

图书基本信息

书名：<<食品、农产品无损检测中的数据处理和分析方法>>

13位ISBN编号：9787030342744

10位ISBN编号：7030342747

出版时间：2012-5

出版单位：科学出版社

作者：赵杰文、林颢

页数：231

字数：315250

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

内容概要

《食品、农产品检测中的数据处理和分析方法》是一部系统介绍和归纳食品、农产品无损检测数据处理和分析的科研论著，基于对所获数据进行处理和分析，从而建立有效的识别模型。

各章节介绍了各种数据处理和分析方法的基本原理，并通过大量的实例阐述这些方法如何应用于食品、农产品的品质检测和分析。

《食品、农产品检测中的数据处理和分析方法》精华部分大多来自作者在食品、农产品检测方面多年研究成果的积累，并结合国内外食品、农产品无损检测中数据处理和分析的最新方法，为相关领域科研人员接触数据处理和分析方法的最新动态提供帮助，具有鲜明的特征和实用性。

《食品、农产品检测中的数据处理和分析方法》可供从事食品工程、食品分析、农业工程方面的教学及科研工作者参考。

作者简介

无

书籍目录

序前言第一章 绪论第一节 食品、农产品品质无损检测技术及其特点第二节 食品、农产品品质无损检测中的数据处与分析一、数据前处理二、变量筛选三、特征提取四、定性识别五、定量分析第三节 数据处和分析在食品、农产品无损检测中的应用趋势一、多学科知识交叉二、计算机和数据处处理软件作用凸显主要参考文献第二章 数据前处理第一节 标准化处处理一、均值中心化二、极小/极大归一化三、标准正态变量变换四、数据标准化处处理中的应用实例第二节 数据平滑与去噪一、数据平滑二、求导去噪三、自适应滤波四、小波分析五、数据平滑、去噪应用实例第三节 其他数据前处处理方法一、净分析物预处处理法二、正交信号校正法主要参考文献第三章 变量筛选第一节 区间筛选法一、区间偏最小二乘二、前向区间偏最小二乘三、后向区间偏最小二乘四、联合区间偏最小二乘五、变量区间筛选法应用实例第二节 遗传算法一、遗传算法基本原理二、遗传偏最小二乘算法三、遗传算法在变量筛选中的应用实例第三节 模拟退火法一、模拟退火法基本原理二、模拟退火法实现过程三、模拟退火法在变量筛选中的应用实例第四节 其他变量筛选方法一、连续投影算法二、无信息变量消除法主要参考文献第四章 特征提取第一节 常规特征提取方法一、数据处处理中的常规特征提取二、应用实例第二节 主成分分析一、二维空间主成分分析二、多维空间主成分分析三、主成分分析在数据特征提取中的应用实例第三节 独立分量分析一、独立分量分析算法原理二、独立分量分析在数据特征提取中的应用实例主要参考文献第五章 定性识别第一节 聚类分析一、聚类分析方法原理二、聚类分析在数据分析中的应用实例第二节 线性判别分析一、欧氏距离线性判别二、马氏距离线性判别三、Fischer线性判别四、线性判别在数据分析中的应用实例第三节 K最近邻法一、K最近邻法算法原理二、K最近邻法在数据分析中的应用实例第四节 人工神经网络一、人工神经网络概述二、BP神经网络三、RBF神经网络四、遗传神经网络五、人工神经网络模型应用实例第五节 支持向量机一、支持向量机的原理二、支持向量机的构造三、支持向量机识别的应用实例第六节 一类分类器一、常用一类分类器二、一类支持向量机三、支持向量数据描述四、一类分类器在数据分析中的应用实例主要参考文献第六章 定量分析第一节 回归分析一、一元线性回归二、多元线性回归三、主成分回归四、回归在数据分析中的应用实例第二节 偏最小二乘法一、偏最小二乘法原理二、模型的评价指标三、偏最小二乘在数据分析中的应用实例第三节 神经网络回归一、BP算法神经网络回归分析二、前馈神经网络模型模糊感知器回归分析三、神经网络回归在数据分析中的应用实例第四节 支持向量机回归一、支持向量机回归原理二、最小二乘支持向量机三、最小二乘支持向量机在数据分析中的应用实例主要参考文献第七章 数据处处理和分析中几种常用软件简介第一节 Excel一、Excel软件在数据处处理和分析中的应用简介二、Excel软件处处理数据实例第二节 SPSS一、SPSS数据文件的建立与操作二、SPSS软件处处理数据实例第三节 MATLAB一、MATLAB 7.x的运行环境二、MATLAB软件处处理数据实例主要参考文献附录A 食品、农产品无损检测数据处处理和分析相关期刊简介附录B 数据处处理与分析算法及源代码下载相关网站简介

