

doi: 10.7621/cjarrp.1005-9121.20190821

· 资源利用 ·

# 近 20 年中国耕地数量变化趋势及其驱动因子分析<sup>\*</sup>

王静怡<sup>1</sup>, 李晓明<sup>2\*</sup>

(1. 西安石油大学计算机学院, 陕西西安 710065; 2. 陕西省土地工程建设集团有限责任公司, 西安 710075)

**摘要** [目的] 研究“二调”前后近 20 年我国耕地面积的变化趋势及其影响因子, 为促进耕地保护和土地资源的高效集约节约利用提供决策支撑。[方法] 基于 1996—2016 年耕地数量的原始统计数据, 利用数据挖掘技术选用 ARIMA 模型, 基于两次全国土地调查时间段耕地数量数据进行相互预测和反推, 即利用 1996—2008 年耕地数量数据预测 2009 年耕地数量, 利用 2009—2016 年数据反推 2008 年耕地数量, 从而实现数据的校正和比值归一化的目的, 进一步研究近 20 年内我国耕地数量的变化趋势, 并通过相关性分析研究遴选出其驱动因子。[结果] 趋势研究结果表明, 近 20 年内我国耕地数量呈现逐年递减的趋势, 其中 2004 年之前递减速度较快, 之后递减速度趋缓。相关性分析结果显示, 第一产值增加值占比、年末总人口数, 以及城镇人口数是影响耕地数量变化的重要驱动因子, 其相关系数分别为 0.959, -0.918, -0.896。[结论] 从研究结果可以推论, 我国在社会经济发展中的产业结构布局, 以及城镇化建设进程对耕地数量变化产生了重要的影响。

**关键词** 数据挖掘 ARIMA 模型 耕地数量 变化趋势 驱动因子

**中图分类号**: P96 **文献标识码**: A **文章编号**: 1005-9121[2019]08171-06

## 0 引言

我国是一个农业大国, 耕地资源直接关系到国家粮食安全问题, 是确保“中国人的饭碗牢牢端在自己手中”的重要基础。同时, 在加快推进我国社会主义现代化建设进程中, 经济建设对土地的旺盛需求也对耕地保护带来了极大的挑战<sup>[1]</sup>。对于我国耕地的变化趋势及其影响驱动力的研究, 一直都是一个热点问题, 甚至有专家学者通过史册文献资料的查证对过去 300 年我国部分省区耕地资源数量变化及驱动因素进行了研究<sup>[2]</sup>。

目前, 对于我国耕地变化趋势的研究, 多基于对统计数据进行分析, 从空间尺度上可以分为国家尺度和区域尺度的研究。其中, 对于国家尺度上的研究, 由于统计标准、技术手段等多种因素的限制, 2008—2009 年数据发生明显的跳跃式变化, 所以绝大部分的研究均处于 2008 年之前的部分时间段内, 如毕于运等<sup>[3]</sup>通过对我国建国至 1996 年实有耕地面积的收集整理研究分析了其变化分析; 李秀彬<sup>[4]</sup>对 1999 年我国前 20 年耕地面积变化的总体趋势、空间特征和驱动因子进行了研究; 曹银贵<sup>[5]</sup>研究分析了 1996—2005 年全国耕地变化的情况及其影响的主要因子; 黄忠华、吴次芳等利用 1978—2007 年我国耕地、经济增长、城镇化、地方财政收入数据分析了社会经济因素对耕地变化的影响<sup>[6]</sup>。以上研究均将时间局限在 2008 年之前, 而回避了 2008—2009 年耕地统计数据跳跃的现象, 这不利于科学研究我国耕地的变化趋势和指导耕地保护工作。对于区域尺度, 如省级, 甚至市县级尺度, 由于统计基数较小, 受统计方法、技术标准等因素的影响不大, 几乎各个省份区域都有专家对其耕地变化和驱动因子进行了大量研究, 如郭丽英等<sup>[7]</sup>

