



中国农业普查
CHINA AGRICULTURAL CENSUS

第二次全国农业普查 数据处理aPRAS软件 技术手册

国务院第二次全国农业普查领导小组办公室 编



中国统计出版社
China Statistics Press



中国地质大学
北京

中国地质大学北京

数据处理AFRAS软件 技术手册

中国地质大学北京地质研究所地质研究所编 编

地质出版社

第二次全国农业普查数据处理 aPRAS 软件技术手册

国务院第二次全国农业普查领导小组办公室 编

 中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

第二次全国农业普查数据处理 aPRAS 软件技术手册/
国务院第二次全国农业普查领导小组办公室编.
—北京:中国统计出版社,2007.10
ISBN 978-7-5037-5274-2

I. 第…
II. 国…
III. 农业—普查—数据处理—应用软件—中国—手册
IV. F322-39

中国版本图书馆 CIP 数据核字(2007)第 150455 号

第二次全国农业普查数据处理 aPRAS 软件技术手册

作 者/国务院第二次全国农业普查领导小组办公室
责任编辑/许立舫
装帧设计/艺编广告公司
出版发行/中国统计出版社
通信地址/北京市西城区月坛南街 57 号 邮政编码/100826
办公地址/北京市丰台区西三环南路甲 6 号
网 址/www.stats.gov.cn/tjshujia
电 话/邮购(010)63376907 书店(010)68783172
印 刷/河北天普润印刷厂
经 销/新华书店
开 本/890×1240 mm 1/16
字 数/370 千字
印 张/19.5
印 数/1—2200
版 别/2007 年 10 月第 1 版
版 次/2007 年 10 月第 1 次印刷
书 号/ISBN 978-7-5037-5274-2/F·2566
定 价/59.00 元

版权所有。未经许可,本书的任何部分不准以任何方式在世界任何地区
以任何文字翻印、拷贝、仿制或转载。
中国统计版图书,如有印装错误,本社发行部负责调换。

《第二次全国农业普查数据处理 aPRAS 软件技术手册》

编辑人员名单

顾 问：张为民

主 编：文兼武

副 主 编：倪志良 盛来运 赵建华

编 辑：（按姓氏笔划为序）

王文娜	王允铎	王忠良	王 波	王挺漳	卢 涛
叶 涛	叶力强	阳 光	孙 婧	李华丽	李险峰
余新华	何小平	张 劲	张 柳	张建华	张俊逸
房 波	杨成实	杨晓光	荀 策	贾书民	夏雨春
高 珂	高永岗	袁晓勇	梁卫华		

责任编辑：夏雨春 阳 光 张俊逸 卢 涛

前 言

aPRAS 软件是为完成第二次全国农业普查数据处理工作而开发的应用系统，其采用了客户机/服务器（C/S）体系架构。客户端运行在 Windows 平台上，提供业务操作的界面，负责同用户的交互、向服务器端发出请求、接收并处理服务器端的返回结果。服务器端采用 J2EE 架构，在 Windows 平台采用的 PRAS Server 集成了 JBoss 应用服务器；对于 HP-UX (Unix) 操作系统环境，应用部署于 WebLogic 8/9 应用服务器和 Tuxedo 8/9 中间件之上，可访问 MySQL、SQL Server 2000、SQL Server 2005、Oracle 9i、Oracle 10g 多种数据库，负责同客户端进行交互，接收客户端的请求，处理相关的业务逻辑，完成对数据资源的存取。

本书介绍 aPRAS 农业普查数据处理软件的系统设计和实现技术以及在 HP-UX 小型机环境下的安装部署步骤，讲解 aPRAS 软件的运行维护知识和公式语法规则，还介绍 Oracle 数据库以及 SQL Server 数据库管理系统的安装与配置、应用服务器 WebLogic 的安装与配置等内容。