章节摘录

第一章 绪论第一节 食品、农产品品质无损检测技术及其特点食品、农产品的品质检测是指运用数学、物理、化学、生物等学科的基本理论及各种科学技术,对检测对象包括生产原料、辅助材料、半成品、成品、副产品等的状态和主要成分含量及微生物状况进行分析检测。

在对食品、农产品品质进行检测时,因检测目的不同,且检测对象的性质和状态差异较大,所选择的检测方法也各不相同。

无损检测又称为非破坏检测,是近年发展起来的一种新技术,是指在不破坏样品的情况下对其进行品质评价的方法。

食品、农产品无损检测技术涉及光学、力学、电学、磁学等学科,范围广泛,其基础更是涉及材料科学、计算机技术、生物技术、信息技术等诸多领域,其中以光学检测发展最快。

近年来,随着科学技术的发展,无损检测技术也得到长足发展,目前已呈现出两个重要的发展趋势。

其中一个重要发展趋势即无损检测技术与传感器技术、纳米技术及计算机技术结合得越来越紧密,使无损检测分析仪器不但具有越来越强大的“智能”,而且正沿着落地式|台式|移动式|便携式|手持式|芯片实验室的方向发展,越来越小型化、微型化、智能化;检测分析仪器和专用计算机的界限在今后也将变得越来越模糊,许多检测分析仪器实际上是具有某种检测分析功能的计算机。

目前,食品、农产品检测系统多数仍处于智能化的低级阶段,系统只能把计算机技术与传统的食品、农产品检测分析结合起来,仅能适应被测参数的变化、自动补偿、自动选择量程等。

因此,将无损检测技术与传感器技术、纳米技术及计算机技术等多学科交叉结合,开发出能识别与解释各种光学、力学谱图的食品、农产品“智能化”检测系统,成为当前食品、农产品检测智能化研究的热点。

无损检测技术发展的另一个重要趋势,是数据处理与分析方法在检测中显示出越来越重要的作用。

科学技术的发展对现代食品、农产品无损检测技术提出了更高的要求,人们不仅要求及时、精确、可靠地获得有关待测样品品质的数据,而且要求全面快速地分析。

无损检测技术不仅要解决有关测量数据的获取问题,更需要解决从大量数据中提取有用信息的问题并建立相应的模型。

尤其是近年来在食品、农产品无损检测技术中所采用仪器的精密度越来越高,所获取的数据量也越来越大。

在一些实验中,一个样品测试一次即可获取几千甚至数十万个数据,对数据信号的前处理、数据的精简、数据变量的筛选、特征信息的提取以及识别模型的建立成为无损检测技术研究的热门课题,模糊数学、统计学、信号处理、化学计量学及模式识别方法等数学方法也越来越多地应用于无损检测数据处理中。

第二节 食品、农产品品质无损检测中的数据处理与分析随着仪器精密度不断提高,食品、农产品无损检测仪器所获得的数据量不断增大,如何运用恰当的数据处理方法,在庞大的数据量中挖掘出能准确描述检测对象的有用信息,建立鲁棒性强的数学模型,已成为无损检测研究的热门课题。

