

中国无公害农产品信息数据库的设计

——以种植业为例

高添^{1,2}, 龙怀玉², 汪景宽¹, 廖超子³, 雷秋良², 张认连², 张燕², 杨柳^{1,2}

(¹沈阳农业大学土地与环境学院, 沈阳 110161;

²中国农业科学院农业资源与农业区划研究所, 农业部作物营养与养分循环重点开放实验室, 北京 100081;

³农业部农产品质量安全中心, 北京 100081)

摘要:通过数据库技术对无公害农产品认证的产地、产品及相关信息进行管理是提高认证效率的一个重要手段。从中国无公害农产品认证的现状出发, 针对认证过程中信息的多源性、时效性、连续性、关联性和多类性的数据特征, 采用实体-联系的建模方法, 分析认证过程中的产品、产地、标准等实体的属性及其关系, 确立农产品认证分类信息、产地信息、产品信息等六类信息, 并建立字段和表, 构建了基于关系模型的无公害农产品信息数据库, 旨在为中国无公害农产品信息的管理提供高效支持。

关键词:无公害农产品; 关系型数据库; 数据特征

中图分类号: S126

文献标识码: A

论文编号: 2009-0731

Design and Realization of Database for Nuisanceless Agro-Food Information in China—A Case Study in Planting

Gao Tian^{1,2}, Long Huaiyu², Wang Jingkuan¹, Liao Chaozi³, Lei Qiuliang²,
Zhang Renlian², Zhang Yan², Yang Liu^{1,2}

(¹ShenYang Agricultural University, College of Land and Environmental Science, Shenyang 110161;

²The Chinese Academy of Agriculture Sciences, Institute of agricultural Resources and

Regional Planning, Ministry of Agriculture Key Laboratory of Crop Nutrition and Fertilization, Beijing 100081;

³The Center For Agro-food Quality & Safety, Ministry of Agriculture, Beijing 100081)

Abstract: It is very helpful to establish database managing producing area, production quality, environment standard and other relative information of nuisanceless agro-food, for the improvement of nuisanceless agro-food certificating. In this paper, the method of ER was utilized to analyze the properties and interrelations of entities such as agro-food, agro-food area and standard, etc. based on status and data characteristics as multi-sources, time effectiveness, continuity, relationship and multi-category of nuisanceless agro-food certification. The relational database of nuisanceless agro-food was accomplished by the classification of the information of nuisanceless agro-food and establishment of tables and the fields, and the objective was to provide an efficient support for the management of nuisanceless agro-food information.

Key words: nuisanceless agro-food, relational database, characteristic of data

基金项目: 国家 863 计划“农产品数字化认证平台”(2006AA10Z270); 国家科技支撑计划“西南地区地力耗竭型种植系统耕地保育综合技术与示范”(2006BAD05B06-03)

第一作者简介: 高添, 男, 1983 年出生, 辽宁锦州人, 硕士研究生, 主要从事土地利用信息技术与农产品质量安全方面的研究。

通讯作者: 龙怀玉, 男, 苗族, 1969 年出生, 湖南绥宁人, 研究员, 博士, 硕士生导师, 主要从事数字土壤、农业信息技术、农田土壤和滨海盐土的质量演变等方面的研究。通信地址: 100081 北京市中关村南大街 12 号中国农业科学院农业资源与农业区划研究所, Tel: 010-82108685, E-mail: hylong@caas.ac.cn。

收稿日期: 2009-04-09, **修回日期:** 2009-09-01。

0 引言

农产品质量安全关系着国民的健康,影响着社会稳定^[1]。对农产品进行认证是保障其质量安全的有效手段,而无公害农产品是保障中国农产品基本质量安全的农产品认证形式,尽管起步晚,但发展快,认证数量巨大。截止到2008年11月底,全国范围内已累计认证的无公害农产品达47000多个,认证产地44000多个,而且数量还在迅速增长。然而,由于中国无公害农产品认证从申报到发证经过环节较多,认证周期较长,产生的文字材料很多^[2],致使无公害农产品信息复杂,种类繁多,既有描述性定性信息,又有检测的定量信息,管理较为困难。

建立无公害农产品信息数据库是提高无公害农产品管理能力的基础技术支持。数据库技术是上个世纪逐渐形成的、目前仍在继续发展的先进计算机技术之一,它为海量信息的管理提供了重要的支持。如今,数据库技术已经应用到农业气象^[3]、土壤^[4]、农产品^[5-8]等诸多农业领域,在信息的储存、管理、应用以及科学研究中发挥着巨大的作用^[9]。然而,数据库技术的优越性在中国无公害农产品认证中却没有得到充分发挥。目前,中国无公害农产品信息管理还停留在传统的纸质文本的管理,不但不利于信息的高效管理与统计,而且难于长期储存与快速传输;数据不规范,缺乏产地的空间信息,数据信息的管理效率较低,这在很大程度上制约了中国无公害农产品的发展。因此,通过构建无公害农产品信息数据库对大量的农产品信息进行高效管理是非常有意义的。针对以上需求,笔者初步设计且实现了无公害农产品信息数据库。