收稿日期: 2018-03-26

作者简介: 王静怡 (1984—), 女, 陕西榆林人, 博士、讲师。研究方向: 资源利用和数据科学

\*通讯作者: 李晓明 (1983—), 男, 山东烟台人, 博士、高级工程师, 硕士生导师。研究方向: 土地资源利用和遥感信息。Email: xmlisde@126.com

\*资助项目: 陕西省教育厅科研计划项目“无线 MARC 系统中联合网络—信道编码的优化设计研究”(18JK0626)

分析了1996—2008年我国沿海地区耕地变化的基本态势、时空特征及其主导成因；吴霞等<sup>[8]</sup>选取1997—2014年宁夏耕地面积数据分析耕地面积变化趋势及其驱动因子；武江民等<sup>[9]</sup>专门针对甘肃兰州市研究其耕地动态变化与驱动力的定量关系。还有学者则基于国家、区域、省域、市县不同尺度研究了中国耕地变化的驱动力<sup>[10]</sup>；

20世纪80年代以来，随着遥感等3S技术的发展应用，诸多学者开始不仅仅局限于统计数据的应用，开始借助遥感信息技术对耕地变化进行研究。如张国平、刘纪元等<sup>[11]</sup>利用中国资源环境数据库及遥感影像对20世纪80年代末至2000年的中国耕地时空变化进行了分析；田光进、庄大方等<sup>[12]</sup>通过对遥感图像解译研究80年代末与90年代末10年间耕地时空动态变化特征；刘旭华，王劲峰等使用中国1987—1989年和2000年两期土地利用遥感调查数据，探索定量化研究国家尺度土地利用变化的驱动力<sup>[13, 14]</sup>；许丽丽等<sup>[15]</sup>基于中分辨率卫星遥感数据，深入分析了2000—2010年中国耕地变化及其空间差异，并对中国耕地占补平衡政策的实施效果进行了评估；李景刚等<sup>[16]</sup>选取1983—2001年NOAA等遥感数据分析了北方13省的耕地变化与驱动力。遥感信息技术的应用开拓了耕地变化研究的新途径，但是由于遥感解译的精度仍有待于进一步提高，所以一定程度上制约了对于耕地变化的研究。

多来以来，我国对于耕地数量的数据统计有多种来源，其中以国土资源部的统计数据被认为最权威<sup>[17]</sup>。由于遥感影像解译精度，与实地调查统计相比仍有一定差距，所以大部分学者仍采用统计数据进行研究，而且大多避开了数据有明显变化的过渡年份。目前仅有极少数学者对过渡年份前后时间段的统计数据变化一并进行了相关的分析<sup>[17, 18]</sup>。针对两次全国土地调查的统计数据的跳跃式变化，文章研究提出了基于数据挖掘技术研究分析两个时间段的数量关系，以更科学地反映耕地数量的变化趋势，同时基于校正数值研究影响耕地变化的驱动因子，这在目前的研究中尚未见到，对于科学反映我国耕地变化趋势具有积极的意义。

## 1 数据与方法

### 1.1 数据获取

#### 1.1.1 全国耕地数量数据

我国耕地数量，尤其是历史时期的耕地数量，长期以来难以确定。国家土地管理部门从1986年开始，每年统计并公布我国耕地的增减情况，虽然还有其他职能部门的统计数据，但仍以国土资源部的数据最为权威<sup>[17]</sup>。为准确掌握全国耕地数量，1984—1996年国务院安排开展了全国第一次大规模的土地利用现状调查，数据成果统计到了1996年10月31日，结果表明，1996年全国耕地面积约为1.300亿 $\text{hm}^2$ （19.51亿亩）<sup>[18]</sup>，这一统计数据被认为具有最高可信度<sup>[19]</sup>。随后，1996—2008年耕地面积数据均可直接从统计公报中查得。

为进一步准确掌握耕地土地基础数据，国家于2007年7月1日起，启动了第二次全国土地调查，数据统计时间节点为2009年12月31日，统计结果显示2009年我国耕地面积为1.354亿 $\text{hm}^2$ （20.31亿亩）<sup>[18]</sup>。随后的每年都有二调变更调查，2007—2016年每年的全国耕地面积都可以从国家土地调查成果服务平台直接查得。由于第二次全国土地调查采用的技术方法的改进等原因，2008年与2009年的耕地数量存在“跳跃”现象，为科学分析耕地的变化趋势，这一现象问题在后文中进行了着重研究处理。

