

基于标准差变换的农用地产能核算可实现产量校正与建模分析

尹晋磊, 黄宗维 (广西壮族自治区国土测绘院, 广西南宁 530023)

摘要 简要介绍了农用地产能层次划分和核算的基本原理;以柳州市区农用地产能核算项目数据为例,重点分析了可实现产量的数据收集过程和误差来源,运用标准差变换手段对外业数据进行分析处理,详细验证了处理前后数据分布和变化的情况,并对异常数据溯源,简要说明其产生的原因。

关键词 农用地;产能核算;产量校正;评价

中图分类号 F301 **文献标识码** A **文章编号** 0517-6611(2016)20-206-03

Farmland Productivity Calculation Correction Analysis and Modeling Based on Standard Deviation Transformation

YIN Jin-lei, HUANG Zong-wei (Land Surveying and Mapping Institute in Guangxi Zhuang Autonomous Region, Nanning, Guangxi 530023)

Abstract The basic principles of farmland productivity calculation and its levels division were introduced briefly. Based on the project data of farmland productive calculation in Liuzhou City, the collection procedure of production achievable data and related error source were analyzed selectively. Standard deviation was applied on the field data processing, and the data distribution and variation of pre-processing and post-processing were also verified in details. The reason of anomaly data was explained by tracing back to the source.

Key words Farmland; Productivity calculation; Yield correction; Evaluation

农用地综合生产能力调查与评价是国土资源大调查的重要内容之一,也是农用地分等项目的深化与继续。农用地综合生产能力是指在一定地域、一定时期和一定经济、社会、技术条件下所形成的耕地产能,形式上可以用产能单产和总产表示。开展农用地综合生产能力调查与评价,特别是农用地产能核算,对于全面了解和掌握农用地资源状况,科学评价农用地综合生产能力,分析评价农用地利用强度和潜力,保障国家粮食安全具有重要意义。

1 农用地产能层次划分

《农用地综合生产能力调查与评价技术规范》将农用地产能划分为3个层次:理论产能、可实现产能和实际产能。其中,理论产能基础数据主要来源于农业科技部门在试验田获得的各种测产数据,意指使用最好的种苗,在最好的生产条件下获得的产量,是一种理想状态下的产量,也是未来努力实现的目标;实际产量是基于统计数据获得的多数农田的产量,是现实条件下多数农田已获得的产量;而可实现产量介于两者之间,是在农业生产条件得到基本保证,光、热、水、土等环境因素均处于正常状态,技术条件可以满足,由政策、投入等因素决定的正常年景下农作物所能够达到的最高产量^[1]。该研究主要讨论可实现产能基础数据的获得和校验。

2 产能核算的原理

耕地现实生产能力是在当前生产技术条件及社会经济水平下,可以获得的耕地生产能力,从理论上理解,必定有耕地资源现实生产能力得到最大程度实现的地块存在,因此可将这些地块的粮食作物年产出量近似等同于该地块的现实产能^[2]。若这样的地块数量达到一定程度,空间分布合理,就可应用这些样本建立产量与地块分等指数的关系模型,运用空间地理插值法,得到整个区域的耕地现实产能分布情况。而根据农用地分等理论,农用地利用等指数可以认为是

耕地在该层次的产能理论上限的量化表达,因此产能核算的基本方法就是对调查地块的最大产量与其对应的利用等指数做回归分析,建立数学模型,测算整个区域耕地生产能力。

3 可实现产量数据调查方法和局限

由于目前实践中无法直接获取农用地分等单元的可实现产能单产,只能将抽样单元农用地指定作物近3~5 a正常年景下的实际最高单产作为该单元的可实现产能单产^[1],因此可实现产量主要是通过外业调查、实地咨询农民的方式获取。这种调查方法的优点是能够得到田间最真实直接的数据,但调查发现,通过实地向农民直接咨询的方式获取的可实现单产样本,由于生产者本身的局限性,提供的数据通常是凭借生产经验得出的估计值,这个数值通常反映的是该农用地近年的实际产量最高值,不能直接用于农用地产能核算。

4 外业调查样本分析

以柳州市区为研究区域,采用分层抽样法,在综合考虑样点空间覆盖均匀与等别覆盖齐全的基础上,预设调查样点95个(水田58个、旱地37个),共涉及全部12个有现状耕地的乡镇,调查范围达到总耕地面积的65%。在调查过程中,根据具体情况,如因现状已转变为建设用地,随机增补部分调查样点,实际调查了水田60个样点(新增2个水田)、旱地37个样点,共计97个样点。其中各等级分布样点个数及地类见表1。