1 材料与方法

无公害农产品及产地的相关信息主要从由农业部农产品质量安全中心提供的相关资料中获取,这些资料包括:无公害农产品认证申报材料、无公害农产品产地认定申请书、无公害农产品产地认定与产品认证现场检查报告、无公害农产品产地与产品认证报告等。无公害农产品标准信息主要从中国标准出版社出版的《无公害食品标准汇编》中提取。

在研究过程中,首先,在充分遵循原始数据的基础上,将以上材料录入到计算机中、进行数字化处理;然后逐步划分出最小信息单元,并分析出这些信息单元的数据特征,同时将数据转化为适于数据库存储的形式(如将字段“产地规模”的数值与单位分开存储),建立字段,并定义字段属性;最后将所有信息按照认证流程与描述的实体分类建表,依据数据库设计规则^[10],建立数据结构与表关系。

空间数据的建立。中国现有的绝大多数无公害农产品产地数据、产品数据均缺乏数值化的空间信息,只有对其所在地方的地名描述,而数值化的空间信息在产地与产品的信息发布、管理中是至关重要的。因此,利用ESRI公司的ArcInfo对产地地址的名称与中国69万个居民点名称(Geodatabase格式)进行空间匹配,以赋予其地理坐标。

2 中国无公害农产品信息的几个基本特征

对数据进行特征分析是建立数据库的重要前提。中国无公害农产品认证的环节较多,周期比较长,涉及信息的数据项纷杂。通过对以上提到的材料进行分析,可以发现中国无公害农产品的数据信息具有以下几个明显特征:

(1)多源性

从数据来源上来看,无公害农产品信息具有多源性。中国无公害农产品信息来源十分复杂,如申报人提交的产品产地申报信息、检测机构产生的检测数据与检测报告、农产品质量安全中心颁发的无公害农产品证书信息以及认证过程中遵循的无公害农产品标准信息等。数据的多源性也在一定程度上加大了数据库建设的难度。

(2)时效性

从信息利用的角度上来看,无公害农产品信息具有很强的时效性。中国无公害农产品证书有效期为三年,即每种被认证的农产品的大部分信息,如检测信息、现场检查信息、产品基本信息等只在颁证后三年内是有效的,证书过期后如需继续使用的无公害农产品生产者应申请复查^[9]。信息的高时效性要求数据库具有灵活的结构以支持信息的快速更新。

(3)连续性。

从信息的发生角度来看,无公害农产品的数据信息具有动态连续性。如认证的产品数量不断增加、无公害农产品证书到期后的复查换证等,这些都直接导致了产品、产地记录的快速增加以及前后信息较强的连续性,这要求设计的数据库有良好的扩展性与移植性。

(4)关联性。

从信息整体结构上来看,无公害农产品数据信息之间是有机关联的。由于信息是随着无公害农产品认证的过程逐步产生的,所以每种被认证的产品的各项信息以产品为纽带相互关联,如产地与产品的关系(多对多),产地、产品与标准的关系(多对一),产地与申报和现场检查的关系(一对一)等,关系较为复杂。这要求数据库具有良好的数据模型以保证数据库运行的高

效性。

(5)多类性。

从数据的组织和表达形式上来看,认证信息的数据主要通过文本型、数字型、逻辑型等多种方式表达。其中文本型的字符长度主要集中在10~50个字节的区间,数字型的字符长度为2~4个字节;一些标准信息定义为数字区间,如无公害大豆油碘值的区间在860~1070g/kg的区间内;信息中含有一些长文本的描述性信息,以备注型进行表示,如茉莉花茶加工技术规程的描述信息达上百个字节^[11];一些备案信息以图片形式存储,如证书原件、产地环境等,信息的表达形式较多。

综上所述,中国无公害农产品认证信息具有多源性、时效性、连续性、关联性以及表达方式多类性等特征。针对这些特点,研究选用关系型数据库^[12]。

3 中国无公害农产品认证的流程与实体分析

实体的确立是数据库设计的首要环节。中国的无公害农产品认证需要经历申请、审查、检测评估和发证四个依次开展的基本程序^[2]。根据这个认证流程,无公害农产品数据库共涉及五个实体^[9]:

(1)产品。已经认证的无公害农产品。该实体的信息包括无公害农产品在接受检查时的检测信息、产品详细信息以及颁发的无公害农产品证书的基本信息等。

(2)产地。已经认证的无公害农产品产地。该实体的信息包括无公害农产品产地在接受检查时的检测信息、产地环境背景信息以及无公害农产品产地证书的基本信息等。

(3)申报人。通过无公害农产品认证的申报者。

该实体的信息为申报人在申报时递交的信息,包括申报人的基本信息,申报产品产地基本信息以及肥料农药的使用信息等。

(4)现场检查。通过无公害农产品认证的现场检查。该实体的信息为审查过程中现场检查产生的信息。包括符合性判定信息,土壤质量评价信息以及检查员、工作记录信息等。

(5)标准。无公害农产品执行标准。包括无公害农产品在进行认证时所执行的各类标准。

4 数据关系模型

数据关系模型是数据库的核心部分,关系模型的好坏决定着数据库的优劣,同时也为以后不可避免的修改、扩充、重新整合提供基础支持^[13]。笔者采用实体-联系的信息建模方法,根据认证流程中所涉及的信息实体设计其关系模型,通过ERD(实体关系图)鸭脚表示法阐述数据库的实体关系^[10]。

4.1 产地与产品及分类目录

产地与产品、分类目录的关系是数据库中最基础的关系,它主要反应了产品与产地的从属情况,因为产地具有空间坐标,通过产地与产品的关联就赋予了产品空间信息,这样就为实现了产品的空间可视化提供了数据支持。由于中国无公害农产品管理体系还不够完善,存在一种无公害农产品对应多个产地,而一个无公害产地又对应多种农产品的复杂情况,即为多对多关系,研究通过建立中间表建立关联。图1是产品与产地的实体关联图,通过产地ID与产品ID链接,图中包含了实体的部分属性(为了表示简洁,未列出多对多关系中的中间表,部分属性使用英文简写),具体字段将在下文详述。

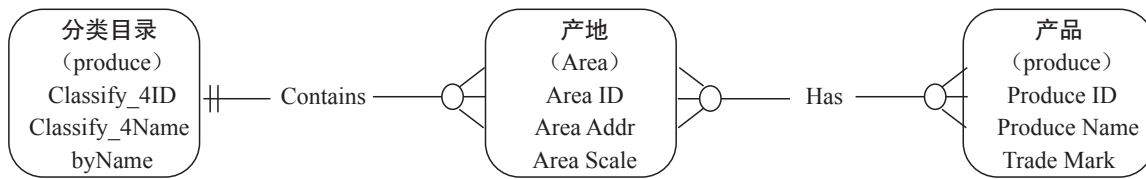


图1 产品与产地实体关联图

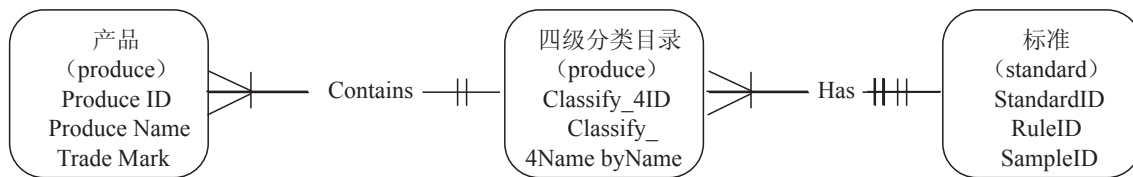


图2 产品与标准实体关联图

4.2 产品与标准

标准即为被认证的无公害农产品所执行的标准。产品与标准以产品分类目录为桥梁进行链接,产品与

分类目录是多对一的关系;从整体上来看,分类目录与标准是多对一的关系,通过产品分类编号“Classify_4ID”进行链接,图2是产品与标准的实体关联图。具

体字段将在下文详述。

4.3 产地与现场检查信息

现场检查与产地是一一对应的关系,通过现场检查编号进行链接(Inspect ID),图3是现场检查与产地的实体关联图。图中列出了现场检查信息中的几个重要属性字段名。

4.4 产品、产地与申报信息

申请人申报信息包括申报产品信息与申报产地信

息等,与相应的产品产地对应关联,图4是产品、产地与申报信息的实体关联图。图中列出了申请人申报信息中的几个重要属性字段名。

4.5 全局的关系模式图

根据系统中实体与关系分析的结果,进行系统关系的集成,生成全局的关系模式图。如图5所示:

5 属性表的设计

根据上述实体关系模型及无公害农产品认证的资

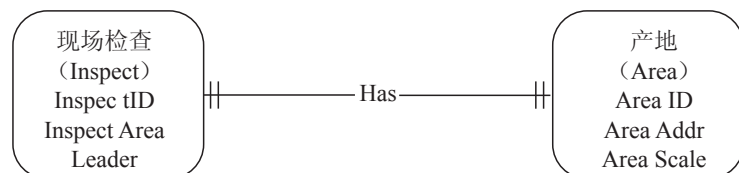


图3 现场检查与产地实体关联图



图4 产品、产地与申报信息的实体关联图

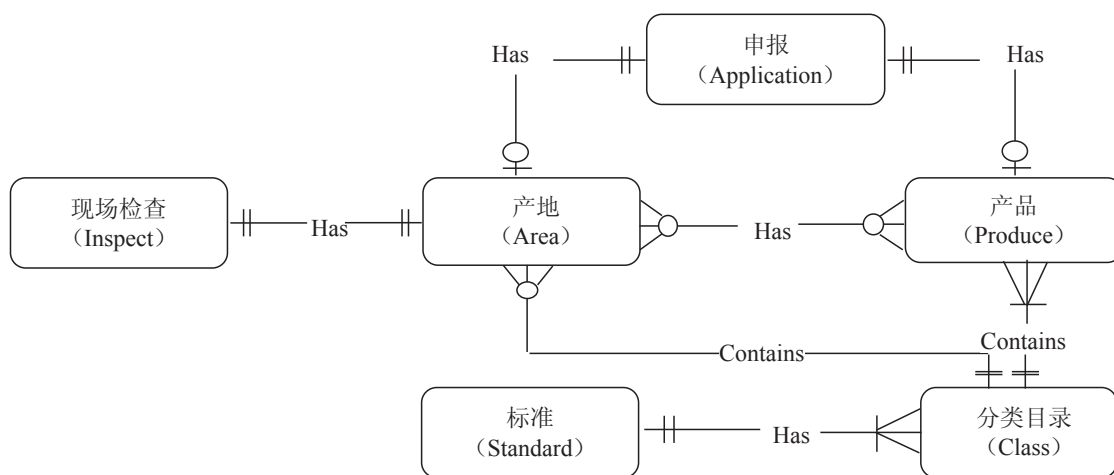


图5 无公害农产品数据库全局关系图

料,按照信息描述的实体与内容把属性信息分为五大类,分别为认定的产地信息、认证的产品信息、申请人申报信息、现场检查信息、无公害农产品标准信息,同时建立无公害农产品索引分类表,用于实现产品分类的检索。

5.1 认证农产品分类信息

该类信息是针对认证的无公害农产品的索引目录及产品分类目录所设计的,描述了无公害农产品的分类情况,主要实现信息的检索功能。包括的表有无公害农产品认证目录表、产品一级分类表、产品二级分类表、产品三级分类表、产品四级分类表;包括的字段有

“产品一级分类”、“产品二级分类”、“认证产品ID”、“产品名称”、“产品别名”等,各个表通过各自的分类ID作为主键实现链接。

每种被认证的农产品入库时将根据计算机编号系统自动分配一个编号,即“认证产品ID”,通过这个ID号与分类信息及其它信息建立链接。产品(自然农产品)共有四级分类,第一级分为种植业、畜牧业和渔业;第二类为蔬菜类、粮油类、水果等;第三类根据《无公害食品标准汇编》进行分类,包括白菜类蔬菜、薯芋类蔬菜、洋葱、大蒜、水生蔬菜等;第四类为最细分类,不

能再分。如白菜类蔬菜中的大白菜、小白菜、乌塌菜、日本水菜、薹菜等。“四级产品ID”的设计是针对产品的四级分类进行设计的，一级分类用一位数字可表示(x)；二级分类可用两位数字表示(xx)；三级分类可用两位数字表示(xx)；四级分类可用三位数字表示

(xxx)。综上可用四级产品ID采用8位长整型数字表示(x / xx / xx / xxx)。以“椰子”为例：“种植业”编号为1，“果品类”编号为4，“(常绿果树)坚(壳)果”编号为18，“椰子”编号为3，空余的位数用“0”填充，则“椰子”对应的四级分类ID为“10418003”(表1,图6)。

表1 认证农产品分类信息的一些字段及特征

信息名称	信息说明	数据类型	长度	索引
认证产品ID	编号体系	数字型	20	是(无重复)主键
一级分类名称	种植业、畜牧业、渔业	文本型	20	无索引
二级分类名称	粮油、蔬菜等	文本型	20	无索引
三级分类名称	根据标准汇编分类	文本型	20	无索引
四级分类ID	底层分类,不能再分,如胡萝卜、花椰菜等	数字型	20	是(有重复)外键
产品名称	产品具体名称	文本型	20	无索引
别名	产品其他常见名称	文本型	40	无索引