#### 1.1.2 驱动因子数据

土地与人类社会的经济发展密不可分，全国耕地变化是在地理背景制约下、社会经济条件、灾害等自然条件共同驱动下发生的<sup>[1]</sup>，为探索研究耕地变化的驱动因子，该文主要选取了3方面共17个因子指标做遴选分析：一方面是反映全国经济发展水平的综合类指标，包括国民总收入、国内生产总值、第一产业增加值及比例、第二产业增加值及比例、第三产业增加值及比例、人均国内生产总值、工业增加值、建筑业增加值、农业总产值；另一方面是反映社会城镇化水平的指标，包括年末总人口、城镇人口、乡村人口；还有第三方面是反映自然灾害的指标，包括受灾面积和粮食产量<sup>[20]</sup>。上述因子指标近20年的数据均

直接从国家统计局网站查得。

## 1.2 研究方法

该文研究通过对近 20 年全国耕地数量数据的分析处理,研究其变化趋势,进而通过与选取的因子进行相关性分析,研究其变化的驱动因子。其中,在对 1996—2016 年耕地面积的数据分析中,针对 2008—2009 年数据的跳跃现象,通过数据挖掘,一方面利用 1996—2008 年数据预测 2009 年数据,另一方面利用 2009—2016 年数据反推 2008 年数据,基于上述两种相互预测或反推得出的数值,分析 2008 年和 2009 年耕地数量的数据比例关系,进而对 1996—2016 年耕地面积数据进行比值归一化,以探索研究耕地面积的变化趋势。

该文在数据挖掘预测中,通过遴选对比最终采用的为 ARIMA 模型,即自回归积分滑动平均模型 (Autoregressive Integrated Moving Average Model),是由博克思 (Box) 和詹金斯 (Jenkins) 提出的著名时间序列预测方法,主要是指将非平稳时间序列转化为平稳时间序列,然后将因变量仅对它的滞后值以及随机误差项的现值和滞后值进行回归所建立的模型,适用于该文的研究。

## 2 近 20 年耕地数量变化趋势分析

### 2.1 基于耕地数量原始统计数据的变化趋势及分析

通过对收集整理 1996—2016 年耕地数量数据分析,如图 1,可以看出,1996—2008 年全国耕地面积逐年减少,其中以 2000—2004 年递减速度最快;2008—2009 年耕地面积数据出现跳跃,由 1.217 亿  $\text{hm}^2$  (18.26 亿亩) 增至 1.354 亿  $\text{hm}^2$  (20.31 亿亩);2009—2016 年耕地面积递减速度非常缓慢。

1996 年国家完成了第一次大规模的土地利用现状调查;2009 年第二次全国土地利用现状调查完成。由此可见,2008—2009 年耕地面积数据的跳跃正是一调与二调数据的分界年份。全国耕地面积是否真的增多了约 0.14 亿  $\text{hm}^2$  (2 亿亩),经文献查阅分析<sup>[17, 18, 21]</sup>,主要原因在于二调采用了更为先进的技术手段,将以往少报的耕地面积统计入库,进一步提高了数据的准确度。所以,近 20 年全国耕地数量的统计数据是基于两次不同的技术水平,可直接分别研究各自时段内的变化趋势,但如果要研究连续 20 年时间段内的变化趋势,就须对两个时间段的数据进行校正,使之具有可对比性。

### 2.2 基于数据挖掘的耕地数量校正

数据挖掘技术在时间序列上具有强有力的预测功能,该文首先利用 1996—2008 年耕地数量数据,预测 2009 年的耕地数量,由此得出 2009 年与 2008 年耕地数量的比值关系  $\alpha$ ;然后利用 2009—2016 年耕地数量数据,反推 2008 年的耕地数量,由此得出 2009 年与 2008 年耕地数量的比值关系  $\beta$ ;为更为科学地反映 2009 年与 2008 年耕地数量的比值关系,取利用正向预测与反向反推两种方法得出的  $\alpha$  与  $\beta$  的算数平均值  $\rho$  作为实际的数量比值关系。