本书共分 10 章：

第 1 章 aPRAS 软件的开发背景，简述了第二次全国农业普查的业务流程。

第 2 章 aPRAS 软件安装运行需要的软件硬件环境。

第 3 章 描述 aPRAS 系统支持的数据库管理系统的安装、部署、备份及恢复技术。涉及的数据库包括 Oracle 9i、Oracle 10g、MS SQL Server 2005。

第 4 章 WebLogic Server 8.1.3 和 WebLogic Server 9.2 的安装及使用。

第 5 章 Tuxedo 8.1 和 Tuxedo 9.1 for HP-UX 的安装及使用。

第 6 章 aPRAS for HP-UX 版软件的安装过程。

第7章 aPRAS 软件的启动与关闭、系统配置管理及运行管理。

第8章 aPRAS 公式系统的使用。

第9章 aPRAS 的体系结构与设计思想。

第10章 aPRAS 系统的数据库设计以及数据库优化技术。

附录 农普基层表列表、MySQL 数据库常见命令及常见问题解答、省级数据上报接收方法。

本书的读者对象：

本书将为数据处理人员、系统管理维护人员、数据库管理员等提供帮助，以支持各级农普办完成第二次全国农业普查数据处理工作。同时，也可各大专院校计算机相关专业的学生提供参考，以了解大型普查等统计调查项目数据处理工作的有关技术。

目 录

第 1 章 aPRAS 开发背景	
1.1 第二次全国农业普查业务概述	(1)
1.2 农普业务流程	(1)
1.3 农普数据处理系统业务模型	(2)
1.4 农普数据处理系统功能要求	(3)
1.5 农普数据处理系统非功能要求	(4)
第 2 章 aPRAS 系统环境要求	
2.1 aPRAS 的版本介绍与适用情况	(6)
2.2 服务器端系统环境要求	(7)
2.2.1 Windows 版系统环境要求	(7)
2.2.2 Unix 版系统环境要求	(7)
2.3 客户端系统环境要求	(10)
2.4 网络要求	(10)
第 3 章 数据库管理系统的安装与配置	
3.1 Oracle 10g for HP-UX 的安装	(11)
3.1.1 简要说明	(11)
3.1.2 安装前准备工作	(11)
3.1.3 安装 Oracle 10g 数据库管理系统	(14)
3.2 Oracle 数据库备份和恢复	(21)
3.2.1 项目背景说明	(21)
3.2.2 备份模式说明	(21)
3.2.3 备份恢复策略与实施步骤	(22)
3.3 Oracle 数据库有用的新特性	(26)
3.3.1 采用可传输表空间方式导入和导出	(26)
3.3.2 Oracle10g 新特性数据泵完成数据导入和导出	(28)

3.3.3	利用 Enterprise Manager 10g 管理数据库	(33)
3.4	SQL Server 2005 for Windows 的数据库安装	(35)
3.4.1	简要说明	(35)
3.4.2	系统环境要求	(35)
3.4.3	数据库软件的安装	(36)
3.4.4	数据库补丁 SP2 的安装	(41)
3.5	SQL Server 2005 数据库的备份与恢复	(41)
3.5.1	备份	(41)
3.5.2	恢复	(47)
第4章 WebLogic Server 的安装与配置		
4.1	安装准备	(50)
4.2	安装 WebLogic Server 8.1.3	(51)
4.2.1	用图形化界面方式安装 WebLogic Server 8.1.3	(51)
4.2.2	安装 WebLogic Server 8.1.3 的 license	(51)
4.3	安装 WebLogic Server 9.2	(52)
4.3.1	用图形化界面方式安装 WebLogic 9.2	(52)
4.3.2	安装 WebLogic Server 9.2 的 license	(59)
4.4	WebLogic Server 的使用	(59)
4.4.1	启动和停止服务器	(59)
4.4.2	使用控制台控制 Web 应用 (以 aPRAS 系统为例)	(60)
4.5	WebLogic Server 参数配置	(62)
第5章 Tuxedo 的安装与配置		
5.1	安装准备	(66)
5.2	Tuxedo 8.1 for HP-UX 的安装	(67)
5.3	Tuxedo 9.1 for HP-UX 的安装	(67)
5.4	Tuxedo 的使用	(73)
5.5	Tuxedo 的参数配置	(77)
第6章 aPRAS 系统安装与配置		
6.1	安装前准备工作	(79)
6.1.1	安装包准备与检查	(79)
6.1.2	aPRAS 安装环境检查	(79)
6.1.3	aPRAS 系统磁盘和文件系统规划	(80)
6.2	安装步骤	(80)
6.2.1	安装文件说明	(80)
6.2.2	创建 apras 用户及设定环境变量	(81)