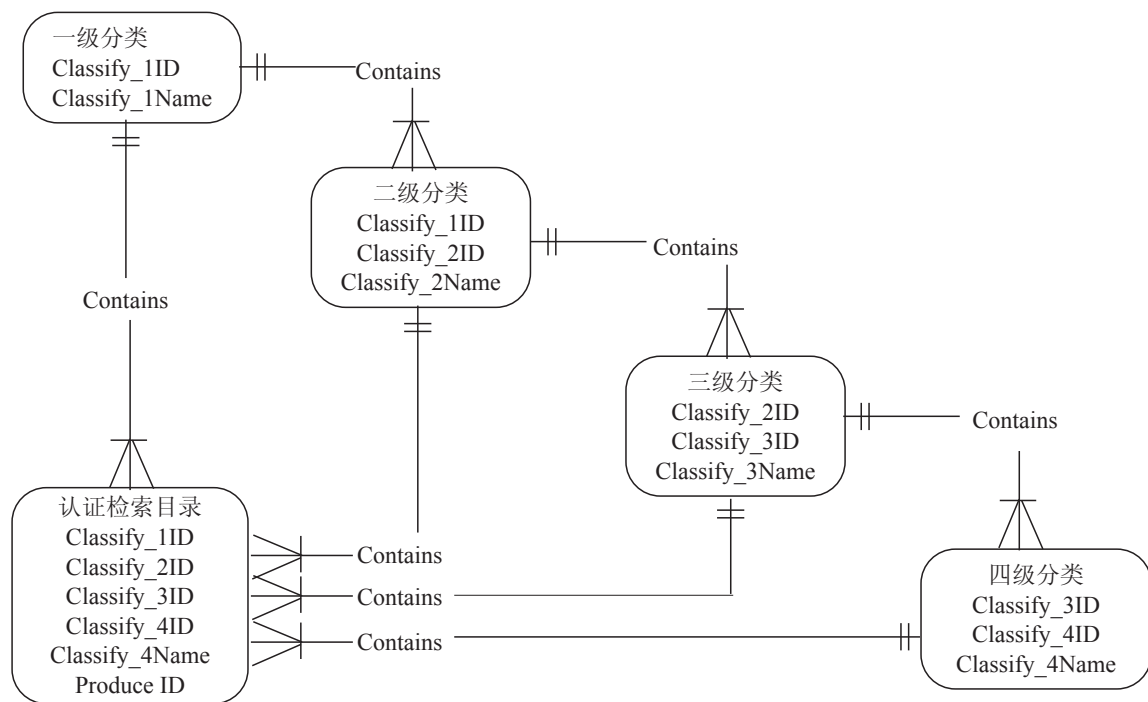


图6 无公害农产品分类目录的实体关联图

5.2 认定的产地信息

该类信息主要针对实体“产地”设计。该部分信息包括已通过认证的无公害农产品产地的基本信息,检查及检测的项目、结果、评价报告及产地检测机构信息等。包括的表有:产地信息表、现场检查报告表、产地环境检测报告表、产地环境检测数据、产地生态环境背景数据表和产地检测机构信息表。该部分信息通过“产地信息表”与其它类信息链接,表中字段“认定产地ID”作为主键(表2,图7)。

5.3 认证的产品信息

该类信息主要针对实体“产品”设计。该部分信息包括已通过认证的无公害农产品的基本信息,检查项目、检测报告及产地检测机构信息等。具体表包括产品信息表、产品检测报告表、产品检测数据表、产品详细信息表、农产品认证信息登记表和产品检测机构信息表。该部分信息通过“产品信息表”与其它类信息链接,表中字段“认证产品ID”作为主键(表3,图8)。

表2 认定的产地信息的一些字段及特征

信息名称	信息说明	数据类型	长度	索引
认定产地ID	编号体系	数字型	20	有(无重复)主键
产地认定证书编号		文本型	15	无索引
产地名称		文本型	20	无索引
产地地址		文本型	50	无索引
生产单位		文本型	50	无索引
证书有效日期		日期/时间	日期型	无索引

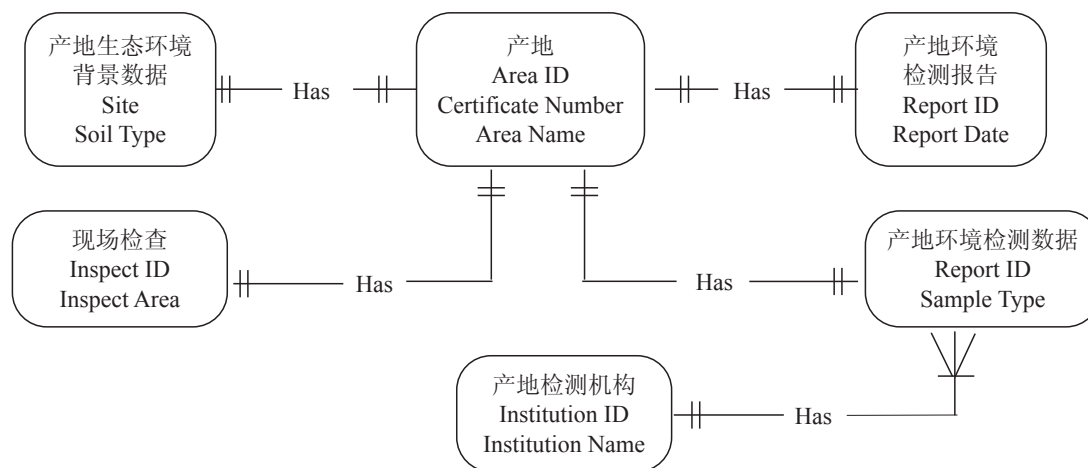


图7 认定产地信息实体关联图

表3 认证的产品信息的一些字段及特征

信息名称	信息说明	数据类型	长度	索引
认证产品ID	编号体系	数字型	20	有(无重复)主键
认定产地ID	编号体系	数字型	20	有(有重复)外键
产品认定证书编号		文本型	15	无索引
产品名称		文本型	20	无索引
有效期限截止日期		日期/时间	50	无索引