在数据的处理与分析中,数据的前处理、变量筛选、特征提取、定性识别模型和定量分析模型的建立是其重要组成部分。

一、数据前处理在食品、农产品品质无损检测中,检测器所获取的数据信号除含样品待测成分信息外,还包括各种仪器噪声,如高频随机噪声、基线漂移、杂散信号、样品背景等。

因此,在数据分析前,首先应针对特定的信号测量和样品体系进行合理的处理,减弱甚至消除各种非目标因素对检测信号信息的影响,为建立稳定、可靠的数学模型奠定基础。

常用的数据前处理方法有数据标准化处理(均值中心化、归一化、标准正态变量变换等)、高频噪声滤除(卷积平滑、傅里叶变换、小波变换等)、信号的微分求导和基线校正等。

数据标准化处理是将原始数据矩阵中各元素减去该列元素均值后,再除以所在列元素的方差。

其特点是数据矩阵的一列元素权重相同,均值都为0,方差和标准化都为1。

卷积平滑法是基于最小二乘法原理,保留分析信号中的有用信息,消除随机噪声,但是过度的平滑

将会造成检测信号中部分有用信息的丢失。

基线校正主要是扣除仪器背景或漂移对信号的影响，可以采取偏置扣减、微分求导处理和基线倾斜等方法。

采用微分求导可以较好地净化谱图信息，在降低噪声的同时也可放大检测信号，但需注意的是，微分求导窗口数据点的大小对结果有一定影响，因此在微分求导的时候需对窗口大小做出合理的选择。

二、变量筛选

变量筛选（也称为变量选择或特征选择）是指从原始变量中挑选出一些有代表性的特征变量，代替原始变量进行数据分析和处理。

在食品、农产品无损检测实验中，检测仪器每次可获取大量的数据，但对应着不同原始变量的数据对待测样品品质信息的贡献率不尽相同，有些变量反映的信息量较为丰富，有些变量反映的信息量较少，甚至与待测样品基本无关。

如果将检测器所获取的数据都用于建模，则建模计算过程将极为烦琐，计算量很大，建立的模型也较为复杂，鲁棒性差。

研究表明，通过特定的变量筛选方法对自变量进行优选，不仅可简化模型，更重要的是可剔除不相关或非线性变量，得到预测能力强、稳健性好的训练模型。

常用的变量筛选方法有穷尽搜索法、区间（如前向区间、后向区间、联合区间）筛选法和以某种算法（如模拟退火、遗传算法、无变量信息消除）为指导思想的随机性质搜索方法。

穷尽搜索法从原始数据集中将所有可能组合都搜索一遍，这种方法一定能得到一个最优子集，但这一方法很少用，因为由它带来的计算量可能让人难以承受。

区间筛选法即将原始数据的变量分为若干个区间，对每个区间或某几个区间的变量建立相应的模型，选取最优区间。

相比穷尽搜索，区间筛选大大减少了工作量，但同样会存在所选取区间的变量间对待测样品信息贡献率不同的问题。

以模拟退火算法为代表的随机性质搜索方法可选择与检测对象品质相关的信息，但该方法在搜索特征变量时搜索的范围太广，带有一定的盲目性，易陷入局部最优。

因此，将区间筛选法和随机搜索方法结合，可减少搜索的盲目性，提高变量筛选效率。

三、特征提取

特征提取是指通过映射（或变换）的方法对原始数据进行重组，以期用较少的特征值描述原始数据中所包含的大部分信息。

由于对所处理的信息了解不深刻，且还有许多因素之间的关系及相关程度亦不能肯定，人们往往先根据化学（或物理）的选择标准，尽可能地把一切相关、又容易获取的特征变量都提取出来，然后借助于数学方法，筛选出对模型的建立起较大作用的特征变量。

在实际计算中，一些不相关的特征变量会降低模型的鲁棒性，因此研究人员总是力图抛弃那些对建模作用不大的特征变量，在保证模型精度的前提下，使特征变量数减到最少。

常用的特征提取方法有直接从原始数据中提取一些特征参数（如均值、标准差、最大值、极差等）法，从原始数据中提取的特征参数比较直观、简便，但所反映的信息较为粗糙；还有以某种算法（如主成分分析法、独立分量分析法等）为依据，从原始数据中提取一些特征变量。

