物理化学属性和感官属性之间的相关性还没有充分建立之前，用于快速了解产品感官属性^[6]。将复杂的风味物质与带有主观性的感官属性进行相关性联系研究，需要通过实验设计，然后从这些实验中获得有意义的信息^[7]。所以统计分析是从所测量事件中获得结果的重要的一步，可以克服主观性，更加客观地获取数据里所包含的信息，总结归纳各个变量或因素间隐藏的关系，为判断、做决定提供重要的现实依据。

在食品领域的统计分析中，常使用单因素方差分析（analysis of variance, ANOVA）比较统计总体，以确定它们之间是否存在统计学显著差异。在化学计量学中，ANOVA用于研究自变量对因变量的影响^[2]。然而使用单变量方法如ANOVA有时不足以发现食品风味物质、感官特征和偏好程度特性之间的关系模式，这是由于食品中存在大量化合物和大量物理化学属性^[8]，由于掩蔽和协同效应的存在，即同时进行两种及两种以上刺激，或刺激之间产生压制或叠加，从而使得消费者对某一刺激的感觉强度低于或超过原有刺激水平，这一现象尤其存在于食物的味道和气味特性之间。食品的属性繁多且相互关联，共同构成了一个高度复杂且多维度的关系网络。这种复杂性超越了传统单变量分析方法的处理能力，使得其无法全面捕捉食品内部的复杂结构和潜在关联。多变量统计方法因其能够同时考虑多个变量之间的相互作用和关系，成为了分析食品复杂性的关键工具。该方法不仅能够深入剖析食品内部的问题，还能够揭示出变量间的潜在联系和模式，为食品科学的研究提供了有力的支撑。因此对于食品复杂属性的分析，需要借助更为全面且先进的多变量统计方法。

在食品科学领域，风味物质的研究对于提升产品品质、满足消费者需求具有重要意义。随着科研技术的发展，降维、分类和回归分析方法在风味物质实验数据处理中得到了广泛应用。同时，近年来神经网络方法在食品科学领域展现出了显著的应用潜力，为风味物质的研究提供了新的视角和方法。然而，目前针对这些方法及其在食品领域应用的系统性综述尚显不足。鉴于此，本

文旨在综述食品领域中化学计量学方法的应用，特别是那些依据风味物质进行消费者喜好预测、质量控制和产品分级分类的统计分析方法。这些统计分析方法包括主成分分析（principal component analysis, PCA）、线性判别分析（linear discriminant analysis, LDA），以及线性回归分析方法如偏最小二乘回归（partial least squares regression, PLSR）和非线性分析方法等。利用这些多变量分析的研究示例总结见表1。

表 1 统计方法在食品领域研究中的应用
Table 1 Application of statistical methods in food field research

产品	目的	统计方法	方法的具体应用
鲈鱼 ^[9]	对鲈鱼中挥发性有机物进行定性定量分析，确定生物标志物	PCA	对鲈鱼中挥发性有机物进行PCA
初榨橄榄油 ^[10]	根据风味特征区分不同类别的初榨橄榄油	PCA、LDA	PCA应用于自然聚类的样品，LDA用于创建橄榄油分类的预测模型
花生油 ^[11]	鉴别不同风味花生油及其掺假	PCA、LDA	PCA用于区分3种风味类型的8种花生油，LDA用于区分花生油中掺杂的植物油
茴香精油 ^[12]	根据挥发性风味物质确定不同提取方法制成的茴香精油	PCA、LDA	对电子鼻数据同时进行两种统计方法的分析对比
葡萄酒 ^[13]	探究感官属性与挥发性组分之间的关系	PCA、PLSR	PCA用于显示基于化学和感官特征的葡萄酒的分布，PLSR研究葡萄酒感官指标与挥发性成分之间的关系
蛋糕模型 ^[14]	材料和时间在烘焙中对于感官属性和消费者对蛋糕新鲜度感知的影响	PCA、PLSR	PCA用于对不同原料和不同烘烤时间的蛋糕进行感官分析，PLSR用于预测所有蛋糕的消费者群的新鲜度
熟羊肉丸 ^[15]	分析熟制羊肉丸在贮藏过程中的风味变化	PLSR	综合分析关键香味活性物质、感官评价和贮藏时间之间的相关性
UHT牛奶 ^[16]	对14种UHT乳制品的感官特征和消费者享乐感知进行研究	PLSR	验证了感官属性与总体喜好之间的相关性
切达奶酪 ^[17]	了解天然切达干酪的风味和香气特征，并确定关键的非挥发性化合物	PLSR	解析了关键非挥发性化合物及其对味觉和香气特征的贡献
豆浆 ^[18]	探讨3种市售豆浆与2种实验室自制豆浆感官特性与消费者接受度	PCA、PLSR	PCA用于分析描述性属性与样品间关系，PLSR用于调查样品的描述性属性和消费者可接受性间关系
橄榄油 ^[19]	分析橄榄油的红外光谱以确定4种主要的感官缺陷	PLS-DA	PLS-DA用于开发分类模型，以区分特级初榨橄榄油和低质量橄榄油
木瓜泥 ^[20]	旨在对4个不同品种的果泥的挥发性成分和感官特性进行表征	PLS-DA	基于不同的挥发物区分果泥
咖啡 ^[21]	研究烘焙咖啡的挥发性化合物与某些感官属性之间的关系	PLS-DA	用PLS-DA对不同质量样品进行判别，解释咖啡化化合物与感官属性的关系
冰激凌 ^[22]	了解感官和仪器风味谱之间的关系	PLSR	解析化合物和感官数据之间的相关性
铁皮石斛花饮料 ^[23]	分析铁皮石斛花饮料的感官评价，并根据其感官特征对不同成分进行排名，以获得最佳饮料	模糊逻辑	根据样品感官特征对不同样品成分进行排名
咖啡 ^[24]	通过感官评价对所选咖啡产品进行排名，并使用模糊决策找出最佳的质量属性	模糊逻辑	对所选择的咖啡产品进行排名
冰激凌 ^[25]	预测冰激凌的总接受度	ANN	用ANN模型进行预测，预测冰激凌在消费者中的总体接受程度
佩科里诺奶酪 ^[26]	解释奶酪挥发性化合物研究中的电子鼻和气相色谱数据	ANN	用ANN模型预测区分不同奶酪类型
冬虫夏草 ^[27]	对野生及人工栽培冬虫夏草进行质量控制和重要标记物筛选，同时进行分类	OPLS-DA	作为训练样本建立原始模型，筛选出重要标记物后同时对供试样品进行分类
牛肉 ^[28]	建立一种可靠有效的特异性肽筛选和牛肉制品真实性鉴定方法	OPLS-DA	用OPLS-DA在短时间内获得所需的特异性肽信息的样本
绿茶 ^[29]	使用模糊逻辑评估不同茶叶样品的感官评分，并根据其感官品质对样品进行分级	模糊逻辑	对茶叶样品评分进行数据分析并排序分级