目前在实践中无法获取抽样单元农用地可实现单产的样本,所以只能调查农用地近3~5 a正常年景下的最高单产,将其看做农用地可实现单产的样本值。在具体实践过程中,可实现单产样本值来自于外业调查中到田间地头及村委会等实际调研所得,由于大多数劳动者没有定量的科学统计,而是根据经验给出近几年的“最高单产的经验平均值”,因此,笔者将获取的数据直接用于建模,拟合效果如图1所示。

作者简介 尹晋磊(1982-),男,四川成都人,工程师,硕士研究生,从事国土资源管理研究。

收稿日期 2016-05-23

表1 产能核算抽样分等单元等级和个数统计

Table 1 Productivity calculation sampling grading unit level and number statistics

农用地利用等 Use of grade of farmland	水田 Paddy field//个数	旱地 Dry land//个数
8等 Grade 8	1	13
9等 Grade 9	8	24
10等 Grade 10	0	0
11等 Grade 11	0	0
12等 Grade 12	0	0
13等 Grade 13	1	0
14等 Grade 14	2	0
15等 Grade 15	18	0
16等 Grade 16	29	0
17等 Grade 17	1	0

经整理后可以看出,将实地调研所得“经验值”以相同利用等为区间观察,在整体上依然呈现一定的梯度变化,并利用等级呈正相关性,但与要建立函数关系的利用等指数之间相关性不高,说明实地调查数值在整体趋势上是与农用地分等一致的;但由于其梯度过大,因此实地调查数据必须经过修正才能应用于综合产能核算中进行建模分析。

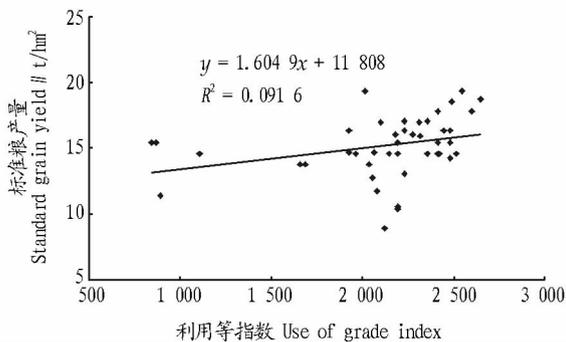


图1 直接拟合建模效果

Fig. 1 Direct fitting modeling effect

5 外业调查样本校正

对可实现单产样本数据校正的原理是:对柳州市区利用等指数进行整体分析,求出各等别间的梯度值,再利用梯度值进行标准化变换形成标准化得分的判断依据,利用标准化得分对可实现单产进行判断。标准化得分过大的作为异常值剔除,在可接受范围内的进行调整,在正常范围内的可以直接使用。

计算出利用等指数标准差和各作物可实现最高单产标准差,再对这2个标准差值做标准化变换,公式为:

$$Z_i = \frac{X_i - \bar{X}}{\sigma} \quad (i=0,1,\dots,n) \quad (1)$$

式中, Z_i 为第*i*个样本的标准化得分,能表明各数据在一组数据分布中的相对位置,是一个以标准差为单位的测量值; σ 为标准差; X_i 表示第*i*个样本的可实现最高单产; \bar{X} 为全部样本的可实现最高单产的平均值。

设 Z_1 为利用等指数的标准化得分, Z_2 为各作物可实现最高单产的标准化得分,令其满足下列关系:

$$Z_1 - Z_2 \leq \alpha \quad (2)$$

以 α 作为阈值来调整各作物可实现最高单产标准化得分,调整后的 Z_2' 进行下列计算:

$$Z_2' \times \sigma_2 + \bar{X}' = X' \quad (3)$$

式中, Z_2' 为调整后的各作物可实现最高单产标准化得分; σ_2 为各作物可实现最高单产标准差; \bar{X}' 为各作物可实现最高单产的平均值; X' 为各作物可实现最高单产校正后的值。

校正的原则是:标准化得分 $Z_i < 0.5$ 为正常值,可直接用于建模; $0.5 \leq Z_i \leq 1.0$ 为正常样本,经标准差变换后用于建模; $Z_i > 1.0$ 为异常值,校正幅度较大。根据这个原则进行筛查可知,柳州市区可实现样点共取得97个有效样本,其中水田60个、旱地37个,以早稻为例,校正过程如表2所示。

表2 早稻可实现单产校正过程(部分)