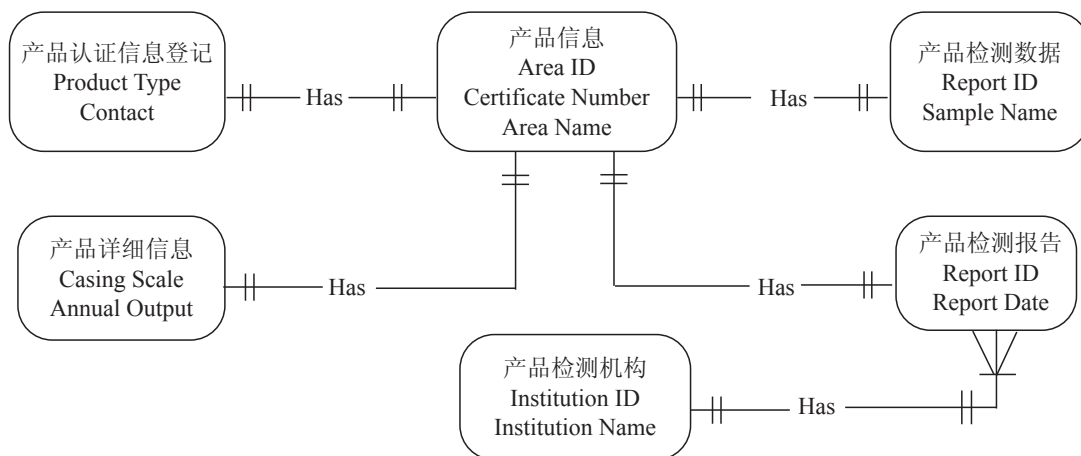


图8 产品信息实体关联图

5.4 申请人申报信息

该类信息主要针对实体“申请人”设计。申请人在申报无公害农产品时首先要提供书面申报材料,针对

申报的材料,分析数据特征设计该类表。本部分包括的表有申请人基本信息表、申报产地基本情况表、申报产品情况表、申报产品农药使用情况表、申报产品标志

表4 申请人申报信息的一些字段及特征

信息名称	信息说明	数据类型	长度	索引
申报ID	编号体系	数字型	20	有(无重复)主键
申请人全称		文本型	50	无索引
申请人类型		文本型	10	无索引
法人代表		文本型	20	无索引
联系电话		文本型	13	无索引
通讯地址		文本型	200	无索引
年总出货量		数字型	单精度型	无索引

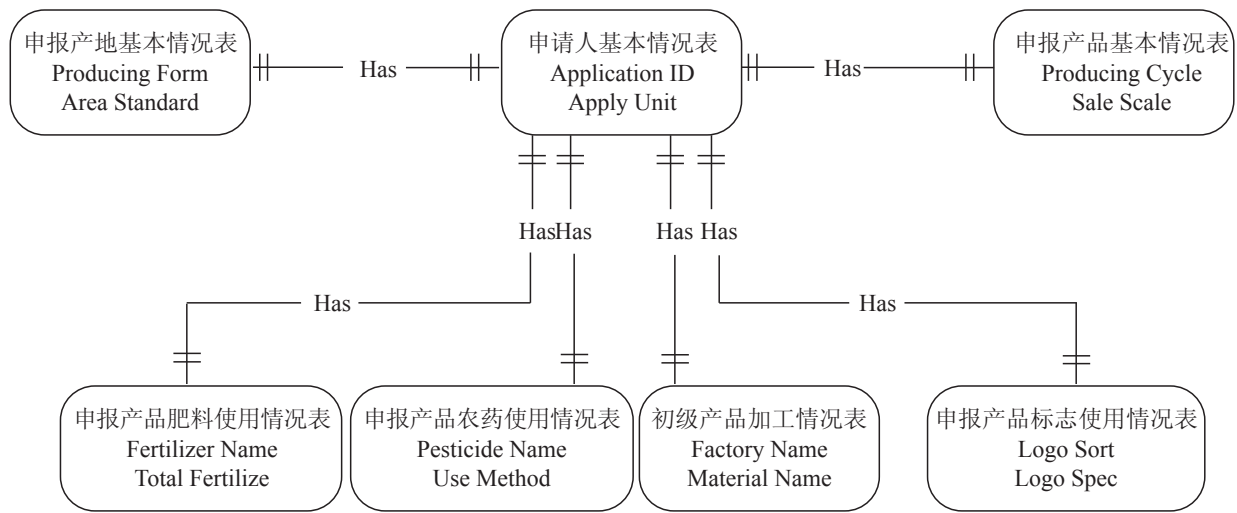


图9 申报信息实体关联图

使用表等信息。设立字段“申报ID”实现该部分表间及与其它类表的链接(表4,图9)。

5.5 现场检查信息

该类信息主要针对实体“现场检查”设计,根据现场检查的要求、具体步骤及内容,该部分的表包括产地

环境背景信息表、产地土壤质量评价信息表、符合性判定信息表、检查员信息表和检查员工作记录表。设立字段“现场检查ID”,实现该部分表间及与其它类表的链接(表5,图10)。