注：PLS-DA.偏最小二乘判别分析（partial least squares-discriminant analysis）；ANN.人工神经网络（artificial neural networks）；OPLS-DA.正交偏最小二乘判别分析（orthogonal partial least squares-discriminant analysis）；UHT.超高温（ultra high temperature）。

1 多变量线性数据分析方法

降维是一种对高维度数据的预处理方法。在多样本多变量的数据中,通过降维将高维度的数据保留下最重要的一些特征,去除不重要的特征,同时利用算法保留易被忽略的重要特征,从而实现提升数据处理速度的目的。但所有多变量方法都不是旨在降维。多变量方法本质上是对原来的变量空间进行移动及旋转,造就一个新的使要素互相垂直的空间。所以PCA和LDA这类多变量分析的原理更多是改变角度观察同样的数据集,使每次看到的信息可以比用原始变量角度观察到的信息更多。在这里介绍PCA和LDA这两种算法。

1.1 PCA

PCA是一种多维数据的投影方法。它是一种线性无监督的识别技术,用于对多元问题中的数据集合进行变量的相关性及个体分布的概括^[30]。无监督方法不需要训练集,它们根据样本之间的相似性对样本进行分组;新样本将被放置在现有组中,或者可能会创建一个新组。在没有事先分类信息的情况下,这些方法适用于初步的、探索性的和定性的分析。它使用正交变换将一组可能相关的变量即不正交的原始变量转换为一组不相关变量的数据^[11]。PCA将原始变量进行不同的线性组合,构造出一组新的互不相关的变量并称为主成分(principal component, PC),这些PC是正交的,每个PC是原始测量变量的线性组合。这种技术提供了一组正交轴,表示数据中最大方差的方向。PC1占总方差的最大值,PC2与PC1不相关,占残差方差的最大值,以此类推,直到占满总方差^[31]。最上面的2个或3个PC构建了一个二维或三维坐标,二维或三维图显示了观察结果之间的关系^[30]。图1为不同成熟时间奶酪的PCA得分图示例,横坐标41.7%为PC1占总方差的百分比,纵坐标34.2%为PC2占总方差的百分比。 u_1 、 u_2 表示原始变量,箭头方向代表原始变量与PC的相关性,箭头直线长度代表原始数据对PC的贡献度。

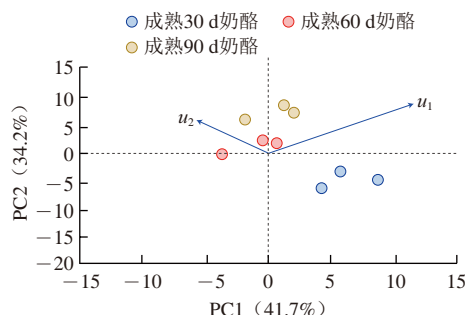


图1 不同成熟时间奶酪样品PCA得分图

Fig. 1 PCA score plot of cheese samples with different ripening periods

Heenan等^[14]在烘焙模型蛋糕系统的研究中,探讨了原料构成与烘焙时长对蛋糕感官质量和消费者新鲜度认

知的具体影响。通过运用PCA和相关负荷图从总计24种蛋糕类型中筛选出了12种在感官变化上表现最为显著的蛋糕配方。并通过PCA进一步明晰了不同甜味剂、脂肪类型及烘焙时间之间的差异与联系。Gonzalez Viejo等^[32]使用人工智能(artificial intelligence, AI)结合化学计量学PCA对啤酒香气和质量性状进行研究。通过PCA研究啤酒挥发性香气成分与泡沫等相关参数的关系,并使用AI模型对啤酒香气进行预测,发现AI能对商业啤酒进行很好地评估。

PCA可以清晰地展示了数据组内的重复性和组间的差异性,评估数据的可重复性,发现可能存在的异常值,比如明显离群的样本点。PCA是最常见的基于特征分解的分析方法,属于探索性分析。但由于数据的复杂性,对于组间差异不够明显的样品,当单纯的无监督分析不能很好地区分样本的组间差异时,就需要用到其他的分析方法。

1.2 LDA

判别分析功能是将样品归类,判别函数是描述符的线性组合^[33]。经典LDA中的最佳变换是通过同时最小化类内距离和最大化类间距离获得的,从而实现最大的类区分度^[34]。LDA是一种有监督的数据分类技术^[11],不仅可以进行数据降维,还可以用于分类。作为有监督的算法,因变量的分类是已知的,输出的数据集的每个样本是有类别的,这点和PCA不同。图2为不同成熟时间奶酪样品LDA得分图示例,从横坐标方向上能看出,相比于PCA,有监督的LDA良好地分离了不同组别。