Table 2 Correction process of yield per unit area of early season rice

所在乡镇(街道办) Towns (subdistrict office)	所在村 Villages	利用等指数 Use of grade index	利用等指数标准差变换 Standard deviation transformation of use of grade index	早稻可实现单产 Yield per unit area of early season rice	早稻标准差变换 Standard deviation transformation of early season rice	异常值判断 Abnormal value judgement	修正参数 Correction parameter	早稻可实现单产修正 Correction of yield per unit area of early season rice
长塘乡 Changtang Country	梳庄	844.78	-2.99	7 500	-0.12	1	-2.49	4 422.9
太阳村镇 Taiyangcun Town	西鹅	892.29	-2.88	5 250	-1.85	1	-2.38	4 566.3
白露街道办 Bailu subdistrict office	园艺	1 658.19	-1.10	6 750	-0.70	0	-0.70	6 750.0
长塘乡 Changtang Country	青茅	1 926.42	-0.48	7 500	-0.12	0	-0.12	7 500.0
沙塘镇 Shatang Town	龙卜	2 015.99	-0.27	10 500	2.19	1	0.23	7 958.4
石碑坪镇 Shipiping Town	石碑坪	2 120.38	-0.03	4 500	-2.43	1	-0.53	6 974.3
∴	∴	∴	∴	∴	∴	∴	∴	∴
均值 Mean		2 131.53		7 657.5				7 590.0
标准差 Standard deviation		430.40	0.00	1 299.3	0.00			

图2显示,校正后各指定作物与对应农用地利用等指数拟合,取得了较理想的效果。

校正后统计样点数据如表3所示,柳州市区农用地可实现单产样本在正常值范围内的个数所占比例为78.4%,其余

不在正常值范围内的样本经小范围校正后建立的模型是合理可靠的。

6 异常点分析

测算过程中,认定为异常点有21个,其中旱地10个、水

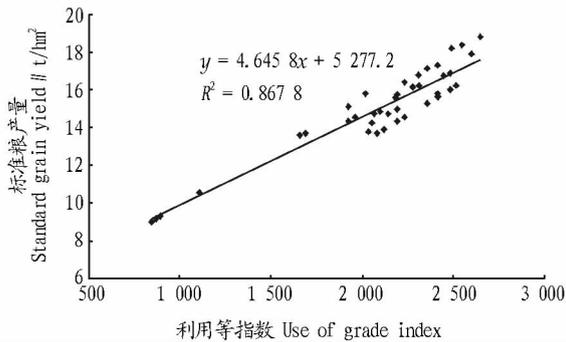


图2 柳州市区可实现单产(校正值)—利用等指数曲线拟合

Fig.2 Yield per unit area(corrected value)—use of grade index curve fitting in Liuzhou region

田 11 个在计算中予以剔除。旱地异常样点调查甘蔗产量达到 120~150 t/hm², 偏离正常产量范围。11 个水田样点中, 7 个经核实为菜地(分等成果采用过渡期分类系统, 产能核算采用全国土地利用现状分类, 将菜地衔接为水浇地), 在前期样点筛选中为调查分类的方便将菜地均归并水田调查。通过查阅柳州市农用地分等成果原始数据发现, 柳州市区菜地的指定作物中, 既有甘蔗也有水稻, 在实际调查中也多次出现了“水田样点, 实为旱地”的情况, 这些以甘蔗为指定作物分等的菜地, 应作为旱地对待, 其余 4 个样点为产量与多数同等级耕地产量差异过大。

表3 柳州市区可实现单产样本检验统计

Table 3 Sample test statistics of yield per unit area in Liuzhou region

标准化得分 Standardized score	水田 Paddy field	旱地 Dry land	占样点比例 Proportion in sample points // %
$Z_i < 0.5$	27	18	46.4
$0.5 \leq Z_i \leq 1.0$	22	9	32.0
$Z_i > 1.0$	11	10	21.6
合计 Total	60	37	100

7 小结

该方法的实践获得了较理想的数学属性, 但其应用是限于各乡镇农用地(耕地)质量和利用水平关系未发生重大变

化的前提下开展的。

农用地自然质量与土地本身属性有关。由表 4 所示的柳州市区耕地分等主导因素及权重可见, 在柳州地区影响耕地质量等级的主导因素中除灌溉保证率受排灌措施影响外, 其余因子是长期相对稳定的, 而水田灌溉保证率得分梯度约为 20 分^[3], 即发生 1 个等级变化, 对总分的影响约为 0.03 分, 旱地只是基准分略高, 即灌溉对总分影响更小, 因此自然质量分是可用的。