该文研究在预测过程中,利用 SPSS 软件,通过比较分析,在正向预测 2009 年数值和反向反推 2008

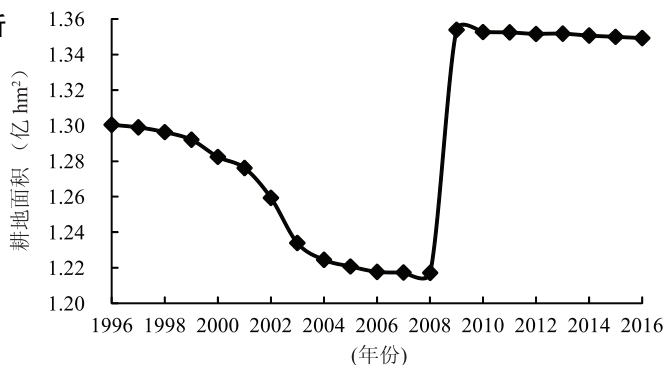


图 1 1996—2016 年基于原始统计数据的耕地数量变化趋势

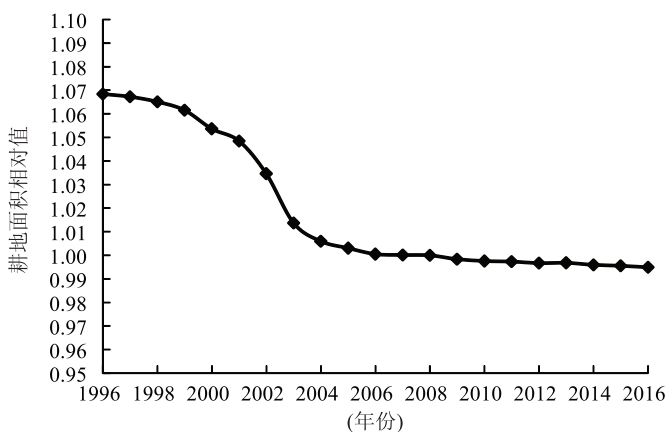


图 2 1996—2016 年基于校正后相对数据的耕地变化趋势

年数值中均选用 ARIMA 模型,其中正向预测选用的模型为 ARIMA (1, 1, 2) ( $r^2 = 0.975^{**}$ ),反向反推选用的模型为 ARIMA (1, 1, 4) ( $r^2 = 0.966^{**}$ ),两者均在 0.01 水平呈现显著性。由此,基于 1996—2008 年数据预测得出的 2009 年耕地数量为 1.213 387 亿  $\text{hm}^2$  (18.200 8 亿亩),基于 2009—2016 年数据反推得出的 2008 年耕地数量为 1.353 987 亿  $\text{hm}^2$  (20.309 8 亿亩),进一步可得出  $\alpha = 996\ 899.887\ \text{E} - 6$ ,  $\beta = 999\ 895.617\ \text{E} - 6$ ,则  $\rho = 998\ 397.752\ \text{E} - 6$ ,由此推测 2009 年与 2008 年耕地数量的比值为 998 397.75 2  $\text{E} - 6$ ,即可为两个时间段数据的校正系数。

### 2.3 基于耕地数量校正相对值的变化趋势分析

为进一步客观的反映近 20 年的耕地数量变化趋势,对全部年份数据进行比值归一化及系数校正,在前文研究基础上,首先对 1996—2008 年全部原始耕地数量数据以 2008 年耕地数量为基数计算得到的比值即为其耕地数量相对值,对 2009—2016 年全部原始耕地数量数据以 2009 年耕地数量为基数计算得到的比值,再乘以校正系数  $\rho$ ,即为其相应年份的耕地数量相对值。由此可以得出 1996—2016 年耕地数量的变化趋势,如图 2 所示,可以看出,近 20 年全国耕地数量是呈现递减趋势的,其中 2004 年以前,递减速度较快,2004 年以后耕地数量递减速度较缓,尤其是 2006 年以后耕地数量递减速度趋于稳定。