主成分分析是把多个指标化为几个综合指标的一种统计方法，它沿着协方差最大方向由多维数据空间向低维数据空间投影，各主成分向量之间相互正交。

通过选择合理的主成分既可以达到降维的目的，又不会过多地丢失原始数据信息，同时可以减少原始数据中的冗余信息。

主成分分析可保证分解出的分量互相正交；独立分量分析利用信号的高阶统计量，要求分解出的各分量尽可能独立，在信号的特征提取中则表现出更大的优势。

四、定性识别

定性识别（也称为模式识别）是指对表征事物或现象的各种形式（如数值、文字和逻辑关系等）的信息进行处理和分析，以便对事物或现象进行描述、辨认、分类和解释的过程。

按照识别时是否需要具有识别样本的先验知识，模式识别方法可分为“监督学习分类”和“非监督学习分类”。

在模式识别的特征空间里，如果所分类的情况是已知的，在此基础上，可以选择一个合适的距离尺度，以得到有关这些类的分布形状以及典型模式的信息，这种方法称为“监督学习分类”。

常见的模式识别方法多为“监督学习分类”方法，如线性判别分析、人工神经网络、偏最小二乘分类、支持向量机等。

在模式识别中，也会遇到不能事先获取任何关于样本的先验知识（很多时候需要在无监督情形下将很多东西分类）的情况。

因此，分类系统必须先通过一种有效的方法去发现样本的内在相似性，然后指导同类检测对象的分类，这种方法称为“非监督学习分类”。

最常见的非监督学习分类法为聚类分析法，包括“树聚类”和“K 均值聚类”等。

此外，按照使用的分类函数，模式识别方法可分为线性和非线性判别分析方法。

线性判别分析方法有欧氏距离、马氏距离、费歇尔投影法和K 最近邻法等；非线性判别分析方法有人工神经网络、支持向量机和支持向量数据描述等。

五、定量分析定量分析是指分析一个被研究对象所包含成分的数量关系或所具备性质间的数量关系；也可以对几个对象的某些性质、特征、相互关系从数量上进行分析比较，研究的结果也用“数量”加以描述。

定量分析是依据统计数据，建立数学模型，并用数学模型计算出分析对象的各项指标及其数值的一种方法。

相比定性识别而言，定量分析更加科学，需运用到更多的数学计算；定性识别虽然较为粗糙，但在数据资料不够充分或分析者数学基础较为薄弱时比较适用。

这两种分析方法对数学知识的要求虽然有高有低，但并不能就此把定性识别与定量分析划分开来。

事实上，现代定性识别方法同样要采用数学工具进行计算，而定量分析则必须建立在定性预测基础上，二者相辅相成，定性是定量的依据，定量是定性的具体化，二者结合起来灵活运用才能取得最佳效果。

线性回归法（包括一元线性回归和多元线性回归）是最早采用的定量分析法。

由于食品、农产品无损检测所获取的数据量比较大，主成分回归及在主成分分析基础上进行的偏最小二乘分析法已越来越多地运用在食品、农产品品质无损检测中；随着非线性方法研究的不断拓展，一些传统的定性识别方法经过改造，也逐步应用到定量分析中，如人工神经网络回归和支持向量回归法。

第三节 数据处理和分析在食品、农产品无损检测中的应用趋势一、多学科知识交叉在食品、农产品无损检测中，数据处理和分析已凸显出越来越重要的作用，其内涵和外延也不断扩展，各门学科中与数据处理相关的最新研究成果经消化、吸收、更新，不断地被应用于食品、农产品无损检测的数据处理和分析中。

在数学领域，应用数学学科的最新研究成果（如小波分析、随机过程相关科研成果）正越来越快地被应用于数据信号的处理和随机信号的分析中；在通信领域，卡尔曼滤波、最小均方自适应滤波、高阶谱分析等方面的最新研究成果也在最短时间内被应用于数据信号的平滑、滤波去噪等信号前处理和分析中；在化学计量学领域，遗传算法和模拟退火算法等方面的最新研究成果也不断被应用于数据矩阵特征提取和变量筛选中；在模式识别领域，各种模式识别方法的最新成果也被应用于数据信号的定性识别中，各种软件也越来越多地被用到食品、农产品无损检测的数据处理和分析中去。