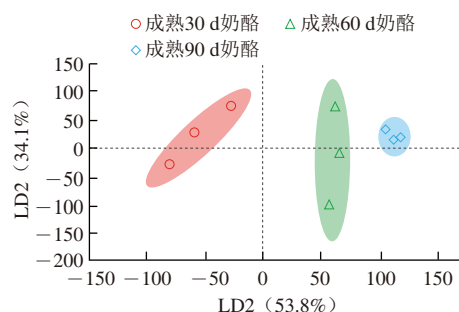


图2 不同成熟时间奶酪样品LDA得分图

Fig. 2 LDA score plot of cheese samples with different ripening periods

LDA的监督方法需要一组初始的分类或标记数据来调整模型的参数,然后将其用作分类器,即建立分类的标准。训练后的分类器可以对现有类别中的新样本进行分类。例如叶美霞等^[35]在LDA中将不同年龄段作为应变量,将不同年龄段乳粉作为已知输出类别,即为有监督分析,分类研究乳粉的挥发性风味物质,并对乳粉中香气特征进行分类,其中婴儿乳粉全部被正确区分,证明LDA可通过风味特征,将同一年龄段不同品牌的乳粉明

索非挥发性代谢物如何协同影响奶酪风味,并确定奶酪风味中关键成分,采用PLSR结合VIP值将感官属性特征与这些非挥发性代谢物关联起来,发现关键非挥发性的成分,并使用PLSR分析味觉特征和味觉活性非挥发性代谢物之间的关系。

2.2 PLS-DA

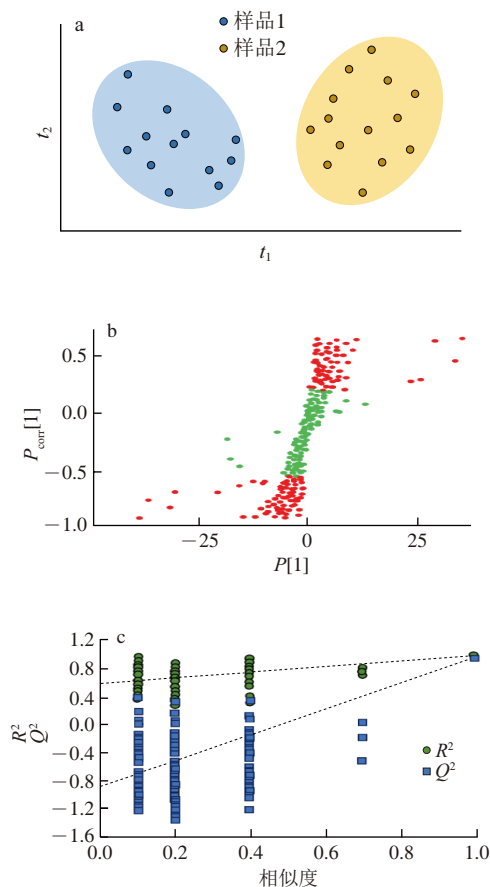
PLS-DA中PLS是多变量分析中的经典方法,是用来判断研究数据如何分类的常用统计分析方法^[46]。PLS-DA是目前应用最广泛的判别分析方法,它是PLSR与分类技术(LDA)的结合^[47]。DA是判别分析,是一种根据观察到的若干变量值,判断样品如何分类的方法。PLS-DA用PLSR的方法,建立了回归模型,并对回归结果进行类别区分分析。PLS-DA得分图外观与PLSR相近(图3),与PLSR相比,当Y即响应值是分类别数据时的技术称为“偏最小二乘判别分析”。可以用PLS-DA建立食品风味物质与样品类别之间的关系模型,实现对样品类别的预测^[48]。Lieb等^[20]对4种木瓜泥的挥发性物质进行分析,并根据不同挥发物来进行木瓜区分,使用PLS-DA将不同木瓜果泥进行了有效区分,PLS-DA的前两个潜在变量(LV1和LV2)解释了Y变量中58.49%的方差,证明使用PLS-DA大大提高了不同木瓜果泥之间的区分度。Borràs等^[19]使用化学计量学方法分析橄榄油的中红外光谱以确定4种主要的感官缺陷。使用PLS-DA开发分类模型,以区分特级初榨橄榄油(不存在缺陷)和低质量橄榄油(存在缺陷)。结果表明PLS-DA模型能够区分有味道缺陷的油和高品质的油,显示出约87%的霉味缺陷和约77%的酒味、发霉和腐臭缺陷的预测能力。

2.3 OPLS-DA

OPLS-DA是一种相对较新的多元统计方法,是在PLS-DA的基础上开发的一种对信息进行分类的算法,近十年来,这种方法在理论和应用方面得到了迅速的发展,并在计量化学中有大量的应用。OPLS是一种多因变量对多自变量的回归建模方法,其大的特点是可以去除自变量X中与分类变量Y无关的数据变异,进而寻找该主成分的正交矫正轴方向,从而使得组间样本分离效果更佳,组间差异最大化凸显,且更适用于两组样本间的分离。使分类信息主要集中在一个主成分中,从而使模型变得简单和易于解释,其判别效果及主成分得分图的可视化效果更加明显。

OPLS-DA结果的评价参数是 R^2 和 Q^2 。 Q^2 是指建模后模型的预测能力, R^2 代表解释能力, R^2 越接近于1,说明应变量对因变量的解释能力越强,回归模型拟合效果越好^[27]。 R^2_x 是X变量中的变化的解释能力, R^2_y 是由预测和正交分量组成的Y变量中的变化的解释能力^[49]。OPLS-DA通常由得分图、S-plot图、置换检验图组成。图4为OPLS-DA结果常见的3种图,图4a为得分图,得分图的