## 3 耕地数量变化驱动因子分析

该文研究共收集了 1996—2016 年共 17 个因子指标,首先进行描述性特征统计,如表 1 所示;然后,通过分析耕地数量与各因子指标的相关性(表 2),进行对比分析,以此遴选出耕地数量变化的驱动因子。

表 1 1996—2016 年各因子指标的描述性统计特征

因子	最小值	最大值	平均值	标准差
国民总收入(亿元)	70 780	741 140	304 446.950	225 956.601
国内生产总值(亿元)	71 814	744 127	305 718.670	226 519.843
第一产业增加值(亿元)	13 878	63 671	31 240.710	17 742.619
第一产业增加值占比(%)	8.556	19.325	12.114	3.280
第二产业增加值(亿元)	33 828	296 236	135 629.240	94 949.577
第二产业增加值占比(%)	39.810	47.559	45.327	1.999
第三产业增加值(亿元)	24 107	384 220	138 848.860	114 605.836
第三产业增加值占比(%)	33.569	51.634	42.560	4.472
人均国内生产总值(亿元)	5 898	53 980	22 872	16 332.770
工业增加值(亿元)	29 530	247 860	116 532.760	79 508.293
建筑业增加值(亿元)	4 393	49 522	19 623	15 992.727
年末总人口(万人)	122 389	138 271	131 047.710	4 702.171
城镇人口(万人)	37 304	79 298	58 413.240	12 989.267
乡村人口(万人)	58 973	85 085	72 634.480	8 313.397
农业总产值(亿元)	13 540	59 288	28 842.620	16 570.347
粮食产量(万 t)	43 070	62 144	52 326.480	5 812.061
受灾面积(万 $\text{hm}^2$ )	2 177	5 468.8	4 100.948 0	1 070.630

由表 2 可以看出,基于原始统计的耕地面积研究其与各因子指标的相关性,由于数据的跳跃不足以反映耕地变化的实际趋势,与各因子指标的相关性均不高。基于校正后的耕地数量的相对值,研究其与各因子指标的相关性,可以发现,与第一产业增加值占比的相关性最高,相关系数高达 0.959;与年末总人口的相关性次之,为负相关,系数为 -0.918,城镇人口和乡村人口与之的相关性也较高,即第一产业增加值占比,总人口以及城镇人口数量与耕地数量的变化相关性较好,是重要的驱动因子。由此,可以推论出国家产业结构布局调整、我国城镇化建设水平是影响我国耕地数量变化的重要因素条件。耕地作为第一产

表 2 耕地面积与各因子的相关系数矩阵

原始统计耕地面积 校正后耕地面积相对值			原始统计耕地面积 校正后耕地面积相对值		
原始统计耕地面积	1	-0.142	人均国内生产总值	0.651	-0.749
校正后耕地面积相对值	-0.142	1	工业增加值	0.638	-0.784
国民总收入	0.649	-0.746	建筑业增加值	0.688	-0.716
国内生产总值	0.652	-0.745	年末总人口	0.433	-0.918
第一产业增加值	0.663	-0.736	城镇人口	0.488	-0.896
第一产业增加值占比	-0.274	0.959	乡村人口	-0.518	0.88
第二产业增加值	0.647	-0.774	农业总产值	0.686	-0.709
第二产业增加值占比	-0.519	0.259	粮食产量	0.752	-0.524
第三产业增加值	0.649	-0.717	受灾面积	-0.514	0.708

业的最重要的载体之一,与第一产业增加值密切正相关,我国产业结构中对于第一产业布局要求会直接诱使我国耕地的变化。而城镇化建设过程中,一方面由于城镇建设会直接占用大量的耕地,另一方面,城镇化建设中大量的农民人口直接进入城镇生活,或者进城务工而导致耕地撂荒,所以城镇化建设水平与耕地变化呈现密切的负相关。