6.2.3	在 Oracle 数据库上建立 aPRAS 系统库	(82)
6.2.4	安装部署 aPRAS 服务器应用系统	(84)
6.3	安装后配置和系统的启动与关闭	(86)
6.3.1	修改 aPRAS 系统配置文件	(86)
6.3.2	启动和关闭 aPRAS 系统	(86)
第 7 章 aPRAS 系统维护与管理		
7.1	系统的启动与关闭	(89)
7.1.1	Windows 版系统的启动	(89)
7.1.2	Windows 版系统启动过程	(89)
7.1.3	Windows 版系统的关闭	(91)
7.2	系统配置介绍	(91)
7.2.1	服务器端的关键配置项	(91)
7.2.2	不同规模下 Oracle 关键参数的推荐配置	(95)
7.3	系统卸载	(96)
7.3.1	aPRAS Windows 版的卸载	(96)
7.3.2	aPRAS Unix 版的卸载	(97)
7.4	日志的查看	(97)
7.4.1	客户端日志	(97)
7.4.2	Windows 版服务器端日志	(98)
7.4.3	Unix 版服务器端日志	(100)
7.5	系统版本升级和部署	(101)
7.6	aPRAS 和 ePRAS 系统部署在同一台主机时注意事项	(101)
7.7	Web 监控页面的使用	(101)
7.8	数据库软件和数据备份、恢复操作流程	(102)
7.8.1	MySQL 数据库软件 and 数据的备份和恢复	(102)
7.8.2	MS SQL Server 数据库软件 and 数据的备份和恢复	(102)
7.8.3	Oracle 数据库软件 and 数据的备份和恢复	(103)
7.9	应用服务器软件和数据备份、恢复操作流程	(103)
7.10	中间件软件备份、恢复操作流程	(104)
7.10.1	WebLogic 软件的备份和恢复	(104)
7.10.2	Tuxedo 软件的备份和恢复	(104)
第 8 章 aPRAS 的公式语法规范		
8.1	aPRAS 公式系统概述	(105)
8.1.1	aPRAS 公式系统的目的	(105)
8.1.2	aPRAS 公式系统组成	(106)

8.1.3	aPRAS 公式系统的基本概念	(106)
8.2	aPRAS 后台公式说明	(109)
8.2.1	aPRAS 后台公式基本结构	(109)
8.2.2	aPRAS 后台公式组成	(116)
8.3	aPRAS 后台公式语法说明	(126)
8.3.1	代码风格	(126)
8.3.2	基本元素	(128)
8.4	aPRAS 后台公式实例	(130)
8.4.1	审核公式	(130)
8.4.2	汇总公式	(135)
8.5	aPRAS 审核表达式说明	(143)
8.5.1	表达式指标说明	(143)
8.5.2	审核关系表达式组成	(144)
8.5.3	审核关系表达式类型	(148)
8.6	aPRAS 汇总表审核说明	(151)
8.6.1	表数据形式	(151)
8.6.2	支持的审核关系表达式	(152)
8.6.3	汇总表审核关系表达式实例	(152)
8.6.4	汇总表审核关系表达式支持的函数	(153)
8.7	aPRAS 录入控制公式说明	(153)
8.7.1	录入控制公式概述	(153)
8.7.2	录入控制公式格式	(154)
8.7.3	录入控制公式事件类型	(154)
8.7.4	录入控制公式语法	(154)
8.7.5	录入控制公式示例	(159)
第 9 章	aPRAS 系统的设计与实现	
9.1	aPRAS 的体系结构与设计思想	(163)
9.1.1	设计思想	(163)
9.1.2	系统架构	(164)
9.1.3	软件开发方法	(164)
9.1.4	技术路线	(165)
9.1.5	应用系统的结构	(167)
9.2	应用系统的实现技术	(169)
9.2.1	服务端	(169)
9.2.2	客户端	(210)