横坐标表示过程中的PC得分值,在横向上可以看出组间差异;纵坐标表示正交成分的得分值,从纵向上可以看出组内差异。图4b为S-plot图,横坐标表示PC与变量的协相关系数,纵坐标表示PC与变量的相关系数,越靠近两角表示其差异越显著,红色的点代表VIP值大于等于1,绿色的点表示VIP值小于1。图4c为置换检验图,有监督的分类模型缺点是可能会出现过拟合现象,即模型可以很好地将样本进行区分,但用来预测新的样本集时却表现很差。因此对于有监督的分类模型,需要验证模型的可靠性。其验证原理是将每个样本的分组标记随机打乱,再来建模和预测。图4c得到 R^2 为0.575, Q^2 为-0.897,虚线为回归线,通常有监督模型要求 Q^2 回归线在Y轴上的截距小于0。原始值(右点)高于左侧置换点时被认为有效,这表明OPLS-DA具有良好的模型拟合度和可预测性。



a. OPLS-DA得分散点图; b. S-plot图; c. 置换检验图。

图4 OPLS-DA常见的3种图

Fig. 4 Three common OPLS-DA plots

OPLS-DA是在PLS-DA的基础上,进行了正交变换的矫正,可以滤除与分类信息无关的噪音,提高了模型的解析能力和有效性^[50]。与PLS相比,OPLS具有相似的预测能力,并改进了对预测成分和系统变异的解释。当

变量数量远大于样品数量时（行数小于列数），PLS或PLS-DA模型容易过拟合，但是加入正交矫正之后数据会更准确^[51]。Wang Jing等^[27]通过研究旨在对野生及人工栽培冬虫夏草进行质量控制和重要标记物筛选，采用多个野生冬虫夏草和人工养殖冬虫夏草作为训练样本建立OPLS-DA原始模型。筛选出重要标记物后对供试试验样品进行分类，分类准确率达到95.5%。近年来的研究中多将化学计量学与机器学习相结合来进一步对食品进行分析。Lu Kuan等^[52]通过研究确定虾酱的来源，剖析虾酱的香气本质。使用OPLS-D和PCA研究挥发性化合物、感官属性和发酵指标之间的相关性，并结合机器学习（ANN、支持向量机（support vector machines, SVM）等）方法建立了虾酱产地识别的分类模型。

关于以上介绍的3种线性回归方法，表3总结了其分别的逻辑原理和特点，可用以进一步区分和了解3种方法。

表3 PLSR、PLS-DA及OPLS-DA的逻辑和特点

名称	逻辑原理	特点	参考文献
PLSR	同时基于目标和数据矩阵的线性回归方法，既考虑了尽量提取Y和X中的PC，又考虑了使从X和Y中提取出的PC之间的相关性最大化	能在自变量之间也存在线性关系的情况下对数据进行建模，也可以应用于样本数远远小于自变量个数的情况	[41,53]
PLS-DA	对不同处理样本（如观测样本、对照样本）的特性分别进行训练并产生训练集，然后再对样品进行预测。当变量数量远大于样品数量时，PLS-DA模型容易过拟合	按照预先定义的分类变量，最大化组间的差异，适合两组以上的组别分类比较，可能出现过度拟合数据	[46-48]
OPLS-DA	使用正交信号校正技术，将X矩阵信息分解成与Y相关和不相关的两类信息，然后过滤掉与分类无关的信息，相关的信息主要集中在第一个预测成分，有效减少模型的复杂性和增强模型的解释能力	加入正交矫正之后可以更好地避免过拟合现象，且更适用于两组样本间的分离	[27,49,53]

3 非线性分析

线性回归模型常用于关联食品系统中的不同数据集。然而，食物中存在的信息本质上可能是非线性的^[54]，需要数据转换或其他技术将非线性关系转换为可用线性技术建模的关系。然而，转换后的变量或模型系数可能难以解释。非线性方法的使用允许将不同的数据集与非线性关系相关联，而不需要进行数据转换。在大多数情况下，例如在食品的化学成分和感官特性中，可能没有预先确定的非线性模型^[40]。如模糊逻辑和ANN等神经网络方法可能会在这类例子中使用。

3.1 模糊逻辑

模糊逻辑是一种用于决策和分类的工具，它是模拟人类思维过程之后通过不精确信息生成复杂决策来建模的^[40]。简单地说，模糊逻辑是指基于概率的逻辑，其中在数据中缺乏确定或绝对值，其中值是假（由0表示）或真（由1表示）。采用了一种计算方法，取消二值之间非

此即彼的对立，用隶属度表示二值间的过度状态。隶属度的值通常介于0和1之间，其中0表示不匹配或完全不相似，1表示完全匹配或完全相似。考虑到这一点，基于概率的模糊逻辑允许值在0和1之间，允许数据是“部分真或假”。绝对值的缺乏是感官评价中的一个重要特点。例如，果汁可能不一定是绝对甜的或酸的，而是可以是部分甜的，带有一丝酸味。由于感官结果通常高度可变，因此使用统计方法的后续分析可能产生具有低准确度、精确度以及不可重复性的结果^[55]。应用模糊数学方法可以更系统、更准确地评价食品的感官质量。模糊逻辑是分析感官数据中经常遇到的处理模糊数据和模糊关系的重要工具，并且可以用于从食物的感官和偏好性特性中得出结论^[29]。图5为模糊逻辑控制器原理图，通过控制器对输入数据进行模糊处理。