表5 现场检查信息的一些字段及特征

信息名称	信息说明	数据类型	长度	索引
现场检查ID	编号体系	数字型	20	有(无重复)主键
土壤结构	依评分体系分级	数字型	整型	无索引
耕层厚度		数字型	整型	无索引
有机质		数字型	整型	无索引
周围植被	评分体系分级	数字型	整型	无索引
土壤颜色	评分体系分级	数字型	整型	无索引
受控污染源与产地间距离		数字型	整型	无索引

5.6 无公害农产品标准信息

无公害农产品标准是中国无公害农产品认证的依据,包括了无公害农产品术语定义、检验规则、包装、运输和贮存等^[1]。无公害农产品标准信息因农产品种类不同而有所区别,种植业无公害农产品标准信息按性

质和用途可分为以下四类:

(1)无公害农产品质量标准信息。该部分包括感官标准信息表、理化标准信息表、安全标准信息表、其他标准信息表。此类信息是无公害农产品认证过程中的必检信息,为农产品质量检验部门提供

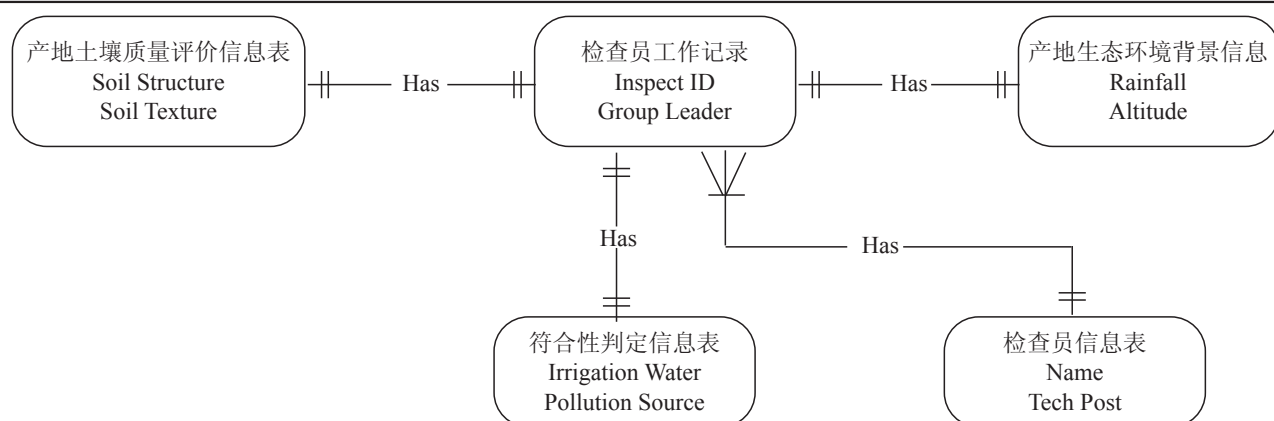


图10 现场检查信息实体关联图

判定依据。

(2)无公害农产品生产规程信息。该部分包括生产技术规程标准信息表、农药安全使用标准信息表、禁用农药标准信息表、病虫害防治标准信息表。此类标准信息主要为生产单位的安全生产提供指导,同时也为检查员进行产地现场检查提供判定依据。

(3)无公害农产品抽样规范信息。该部分只含有无公害农产品抽样规范信息表,此类信息是检测机构为保证检测样本具有代表性而应遵循的抽样方法和规

则,为通则类标准。

(4)无公害农产品产地环境标准信息。农产品产地环境应符合的标准,包括大气环境标准信息表、水环境标准信息表、土壤环境标准信息表。此类信息主要为环境检测机构提供判断依据。

建立标准信息各单位表,索引表以实现表间链接的高效性。标准根据四级分类表与产品信息关联,表6是标准信息表中的部分字段及其特征,图11为无公害农产品标准信息各子表的实体关联。

表6 无公害农产品标准数据信息的一些字段及特征

信息名称	信息说明	数据类型	长度	索引
记录ID	编号体系	数字型	自动编号	有(无重复)
生产规程ID	各产品的质量标准ID	数字型	整型	有(有重复)
农药名称	生产技术规程指标	文本型	20	无索引
剂型	药剂类型	文本型	50	无索引
常用药量	作物常用使用药的量	文本型	50	无索引
最高用药量	每季作物最高使用农药量	文本型	50	无索引
施药方法	农药施用方法	文本型	50	无索引
最多使用次数	每季作物使用药物最多次数	文本型	10	无索引