9.3	关键功能设计与实现——公式引擎	(214)
9.3.1	公式系统机理	(214)
9.3.2	公式与 Python, Jython 的关系	(217)
9.3.3	批量和逐行公式的实现技术	(219)
9.3.4	公式 SQL 语句举例	(221)
9.4	业务实现举例	(227)
9.4.1	对象交互总述	(227)
9.4.2	举例	(228)
9.5	aPRAS 数据接口标准	(233)
9.5.1	文本数据接口的规范	(234)
9.5.2	农普光电系统的接口规范	(238)
9.6	aPRAS 系统导入导出文件列表及数据导入导出实现	(239)
9.7	aPRAS 软件开发及测试工具	(242)
第 10 章	aPRAS 系统的存储设计	
10.1	设计思想	(243)
10.1.1	设计目标	(243)
10.1.2	设计特点	(243)
10.1.3	数据完整性	(243)
10.2	数据库的设计过程	(244)
10.3	命名规范	(246)
10.3.1	数据库存储对象	(246)
10.3.2	数据库表	(248)
10.3.3	数据库表字段	(249)
10.4	可扩展性	(249)
10.5	逻辑模型	(250)
10.5.1	系统库模型	(250)
10.5.2	项目库模型	(251)
10.5.3	动态表生成原理	(255)
10.6	数据字典	(258)
10.6.1	实体清单列表	(258)
10.6.2	数据字典列表	(259)
10.7	数据库差异性	(269)
10.8	数据库业务逻辑	(271)
10.8.1	项目和制度的备份恢复	(272)
10.8.2	数据库并行操作	(273)

10.8.3	SQL Server 2005 和 SQL Server 2000 数据批量插入的差别	(273)
10.8.4	数据库在数据导入时对磁盘空间的要求	(273)
10.9	数据库安全性	(274)
10.10	数据库优化	(275)
10.10.1	分区	(275)
10.10.2	物化视图	(277)
10.10.3	索引的使用	(279)
10.10.4	MySQL 存储引擎	(280)
	附录 A: 第二次全国农业普查基层表列表	(282)
	附录 B: MySQL 数据库常见命令	(283)
	附录 C: aPRAS 常见问题答疑	(286)
	附录 D: 数据上报和接收方法	(293)

第1章 aPRAS 开发背景

aPRAS 数据处理系统是为了完成第二次全国农业普查数据处理工作而开发的软件。它采用现代信息技术,实现第二次全国农业普查数据的录入、编辑、审核、汇总、资料开发等功能,并为后期建立第二次全国农业普查数据库、以及今后搭建新的农村统计调查业务工作平台打下基础。本章介绍 aPRAS 软件的开发背景,即第二次全国农业普查业务流程和功能及非功能要求。

1.1 第二次全国农业普查业务概述

第二次全国农业普查(以下简称农普)是全面掌握我国农业、农村和农民的基本情况,为全面掌握我国农业、农村和农民的基本情况,为国家研究制定经济社会发展规划和科学决策提供依据,为农业生产经营者和社会公众提供统计信息服务。

普查机构由国务院、省、地市、县四级设立的第二次全国农业普查办公室构成。

普查的行业范围是农作物种植业、林业、畜牧业、渔业和农林牧渔服务业。普查的地域范围:农业生产经营者的普查地域范围是中华人民共和国境内的所有农村地域和城镇地域;农村住户和村民委员会的普查地域范围是农村;乡镇人民政府的普查地域范围是所有乡镇。本次普查以村民委员会所辖的地域为普查区,在普查区设立普查小区,以普查小区为最小单位组织入户调查,填写调查表。