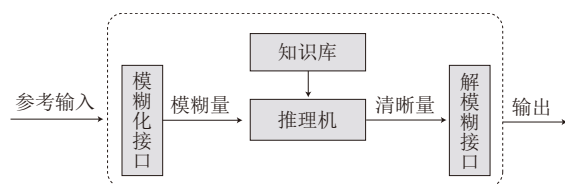


图5 模糊逻辑控制器功能模块

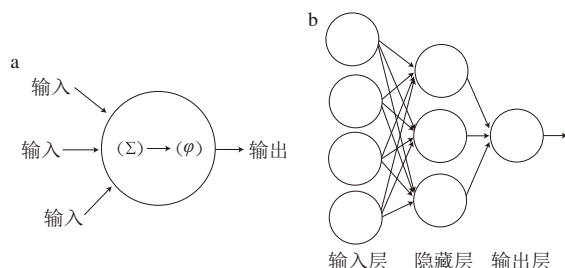
Fig. 5 Functional module of fuzzy logic controller

目前模糊逻辑已被用于感官分析，以得出有关消费者接受度、食品味道排名的结论，以及识别重要的差别因素。在模糊数学中，把一系列需要研究的现象作为考察对象，把一些相互关联的因素称为集合。用模糊数学评价食品的感官质量，食品质量由此可以分级，这就解决了食品产品分级的难题^[23]。模糊逻辑可以被用于分析其他食品和应用中的感官数据，例如使用模糊逻辑对不同茶叶样品的感官评分数据进行分析^[29]，并用于样品排序，找出了一般的绿茶饮料的质量属性的偏好，并根据其感官品质对样品进行分级。Lazim等^[24]也对3种不同咖啡饮料的感官评分使用模糊逻辑，通过比较判断隶属度函数和权重对产品进行排序，最终确定了3种咖啡的排名，并确定最佳质量属性，即味道是最好的质量属性，这一质量属性也反过来证实了感官评价的重要性。

3.2 ANN

ANN是一种大规模并行分布式信息处理系统，具有类似于人脑生物神经网络的某些性能特征。ANN已经发展成为人类认知和神经生物学数学模型的推广^[56]。作为一种非线性方法，其目的是模拟大脑和生物神经系统进行的模式识别和信息存储过程，单个节点模仿生物神经元，加权连接模仿轴突和突触。这可以与模糊逻辑形成对比，模糊逻辑也试图模拟复杂的人类思维过程^[40]。图6a、b分别为单个和多层ANN图示，由连接的神经单元（也称为节点）组成，通常分为输入层、输出层和隐藏

层。输入层往往对应于自变量,输出层生成神经网络的输出,在大多数情况下,它是回归中因变量的预测值,或分类问题中的分类组。隐藏层是介于输入和输出层之间的一层节点,与输入和输出数据没有直接关系。这一层是神经网络的主力,也是ANN解决复杂非线性问题能力的主要驱动力。



a. 单个人工神经单元的图示, 其中求和和变换的主要过程分别由 Σ 和 ϕ 表示; b. 具有4-3-1结构的3层ANN, 每层由人工神经单元组成。

图6 人工神经单元和ANN原理图

Fig. 6 Schematic diagram of artificial neural unit and ANN

ANN已被用于食品领域, 是一种新兴的方法, 用于食品分析中非线性计算的建模, 如食品安全和质量分析、食品图像分析, 以及各种热和非热食品加工操作的建模。Bhagya Raj等^[57]结合ANN在非线性食品工程问题建模中的各种应用, 综述了ANN在乳品加工过程建模中的应用。Bahramparvar等^[25]评估了ANN在预测消费者对由3种不同稳定剂制成的冰淇淋的接受度方面的用途。消费者接受度是基于由训练有素的小组确定的6个感官属性来预测的。结果表明, 风味和质地是用来预测冰激凌总接受度的最敏感的感官属性, 同时ANN模型可以潜在地用于估计冰淇淋的总感官接受程度。ANN作为非线性数字信号与感官特性或享乐特性相关联的工具, 也已用于与电子鼻和电子舌一起分析食品的气味和味道活性化合物^[26]。Cevoli等^[26]用ANN解释电子鼻处理后的奶酪中挥发性化合物数据, 结果证明所有的ANN模型(不同的预处理数据)有预测不同奶酪类别的能力。

3.3 其他神经网络方法

除上述两种方法外, 常用于食品风味领域的还有神经网络中的SVM、自组织映射(self-organizing maps, SOM)。SVM是最大区间分类器, 它使经验误差最小化, 边缘面积最大化。它主要用于解决线性不可分问题, 数据通常处于非线性和不可分离状态^[58]。Liu Ming等^[59]对白酒的不同香型进行分类, 发现使用SVM分类的准确率高于LDA, 且测试准确率几乎匹配所有白酒类型, 说明SVM有着更强的泛化能力, 准确地管理分类任务。SOM中没有隐藏层, 这同时允许非线性建模和降维。High等^[60]研究不同成熟时间的蓝纹奶酪, 识别挥发性风味化合物。他们研究了SOM和基于熵的特

征选择(entropy-based feature selection, EFS)模型对奶酪的分类, 发现SOM-EFS方法的预测性能优于PLS-DA模型。现多将机器学习与化学计量学结合对食品模型进行预测。机器学习中的预测模型是使用大量已知标记数据(训练数据)构建的, 在训练后可以使用该模型来预测未知和未标记数据。Shen Che等^[61]将PCA、OPLS-DA与卷积神经网络-支持向量机(convolutional neural network-support vector machine, CNN-SVM)结合, 因为CNN提取更抽象和不变的数据特征, 并且比浅层分类器表现更好, 因此模型比普通SVM更有效。文章研究并预测了羊肉串中的挥发性化合物。此外, 该模型对不同烘烤方式羊肉串的风味特征识别效果较好, 能够更清晰地区分羊肉串的不同烘烤方式。虽然神经网络被认为是处理复杂的问题与非线性数据集的较好的工具, 但因为神经网络模型看起来像一个“黑匣子”, 无法确切地看到模型计算过程中发生了什么, 因此, 提供误导性输出的可能性很大^[62], 不像线性回归方法这样容易实现, 因此上述非线性的方法在食品领域应用还较少。