表4 柳州地区耕地分等因素权重

Table 4 Cultivated land grading factor weight in Liuzhou region

评价因子 Evaluation factors	水田 Paddy field	旱地 Dry land
表层质地 Surface texture	0.18	0.12
有机质含量 Organic matter content	0.20	0.15
pH 值 pH value	0.15	0.1
坡度 Slope	0.09	0.21
灌溉保证率 Ensurance probability of irrigation water	0.16	0.17
土层厚度 Soil thickness	0.22	0.25

广西农用地分等工作以 2002 年为基期年, 农用地产能核算以 2009 年为基期年, 7 a 来, 自治区广大农村地区的经济发展取得了重大的成果, 但从整体上看, 除个别极端情况(如贫困或近郊地区由于政策扶持或产业结构改变), 发展过程是相对平稳的, 即以乡(镇)为单位其发展水平是基本同步的。试点区为柳州市辖区, 无极端贫困的情况, 而城市近郊农村地区虽因城市发展经济水平得到提高, 但这并非由农业发展所带来的, 利用水平的绝对值可能有所提高, 但其相对关系是基本稳定的; 因此在 2 个前提下进行的数据校正, 和在此基础上所建立的产能核算模型是可以接受的。

参考文献

- [1] 国土资源部. 农用地生产能力调查与评价技术规范[Z]. 2009.
- [2] 伍育鹏, 郎文聚, 邹如. 耕地产能核算模型的研究[J]. 农业工程学报, 2008, 24(S2): 108-113.
- [3] 广西国土测绘院. 广西耕地质量等级评价技术报告[R]. 2004.

(上接第 198 页)

- [29] 杨赛明, 徐跃通, 张邦花. 区域土地资源可持续利用的生态安全评价[J]. 中国人口·资源与环境, 2010, 20(3): 325-328.
- [30] 汪慧玲, 朱震. 区域土地资源生态安全分析: 以甘肃省为例[J]. 重庆理工大学学报(社会科学版), 2015, 29(4): 15-20.
- [31] 高桂芹, 韩美. 区域土地资源生态安全评价: 以山东省枣庄市中区为例[J]. 水土保持研究, 2005, 12(5): 271-274.
- [32] 叶明武, 陈振楼, 王军, 等. 城镇地区生态安全评价及对策: 以上海崇明岛为例[J]. 城市环境与城市生态, 2007, 20(1): 9-13.
- [33] 张军以, 苏维词, 张婕. 2000-2009 年重庆土地资源生态安全评价及趋势分析[J]. 地域研究与开发, 2011, 30(4): 128-131.
- [34] 张志勇, 刘希玉. 基于 SVM 的区域土地资源生态安全评价研究[J]. 计算机工程与应用, 2009, 45(10): 245-251.
- [35] 陈书广. 河南省土地生态安全评价研究[D]. 郑州: 河南大学, 2006: 6-15.
- [36] 苏臣. 常德市土地资源生态安全评价[D]. 长沙: 中南林业科技大学, 2009: 6-10.
- [37] 王韩民. 国家生态安全评价体系及其战略研究[D]. 西安: 西北工业大

- 学, 2006: 9-17.
- [38] 类淑霞, 郝晋珉, 尚韬. 基于土地利用变化的区域土地生态安全动态评价: 以北京市朝阳区为例[J]. 生态经济, 2011, 241(7): 38-42.
- [39] 王鹏, 唐丽, 田亚平, 等. 湖南省土地资源生态安全评价[J]. 资源开发与市场, 2012, 28(3): 257-261.
- [40] 卢艳敏. 东江源区土地可持续利用现状与生态安全评价研究[D]. 南昌: 江西理工大学, 2012.
- [41] 李锐. 中卫市沙坡头区土地利用生态安全评价研究[D]. 银川: 宁夏大学, 2012.
- [42] 史永龙, 贾宪威, 蔡昕, 等. 省域土地生态安全评价研究: 以江苏省为例[J]. 安徽农业科学, 2011, 39(4): 2365-2368.
- [43] 李茜, 任志远. 区域土地生态环境安全评价[J]. 干旱区资源与环境, 2007, 21(5): 75-79.
- [44] 四川地理环境概况[EB/OL]. [2016-04-01]. <http://wenku.baidu.com/view/7382cf154431b90d6e85c798.html>.
- [45] 张莉. 四川盆周山地土地整理的生态安全研究[D]. 成都: 四川大学, 2009.