普查对象是农村住户、城镇农业生产经营户、农业生产经营单位、村民委员会和乡镇人民政府。

普查的内容为农户人口与劳动力就业状况;农户家庭生活设施和生活条件;农户承包与经营的农业用地;农户农业生产条件和农业生产情况;各种企事业单位的农业生产经营情况;以及乡镇和村的社会经济状况与发展情况。

1.2 农普业务流程

为完成第二次全国农业普查工作,各级普查机构(农业普查办公室)要分工协作,共

同努力。第二次全国农业普查数据处理工作在国家、省、地（市）三级普查机构组织实施。

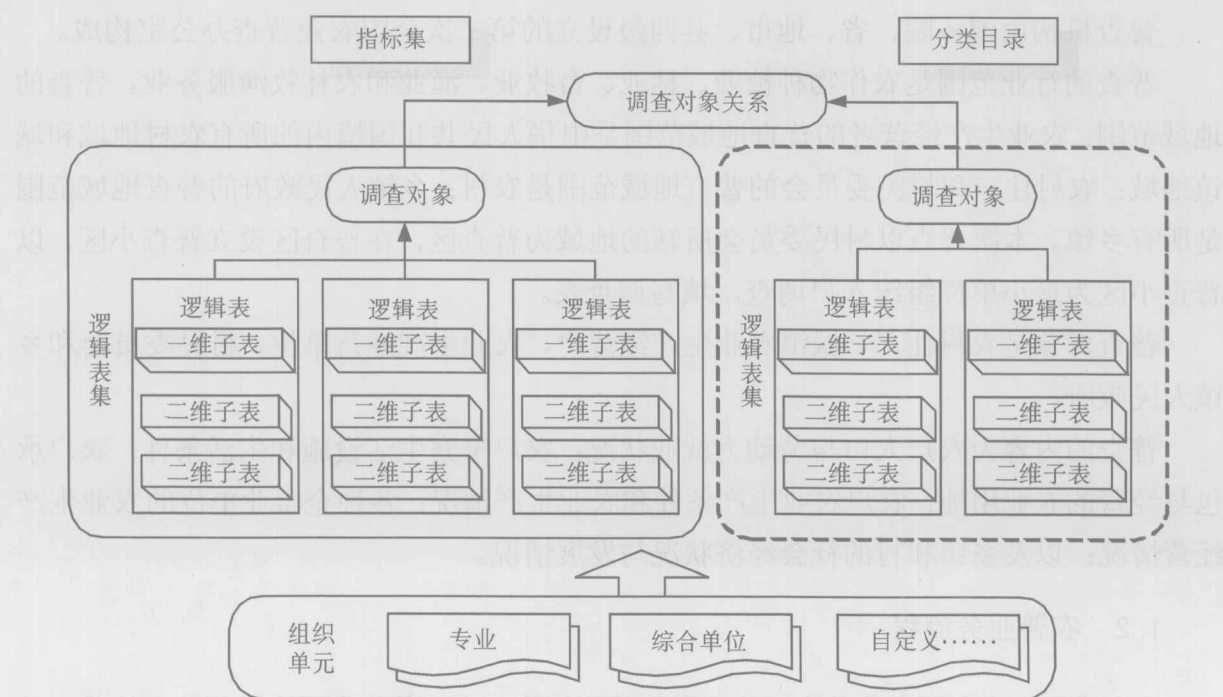
国务院农普办制定第二次全国农业普查方案，并将方案向省级农普办下发。省级农普办按照本省的实际情况，在第二次全国农业普查方案的基础上进行制度扩充。在入户调查结束后，由县级农普办对填写好的普查表进行初审、编码、整理。地市农普办进行光电录入或手工录入等工作，将经 OCR 识别后的数据载入到数据处理程序 aPRAS 中，进行数据编辑与审核，在完成全地市的数据编审工作后，汇总出本地市农业普查综合数据，进行数据质量评估，并将评估审定后的本地市基层数据、综合数据、普查表影像数据一并上报省级农普办数据处理机构。