4 结 语

PCA、LDA以及PLS系列等线性回归分析方法在食品风味及感官领域已经广泛应用, 本文对其应用进行了示例讲解与总结, 但这些技术可能不足以完全描述食品中发现的复杂且潜在的非线性系统。部分非线性分析方法, 如神经网络系列的方法, 随着计算机技术的不断发展, 也逐渐被发现并更多地应用在食品领域。本文对多种非线性方法进行了应用介绍, 但目前非线性方法还存在未知领域, 在应用中不像线性方法容易实现。本文讨论的统计学方法的基本原理和应用可能能够为实验者提供更深入有效的信息。最新的统计学数据处理方法仍在不断开发中, 为了使其在食品研究领域更容易建立联系并被接受, 仍需进一步探索。

参考文献:

- [1] BYSTRZANOWSKA M, TOBISZEWSKI M. Chemometrics for selection, prediction, and classification of sustainable solutions for green chemistry: a review[J]. Symmetry, 2020, 12(12): 2055. DOI:10.3390/sym12122055.
- [2] TARAPOULOUZI M, MIRONESCU M, DROUZA C, et al. Insight into the recent application of chemometrics in quality analysis and characterization of bee honey during processing and storage[J]. Foods, 2023, 12(3): 473. DOI:10.3390/foods12030473.
- [3] ANDREWES P, BULLOCK S, TURNBULL R, et al. Chemical instrumental analysis versus human evaluation to measure sensory properties of dairy products: what is fit for purpose?[J]. International Dairy Journal, 2021, 121: 105098. DOI:10.1016/j.idairyj.2021.105098.

- [4] 田怀香, 郑国茂, 于海燕, 等. 气味与滋味间相互作用对食品风味感知影响研究进展[J]. 食品科学, 2023, 44(9): 259-269. DOI:10.7506/spkx1002-6630-20220515-194.
- [5] 王铁龙, 许凌云, 杨冠山, 等. 智能感官分析技术在食品风味中的研究进展[J]. 食品安全质量检测学报, 2023, 14(8): 37-43.
- [6] JOHNSON A J, HEYMANN H, EBELER S E. Volatile and sensory profiling of cocktail bitters[J]. Food Chemistry, 2015, 179: 343-354. DOI:10.1016/j.foodchem.2015.01.114.
- [7] PERROT N, IOANNOU I, ALLAIS I, et al. Fuzzy concepts applied to food product quality control: a review[J]. Fuzzy Sets and Systems, 2006, 157(9): 1145-1154. DOI:10.1016/j.fss.2005.12.013.
- [8] ZIELINSKI A A F, HAMINIUK C W I, NUNES C A, et al. Chemical composition, sensory properties, provenance, and bioactivity of fruit juices as assessed by chemometrics: a critical review and guideline[J]. Comprehensive Reviews in Food Science and Food Safety, 2014, 13(3): 300-316. DOI:10.1111/1541-4337.12060.
- [9] NIE S, LI L H, WANG Y Q, et al. Discrimination and characterization of volatile organic compound fingerprints during sea bass (*Lateolabrax japonicus*) fermentation by combining GC-IMS and GC-MS[J]. Food Bioscience, 2022, 50: 102048. DOI:10.1016/j.fbio.2022.102048.
- [10] GERHARDT N, SCHWOLOW S, ROHN S, et al. Quality assessment of olive oils based on temperature-ramped HS-GC-IMS and sensory evaluation: comparison of different processing approaches by LDA, kNN, and SVM[J]. Food Chemistry, 2019, 278: 720-728. DOI:10.1016/j.foodchem.2018.11.095.
- [11] HUYAN Z Y, DING S X, LIU X L, et al. Authentication and adulteration detection of peanut oils of three flavor types using synchronous fluorescence spectroscopy[J]. Analytical Methods, 2018, 10(26): 3207-3214. DOI:10.1039/C8AY00837J.
- [12] CHAI X H, HUANG X W, ZHANG T, et al. Comparative study of E-nose, GC-MS, and GC-IMS to distinguish star anise essential oil extracted using different extraction methods[J]. Separations, 2023, 10(4): 256. DOI:10.3390/separations10040256.
- [13] VILANOVA M, GENISHEVA Z, MASA A, et al. Correlation between volatile composition and sensory properties in Spanish Albariño wines[J]. Microchemical Journal, 2010, 95(2): 240-246. DOI:10.1016/j.microc.2009.12.007.
- [14] HEENAN S P, DUFOUR J P, HAMID N, et al. The influence of ingredients and time from baking on sensory quality and consumer freshness perceptions in a baked model cake system[J]. LWT-Food Science and Technology, 2010, 43(7): 1032-1041. DOI:10.1016/j.lwt.2009.12.009.
- [15] ZHANG Y, SUN Y W, SONG H L. Variation in volatile flavor compounds of cooked mutton meatballs during storage[J]. Foods, 2021, 10(10): 2430. DOI:10.3390/foods10102430.
- [16] SU Y F, WANG H Y, WU Z Y, et al. Sensory description and consumer hedonic perception of ultra-high temperature (UHT) milk[J]. Foods, 2022, 11(9): 1350. DOI:10.3390/foods11091350.