各个学科间的相互交叉是现代科学技术发展的重要趋势，各个学科之间的界限越来越模糊，在食品、农产品无损检测数据处理和分析中表现得尤为明显。

计算机、通信、化学计量学等领域的数据处理和分析都是以数学为基础，有一定的交叉但又明显带有各自学科的特色，这些学科的知识又可以相互交叉应用于无损检测中的数据处理与分析，如一批实验样品的检测数据，其数据前处理部分采用的是通信领域的信号处理，特征变量提取与筛选则应用到化学计量学的知识，而数据的识别则需采用模式识别方法。

随着科技的发展，各个学科之间的交叉将会越来越明显，且各个学科中的最新研究成果也将会越来越快地被应用到食品、农产品无损检测数据处理与分析中。

二、计算机和数据处理软件作用凸显在食品、农产品无损检测中，实验所获取的数据量越来越大，一批实验结果获取到几十或几百万个数据量的现象十分常见，且随着仪器检测精度、速度进一步提高，所获取的数据量将会更庞大。

检测仪器获取的大量（甚至可谓海量）数据必将伴随着极为烦琐的数据处理过程，纯粹依赖人工不可能完成如此庞大的数据计算量，借助计算机来处理数据已经成为必然趋势，且由于处理数据量过大，一些简易的计算软件也难以处理如此庞大的数据。

因此，现代数据处理和分析对计算机提出了更高的要求，即选择快速、简便、适宜的软件处理数据以缩短计算过程和时间，并保证计算结果的准确性。

目前，可用于数据处理和分析的通用软件有Excel、SPSS（statistical product and service solutions）、SAS（statistical analysis system）、MATLAB（matrix laboratory）、VB（visual basic）、VC++（microsoft visual C++）等，还有一些专用型的实验数据优化、图形可视化和数值分析软件，如Design-expert、Origin、Probit等。

其中，最常用的是Excel、SPSS、SAS和MATLAB。

Excel以表格的方式输入与管理数据，能进行简单的数据处理和分析，如数据排序、回归分析等，操作比较简单、易学；SPSS是IBM公司开发的最早采用图形菜单驱动界面的统计软件，采用类似Excel表格的方式输入与管理数据，数据接口较为通用，能方便地从其他数据库中读入数据，其统计过程包括常用的、较为成熟的统计过程，可以满足非统计专业人士的工作需要；SAS是由美国北卡罗来纳州州立大学开发的统计软件，功能及操作方法与SPSS基本一致；MATLAB是矩阵实验室（matrix laboratory）的简称，是美国MathWorks公司出品的商业数学软件，用于算法开发、数据可视化、数据分析以及数值计算的高级技术计算语言和交互式环境。

MATLAB可以进行矩阵运算、绘制函数和数据、实现算法、创建用户界面、连接其他编程语言的程序。

几乎所有食品、农产品无损检测中数据处理与分析，都可以在MATLAB中实现。

随着科学技术的发展，各种软件版本也不断更新，目前SPSS和SAS软件已更新了十多个版本，而MATLAB软件更新了二十多个版本。

在食品、农产品数据处理量不断增大且精度要求越来越高的趋势下，各种软件将不断优化更新，而另外一些专用型软件也将不断推出。

此外，对计算机硬件也将提出更高的要求，或许不久的将来，处理食品、农产品无损检测数据的专用型计算机也会诞生。

……

编辑推荐

《食品农产品检测中的数据处理和分析方法》该书围绕如何对所获取的数据进行处理和分析，从而建立有效的识别模型展开讨论，分为数据前处理、变量筛选、特征提取、定性识别模型和定量分析模型等几部分，是作者在食品、农产品检测方面多年研究成果的积累，同时也反映了计算机科学和应用数学的最新研究成果在食品、农产品检测方面的应用动态，该书是国内外第一本系统介绍和归纳关于食品、农产品无损检测数据处理和分析的科研论著，该书的出版将对食品、农产品品质检测技术的发展起到促进作用，同时也将为广大从事该领域研究的科学工作者、技术人员和研究生们提供一本内容全面、反映数据处理和分析方法最新动态的技术参考书。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>