省级农普办数据处理机构负责本省农业普查的数据处理工作。负责对所辖各地（市）农普办上报的普查数据及资料进行审核、验收，并及时将验收结果反馈给各地市农普办。在验收通过的基础上，汇总生成本省农业普查综合数据，进行数据评估审定后，将数据上报国务院农普办数据处理机构。

国务院农普办数据处理机构负责通过网络或存储介质接收省级上报的农普数据和相关资料，进行审核验收，向各省农普办进行信息反馈，汇总生成全国农业普查综合数据，并向社会公布普查信息等。

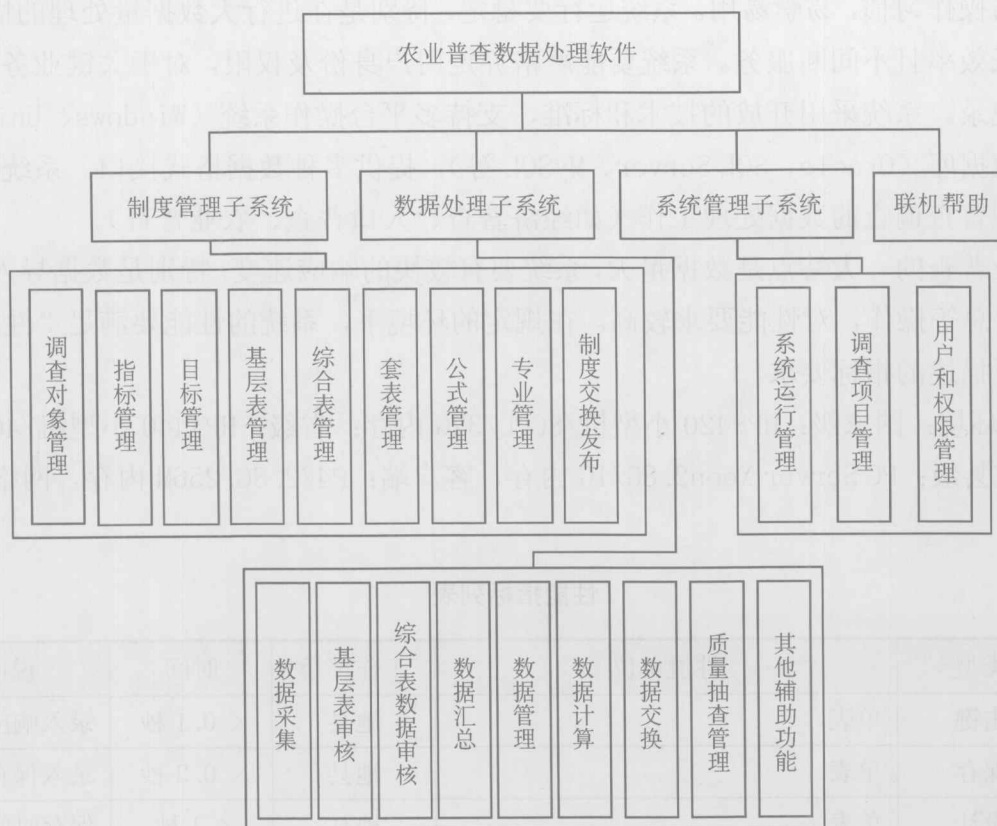
1.3 农普数据处理系统业务模型

通过对农普数据处理工作的业务分析和抽象，形成农普数据处理系统的业务模型，该业务模型是整个软件设计的基础。aPRAS 的业务模型由调查对象类、指标、目录、逻辑表、组织单元几个部分组成，各种业务元素之间存在着一定的关系，调查对象之间有父子关系，乡和村为父子关系；村和小区为父子关系；调查对象与表之间的关系是调查对象可以关联一张或多张基层表。业务模型如图所示。