- [17] XIANG Q, XIA Y X, SONG J H, et al. Characterization of the key nonvolatile metabolites in Cheddar cheese by partial least squares regression (PLSR), reconstitution, and omission[J]. Food Chemistry, 2023, 403: 134034. DOI:10.1016/j.foodchem.2022.134034.
- [18] JU S, SONG S, LEE J, et al. Development of nano soy milk through sensory attributes and consumer acceptability[J]. Foods, 2021, 10(12): 3014. DOI:10.3390/foods10123014.
- [19] BORRÀS E, MESTRES M, ACEÑA L, et al. Identification of olive oil sensory defects by multivariate analysis of mid infrared spectra[J]. Food Chemistry, 2015, 187: 197-203. DOI:10.1016/j.foodchem.2015.04.030.
- [20] LIEB V M, ESQUIVEL P, CUBERO CASTILLO E, et al. GC-MS profiling, descriptive sensory analysis, and consumer acceptance of Costa Rican papaya (*Carica papaya* L.) fruit purees[J]. Food Chemistry, 2018, 248: 238-246. DOI:10.1016/j.foodchem.2017.12.027.
- [21] RIBEIRO J S, AUGUSTO F, SALVA T J G, et al. Prediction of sensory properties of Brazilian *Arabica* roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares[J]. Analytica Chimica Acta, 2009, 634(2): 172-179. DOI:10.1016/j.aca.2008.12.028.
- [22] CHUNG S J, HEYMANN H, GRÜN I U. Application of GPA and PLSR in correlating sensory and chemical data sets[J]. Food Quality and Preference, 2003, 14(5/6): 485-495. DOI:10.1016/S0950-3293(03)00010-7.
- [23] ZHANG Z L. Sensory evaluation of different drinks formulated from dendrobium officinale flowers using fuzzy logic[J]. Journal of Food Measurement and Characterization, 2021, 15(4): 3060-3068. DOI:10.1007/s11694-021-00877-y.
- [24] LAZIM M A, SURIANI M. Sensory evaluation of the selected coffee products using fuzzy approach[J]. World Academy of Science, Engineering and Technology, 2009, 38: 726-729.
- [25] BAHRAMPARVAR M, SALEHI F, RAZAVI S M A. Predicting total acceptance of ice cream using artificial neural network[J]. Journal of Food Processing and Preservation, 2014, 38(3): 1080-1088. DOI:10.1111/jfpp.12066.
- [26] CEVOLI C, CERRETANI L, GORIA A, et al. Classification of Pecorino cheeses using electronic nose combined with artificial neural network and comparison with GC-MS analysis of volatile compounds[J]. Food Chemistry, 2011, 129(3): 1315-1319. DOI:10.1016/j.foodchem.2011.05.126.
- [27] WANG J, XIAO Q Y, HUANG H B, et al. Non-target screening and identification of the significant quality markers in the wild and cultivated *Cordyceps sinensis* using OPLS-DA and feature-based molecular networking[J]. Chinese Journal of Analytical Chemistry, 2023, 51(9): 100302. DOI:10.1016/j.cjac.2023.100302.
- [28] KANG C D, ZHANG Y Y, ZHANG M Y, et al. Screening of specific quantitative peptides of beef by LC-MS/MS coupled with OPLS-DA[J]. Food Chemistry, 2022, 387: 132932. DOI:10.1016/j.foodchem.2022.132932.
- [29] SINIJA V R, MISHRA H N. Fuzzy analysis of sensory data for quality evaluation and ranking of instant green tea powder and granules[J]. Food and Bioprocess Technology, 2011, 4(3): 408-416. DOI:10.1007/s11947-008-0163-x.
- [30] CHEN Q S, SUN C C, OUYANG Q, et al. Classification of different varieties of Oolong tea using novel artificial sensing tools and data fusion[J]. LWT-Food Science and Technology, 2015, 60(2): 781-787. DOI:10.1016/j.lwt.2014.10.017.
- [31] BERRUETA L A, ALONSO-SALCES R M, HÉBERGER K. Supervised pattern recognition in food analysis[J]. Journal of Chromatography A, 2007, 1158(1/2): 196-214. DOI:10.1016/j.chroma.2007.05.024.
- [32] GONZALEZ VIEJO C, FUENTES S. Beer aroma and quality traits assessment using artificial intelligence[J]. Fermentation, 2020, 6(2): 56. DOI:10.3390/fermentation6020056.
- [33] QIN H, HUO D Q, ZHANG L, et al. Colorimetric artificial nose for identification of Chinese liquor with different geographic origins[J]. Food Research International, 2012, 45(1): 45-51. DOI:10.1016/j.foodres.2011.09.008.