#### 4 结论与讨论

土地本身既有自然属性,又有社会属性,对于耕地数量的统计长期以来受社会、经济、政策等诸多因素的影响,导致难以对其进行长时间序列的趋势研究。该文研究的时间序列跨度全国两次土地调查的统计数据,两个时间阶段的数据具有明显的跳跃性,但并不代表 2009 年耕地数量真的比 2008 年大幅度增长。究其原因,主要是由于第二次全国土地调查中应用遥感等新方法、统一了标准,提高了调查的技术水平和准确度,使原来因少报、漏报、瞒报等各种原因未被统计到的耕地重新统计入库,所以要研究跨两个阶段的整体变化趋势,就必须研究建立其数值量化关系。该文通过数据挖掘,并对耕地数量进行校正和比值归一化,研究结果比较准确地反映了我国耕地变化趋势的实际情况。

研究表明,近 20 年,我国耕地数量呈现递减趋势,其中 2004 年以前递减速度较快,2004 年以后递减速度趋缓,我国产业结构布局方面,主要是第一产业增加值占比,以及我国人口总数方面,主要是城镇化人口数量是影响变化的主要驱动因子。由于 1999 年后开始生态退耕导致耕地数量迅速减少<sup>[17, 22]</sup>,2004 年国家《土地管理法》修订后强调要加强耕地保护,实行占补平衡,对我国耕地的数量变化产生了积极作用。该文仅对近 20 年我国耕地数量的变化趋势进行了分析;现在国家加强耕地保护,不仅以数量为基础,更关注质量,以产能为核心,实现数量、质量和生态的三位一体保护<sup>[23]</sup>,下一步将对全国耕地质量的变化趋势进一步分析研究。同时,该文研究通过相关性分析研究提出了主要驱动因子,下一步还将进一步对驱动因子的贡献率进行量化分析研究。

该文研究仅对 1996—2016 年耕地数量的变化趋势进行了分析研究,由于对诸多指标因子数据统计的完整性需求,未对更长时间的耕地变化趋势进行研究。下一步,将在对统计分析的基础上,通过对遥感影像的数据挖掘,深入研究更长时间序列的变化趋势。

#### 参考文献

- [1] 赵其国,周生路,吴绍华,等. 中国耕地资源变化及其可持续利用与保护对策. 土壤学报, 2006, 43 (4): 662-672.
- [2] 葛全胜,戴君虎,何凡能,等. 过去 300 年中国部分省区耕地资源数量变化及驱动因素分析. 自然科学进展, 2003, 13 (8): 825-832.
- [3] 毕于运,郑振源. 建国以来中国实有耕地面积增减变化分析. 资源科学, 2000, 22 (2): 8-12.
- [4] 李秀彬. 中国近 20 年来耕地面积的变化及其政策启示. 自然资源学报, 1999, 14 (4): 329-333.