1.4 农普数据处理系统功能要求

农普数据处理系统需要完成农普制度的定制、数据采集及编辑、用户及权限管理等相关功能。具体分为制度管理、数据处理、系统管理三个子系统。如下图所示。



系统管理子系统：提供农业普查数据处理系统的安全及运行管理功能，主要包括：调查项目的建立、用户及角色管理、权限管理、数据备份和恢复、综合单位管理、系统日志管理等。

制度管理子系统：提供在项目中的制度新建、修改、删除功能；制度数据导入、导出功能；制度内容设计、制度扩充、制度交换、制度版本管理等功能。特别是需要设计灵活的、符合统计系统普查数据处理使用习惯的公式描述智能向导，以便于用户在进行二次开发时，能够进行录入控制、审核、计算、汇总公式的编写等操作。还支持制度包的反复下发和上报。要求软件能够支持农业普查制度中各种复杂关系的正确描述和处理。

数据处理子系统：提供数据采集、数据审核、数据修改、数据汇总、数据查询、数据管理、数据上报与接收、审核结果输出、查询结果及汇总结果输出、质量抽查等功能。特别要提供灵活、多样、高效的外部数据导入导出接口。要支持上级统计机构向下级统计机构的综合数据和基层数据的反馈与接收工作。要提供集成的图像查看与数据编辑界面，支持数据编辑时对光电扫描图像的调用。

1.5 农普数据处理系统非功能要求

除了上述功能性要求外，还要求农业普查数据处理系统能灵活部署在县、地（市）、省、国家四级统计系统的多种软硬件环境下。软件界面友好、实用，符合统计行业用户的术语与操作习惯，易学易用。系统运行要稳定，特别是在进行大数据量处理的情况下，要不降低效率且不间断服务。系统要能严格界定用户身份及权限，对于关键业务操作要有日志记录。系统采用开放的技术和标准，支持多平台操作系统（Windows、Unix 等）及多种数据库（Oracle、SQL Server、MySQL 等），提供多种数据格式接口。系统要适应多种大型普查调查的数据处理工作（如经济普查、人口普查、农业普查）。

农业普查的一大特点是数据量大，系统要有较快的响应速度，特别是数据导入导出、审核、汇总等操作，对性能要求较高。在规定的环境下，系统的性能要满足“性能指标列表”所描述的指标要求。

规定环境：国家级：HP8420 小型机/8CPU/32G 内存；省级：HP4440 小型机/4CPU/8G 内存；地县级：PC Server Xeon2.8G/1G 内存，客户端：P4/2.8G/256M 内存，网络带宽：100M。

性能指标列表

操作类型	性能单位	工作环境	时间	说明
录入击键	单表	地县	< 0.1 秒	录入响应时间
录入保存	单表	地县	< 0.3 秒	录入保存时间
表样设计	单表	地县	< 1 秒	保存时间
界面显示	单表	地县	< 0.2 秒	显示时间
表内审核	1 百万户*50 条审核关系	地县	< 3 分钟	错误率 1%
表间审核	1 百万户*10 条审核关系	地县	< 5 分钟	错误率 1%
表内审核	1 千万户*50 条审核关系*1 个公式	省	< 15 分钟	错误率 1‰
	1 亿户*50 条审核关系*1 个公式	国家	< 45 分钟	错误率 0.1‰
表间审核	1 千万户*10 条审核关系*1 个公式	省	< 20 分钟	错误率 1‰
	1 亿户*10 条审核关系*1 个公式	国家	< 60 分钟	错误率 0.1‰
汇 总	1 百万户*10 综合表（10 个公式）	地县	< 5 分钟	10 个汇总表， 每个汇总表有 30 行 10 列
	1 千万户*10 综合表（10 个公式）	省	< 20 分钟	10 个汇总表， 每个汇总表有 30 行 10 列
	1 亿户*10 综合表（10 个公式）	国家	< 120 分钟	
数据导出	1 百万户的全部数据	地县	< 5 分钟	
	1 千万户的全部数据	省	< 30 分钟	
	1 亿户的全部数据	国家	< 80 分钟	