- [34] LAREF S, LAREF A. Fast determination of a polaron mechanism in beryllium-VI polymorphs using LDA-1/2 approach[J]. Journal of Luminescence, 2013, 135: 248-253. DOI:10.1016/j.jlumin.2012.10.003.
- [35] 叶美霞, 李荣, 姜子涛, 等. 不同年龄乳粉的挥发性风味成分分析及其分类预测[J]. 食品科学, 2022, 43(16): 242-252. DOI:10.7506/spkx1002-6630-20210807-097.
- [36] DE ALMEIDA V E, DE SOUSA FERNANDES D D, DINIZ P H G D, et al. Scores selection via Fisher's discriminant power in PCA-LDA to improve the classification of food data[J]. Food Chemistry, 2021, 363: 130296. DOI:10.1016/j.foodchem.2021.130296.
- [37] GÓMEZ A H, WANG J, HU G X, et al. Monitoring storage shelf life of tomato using electronic nose technique[J]. Journal of Food Engineering, 2008, 85(4): 625-631. DOI:10.1016/j.jfoodeng.2007.06.039.
- [38] MIRICIOIU M G, IONETE R E, COSTINEL D, et al. Classification of *Prunus* genus by botanical origin and harvest year based on carbohydrates profile[J]. Foods, 2022, 11(18): 2838. DOI:10.3390/foods11182838.
- [39] WANG S, NING Y F, SHI H M. A new uncertain linear regression model based on equation deformation[J]. Soft Computing, 2021, 25(20): 12817-12824. DOI:10.1007/s00500-021-06030-7.
- [40] YU P G, LOW M Y, ZHOU W B. Design of experiments and regression modelling in food flavour and sensory analysis: a review[J]. Trends in Food Science & Technology, 2018, 71: 202-215. DOI:10.1016/j.tifs.2017.11.013.
- [41] AFANADOR N L, TRAN T N, BUYDENS L M C. Use of the bootstrap and permutation methods for a more robust variable importance in the projection metric for partial least squares regression[J]. Analytica Chimica Acta, 2013, 768: 49-56. DOI:10.1016/j.aca.2013.01.004.
- [42] SHON D H, PARK S J, YOON S J, et al. Identification of browning in human adipocytes by partial least squares regression (PLSR), infrared spectral biomarkers, and partial least squares discriminant analysis (PLS-DA) using FTIR spectroscopy[J]. Photonics, 2022, 10(1): 2. DOI:10.3390/photonics10010002.
- [43] 刘平, 王雪梅, 向琴, 等. 郫县豆瓣发酵过程中微生物群落与理化指标相关性研究[J]. 西华大学学报(自然科学版), 2020, 39(6): 72-80.
- [44] 贾懿敏, 袁彬宏, 余沛, 等. 基于GC-O-MS及感官评价分析脱皮核桃仁关键风味物质[J]. 食品科学技术学报, 2023, 41(4): 126-134. DOI:10.12301/spxb202200618.
- [45] CHONG I G, JUN C H. Performance of some variable selection methods when multicollinearity is present[J]. Chemometrics and Intelligent Laboratory Systems, 2005, 78(1/2): 103-112. DOI:10.1016/j.chemolab.2004.12.011.
- [46] SONG W R, WANG H, MAGUIRE P, et al. Nearest clusters based partial least squares discriminant analysis for the classification of spectral data[J]. Analytica Chimica Acta, 2018, 1009: 27-38. DOI:10.1016/j.aca.2018.01.023.
- [47] CHEN X J, XU Y L, MENG L W, et al. Non-parametric partial least squares-discriminant analysis model based on sum of ranking difference algorithm for tea grade identification using electronic tongue data[J]. Sensors and Actuators B: Chemical, 2020, 311: 127924. DOI:10.1016/j.snb.2020.127924.
- [48] BALLABIO D, CONSONNI V. Classification tools in chemistry. Part 1: linear models. PLS-DA[J]. Analytical Methods, 2013, 5(16): 3790-3798. DOI:10.1039/C3AY40582F.
- [49] CHUNG I M, KIM J K, HAN J G, et al. Potential geo-discriminative tools to trace the origins of the dried slices of shiitake (*Lentinula edodes*) using stable isotope ratios and OPLS-DA[J]. Food Chemistry, 2019, 295: 505-513. DOI:10.1016/j.foodchem.2019.05.143.
- [50] WORLEY B, POWERS R. PCA as a practical indicator of OPLS-DA model reliability[J]. Current Metabolomics, 2016, 4(2): 97-103. DOI:10.2174/2213235X04666160613122429.
- [51] THÉVENOT E A, ROUX A, XU Y, et al. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses[J]. Journal of Proteome Research, 2015, 14(8): 3322-3335. DOI:10.1021/acs.jproteome.5b00354.
- [52] LU K, LIU L, XU Z Y, et al. The analysis of volatile compounds through flavoromics and machine learning to identify the origin of traditional Chinese fermented shrimp paste from different regions[J]. LWT-Food Science and Technology, 2022, 171: 114096. DOI:10.1016/j.lwt.2022.114096.
- [53] YANG Z Y, BALDERMANN S, WATANABE N. Recent studies of the volatile compounds in tea[J]. Food Research International, 2013, 53(2): 585-599. DOI:10.1016/j.foodres.2013.02.011.
- [54] SEISONEN S, VENE K, KOPPEL K. The current practice in the application of chemometrics for correlation of sensory and gas chromatographic data[J]. Food Chemistry, 2016, 210: 530-540. DOI:10.1016/j.foodchem.2016.04.134.
- [55] MUKHOPADHYAY S, MAJUMDAR G C, GOSWAMI T K, et al. Fuzzy logic (similarity analysis) approach for sensory evaluation of chhana podo[J]. LWT-Food Science and Technology, 2013, 53(1): 204-210. DOI:10.1016/j.lwt.2013.01.013.
- [56] MUSTAFA A. Experimental analysis of artificial neural networks used for function approximation[J]. International Journal of Computer Science and Network Security, 2019, 19(9): 82-90.
- [57] BHAGYA RAJ G V S, DASH K K. Comprehensive study on applications of artificial neural network in food process modeling[J]. Critical Reviews in Food Science and Nutrition, 2022, 62(10): 2756-2783. DOI:10.1080/10408398.2020.1858398.
- [58] ZENG X Q, CAO R, XI Y, et al. Food flavor analysis 4.0: a cross-domain application of machine learning[J]. Trends in Food Science & Technology, 2023, 138: 116-125. DOI:10.1016/j.tifs.2023.06.011.
- [59] LIU M, HAN X M, TU K, et al. Application of electronic nose in Chinese spirits quality control and flavour assessment[J]. Food Control, 2012, 26(2): 564-570. DOI:10.1016/j.foodcont.2012.02.024.
- [60] HIGH R, EYRES G T, BREMER P, et al. Characterization of blue cheese volatiles using fingerprinting, self-organizing maps, and entropy-based feature selection[J]. Food Chemistry, 2021, 347: 128955. DOI:10.1016/j.foodchem.2020.128955.
- [61] SHEN C, CAI Y, DING M Q, et al. Predicting VOCs content and roasting methods of lamb shashliks using deep learning combined with chemometrics and sensory evaluation[J]. Food Chemistry: X, 2023, 19: 100755. DOI:10.1016/j.fochx.2023.100755.
- [62] KHAN M I H, SABLANI S S, JOARDDER M U H, et al. Application of machine learning-based approach in food drying: opportunities and challenges[J]. Drying Technology, 2022, 40(6): 1051-1067. DOI:10.1080/07373937.2020.1853152.