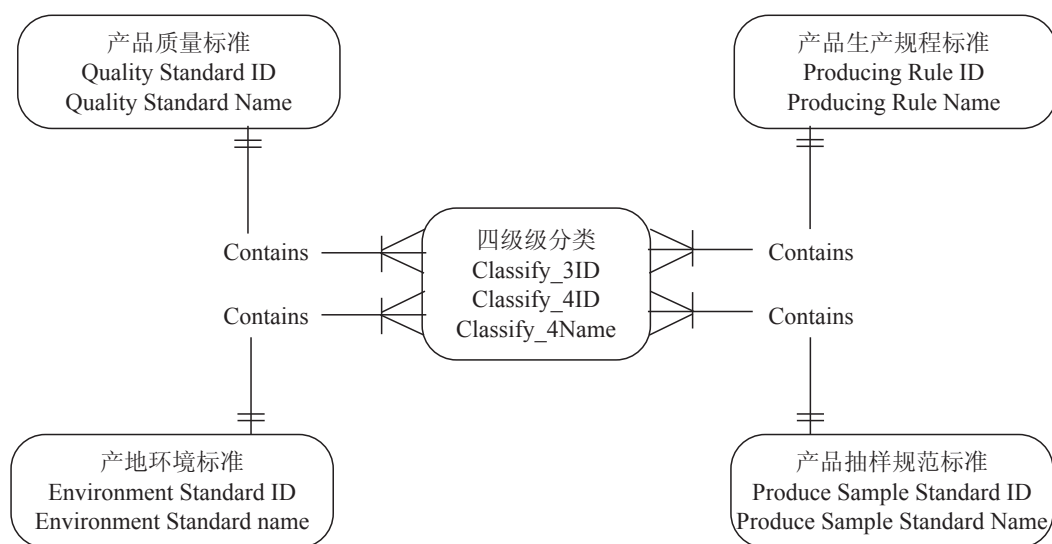


图11 无公害农产品标准信息实体关联图

6 结语

中国无公害农产品的认证正处于高速发展的阶段,但也存在的诸多的限制因素,如认证信息量大、信息传输方式落后、认证程序繁琐等已严重制约了中国无公害农产品的发展。为解决以上问题,笔者从中国无公害农产品认证的现状出发,根据无公害农产品信息的多源性、时效性、连续性、关联性和多类性的数据特征,针对产品认证过程中的产地、产品、申请人、现场检查、标准的五大实体,采用实体-联系的建模方法并依据数据库的设计规则,建立了无公害农产品信息数据库。数据库共包括51个表(含单位表),涉及400多个属性字段,已成功录入无公害农产品、产地、标准等记录共四万多条。但是,由于中国无公害农产品数字化还处于探索阶段,相关的研究较少,借鉴的文献不多,使得研究在数据分析、筛选、组织等方面很可能存在许多不足,笔者提出的数据库结构模型也必然有待进一步完善,但希望拙文能起到一定的启示作用,引发广大学者及相关人员加强该领域的研究,从而促进中国无公害农产品数字化认证快速发展。

参考文献

[1] 刘斌.陕西省农产品质量安全认证发展研究.陕西杨陵:西北农林科技大学,2004.
 [2] 马爱国.无公害农产品管理与技术(第二版).北京:中国农业出版

社.2007:16-36.
 [3] 冯锐,张玉书.辽宁省基本气象资料数据库系统设计与建立.气象科技,2003,31(5):296-299.
 [4] 马友华,胡芹远,转可钦,等.合肥市土壤数据库系统的建立.安徽农学通报,2001,7(1):48-49.
 [5] 肖志勇,董文光,高景红,等.运用数据库技术实现对蔬菜产品“从产地到餐桌”的全程监控.环境管理.2005,(2):48.
 [6] Jaspreet K C, Ahuja, Betty P, Perloff. Quality control procedures for the USDA Food and Nutrient Database for Dietary Studies Nutrient Value[J]. Journal of Food Composition and Analysis. 2008(21): 119-124.
 [7] McMeekin T A, Baranyi J, Bowman J, et al. Information systems in food safety management. International Journal of Food Microbiology. 2006(112):181-194.
 [8] Nesvadba P, Housska M, Wolf W,et al. Database of physical properties of agro-food materials. Journal of Food Engineering. 2004(61):497-503.
 [9] 李翠霞,尹宁园.浅论农业专题数据库的建设.安徽农业科学,2003, 31(4):630-631.
 [10] Michael V, Mannino.数据库设计、应用开发和管理.韩宏志译.北京:清华大学出版社,2007:28-30.
 [11] 农业部农产品质量安全中心.NY5000,无公害农产品标准汇编种植业卷(上、下).北京:农业部农产品质量安全中心,2007.
 [12] Terry Halpin.信息建模与关系数据库.施伯乐,刘国华,陈子军等译.北京:电子工业出版社:4-9.
 [13] 李洪义,程街亮,杨大志,等.浙江省主要农产品安全调查背景数据库的设计与实现.浙江农业学报,2005,17(3):138-142.