- [5] 曹银贵, 袁春, 周伟, 等. 中国耕地变化的驱动因子及其省域差异分析. 中国土地科学, 2008, 22 (2): 17-22.
- [6] 黄忠华, 吴次芳, 杜雪君. 我国耕地变化与社会经济因素的实证分析. 自然资源学报, 2009, 24 (2): 192-199.
- [7] 郭丽英, 王道龙, 王介勇. 中国沿海地区耕地变化及其成因分析. 中国农业资源与区划, 2012, 33 (1): 6-10.
- [8] 吴霞, 王世荣, 尚红莺, 等. 宁夏近18年来耕地面积动态变化及驱动力分析. 中国农业资源与区划, 2017, 38 (8): 98-104.
- [9] 武江民, 赵学茂, 党国锋. 甘肃兰州市耕地动态变化与驱动力关系定量研究. 干旱区资源与环境, 2010, 24 (12): 33-38.
- [10] 于天, 曹银贵, 许婧雪. 基于不同尺度的中国耕地变化驱动力研究进展. 中国农学通报, 2016, 32 (24): 194-198.
- [11] 张国平, 刘纪远, 张增祥. 近10年来中国耕地资源的时空变化分析. 地理学报, 2003, 58 (3): 323-332.
- [12] 田光进, 庄大方, 刘明亮. 近10年中国耕地资源时空变化特征. 地球科学进展, 2003, 18 (1): 30-36.
- [13] 刘旭华, 王劲峰, 刘纪远, 等. 国家尺度耕地变化驱动力的定量分析方法. 农业工程学报, 2005, 21 (4): 56-60.
- [14] 刘旭华, 王劲峰, 刘明亮, 等. 中国耕地变化驱动力分区研究. 中国科学D辑地球科学, 2005, 35 (11): 1087-1095.
- [15] 许丽丽, 李宝林, 袁焯城, 等. 2000—2010年中国耕地变化与耕地占补平衡政策效果分析. 资源科学, 2015, 37 (8): 1543-1551.
- [16] 李景刚, 何春阳, 史培军, 等. 近20年中国北方13省的耕地变化与驱动力. 地理学报, 2004, 59 (2): 274-282.
- [17] 封志明, 刘宝勤, 杨艳昭. 中国耕地资源数量变化的趋势分析与数据重建: 1949—2003. 自然资源学报, 2005, 20 (1): 35-43.
- [18] 谭永忠, 何巨, 岳文泽, 等. 全国第二次土地调查前后中国耕地面积变化的空间格局. 自然资源学报, 2017, 32 (2): 186-197.
- [19] Lin G C S, HO S P S. China's land resources and land-use change: insights from the 1996 land survey. land Use Policy, 2003, 20 (2): 87-107.
- [20] 傅泽强, 蔡运龙, 杨友孝, 等. 中国粮食安全与耕地资源变化的相关分析. 自然资源学报, 2001, 16 (4): 313-319.
- [21] 张雅琼, 杨子生. 我国两次土地资源调查的对比分析. 国土资源科技管理, 2013, 30 (2): 100-104.
- [22] 徐宪立, 蔡玉梅, 张科利, 等. 耕地资源动态变化及其影响因素分析. 中国人口·资源与环境, 2005, 15 (3): 75-79.
- [23] 谭永忠, 吴次芳, 王庆日, 等. “耕地总量动态平衡”政策驱动下中国的耕地变化及其生态环境效应. 自然资源学报, 2005, 20 (5): 727-734.

## RESEARCH ON THE CHANGE TREND OF FARMLAND QUANTITY IN CHINA FOR RECENT 20 YEARS AND ITS DRIVING FACTORS \*

Wang Jingyi<sup>1</sup>, Li Xiaoming<sup>2\*</sup>

(1. Xi'an Shiyou University, Xi'an, Shaanxi 710065, China;

2. Shaanxi Provincial Land Engineering Construction Group, Xi'an, Shaanxi 710075, China)

**Abstract** The research aims to research the changing trend of farmland quantity for recent 20 years before and after the second national land use survey. It could provide decision support for farmland protection and efficient use of land resources. The original data of farmland area from 1996 to 2016 were collected, data mining technology was introduced, and ARIMA model was chosen to predict the farmland quantity each other based on the farmland area of the twice land use survey, which meant to predict the farmland quantity of 2009 with the farmland quantity of the years from 1996 to 2008, and to predict the farmland quantity of 2008 inversely with the farmland quantity of the years from 2009 to 2016, so the original farmland area data could be corrected and normalized. Then the changing trend of farmland quantity for recent 20 years was studied furtherly and correlation analysis was used to study the driving factors. The trend research result showed, the farmland area decreased year by year for recent 20 years, the decreasing rate was faster before 2004, and then it decreased slowly. The correlation analysis result showed, the added value proportion of primary industry, the total population, and the urban population were the most important driving factors which influenced the farmland quantity change, their coefficients were 0.959, -0.918 and -0.896 respectively. It could be concluded by the result that the industry configuration and the urbanization could affect the farmland quantity with the development of social economy.

**Keywords** data mining; ARIMA model; farmland quantity; change trend; driving factors