

粮食产量预测理论、方法与应用

IV.粮食估产理论、模型及其应用

米长虹¹, 王 农¹, 黄治平¹, 侯彦林^{1*}, 刘书田¹, 郑宏艳¹, 蔡彦明¹, 夏 维¹,
王铎今², 任 军³, 王新民⁴, 侯显达⁵

(1.农业部环境保护科研监测所, 天津 300191; 2.北京农业信息技术研究中心, 北京 100089; 3.吉林省农业科学院, 吉林 长春 130033; 4.河南牧业经济学院, 河南 郑州 450011; 5.北京优雅施软件研发服务中心, 北京 100089)

摘 要:粮食估产的“通道-概率”理论:把属于最近通道的历年来的产量划分为5个气候年型通道,即丰产年、偏丰年、平产年、偏欠年、欠产年;计算产量出现在5个气候年型中的频率作为概率使用,估产年的初始估产值等于预测年各通道内平均产量与概率之积的和;估产值等于初始估产值与气候年型修正参数之积,专家根据当年气候条件和作物长势实时确定修正参数。预报单元为全国、省和县。应用结果表明:国家尺度上不需要修正,省和县级尺度需要气候年型参数修正;预测误差在3%以内;所述估产理论严谨、方法简单,参数少,参数来自原始数据本身和专家经验,易于推广使用。

关键词:粮食;估产;理论;方法;应用

中图分类号:S114

文献标志码:A

文章编号:2095-6819(2014)03-0227-06

doi: 10.13254/j.jare.2014.0070

The Theory, Method and Its Application of the Grain Yield Forecast

IV. The Theory, Method and Its Application of Grain Yield Estimation

MI Chang-hong¹, WANG Nong¹, HUANG Zhi-ping¹, HOU Yan-lin^{1*}, LIU Shu-tian¹, ZHENG Hong-yan¹, CAI Yan-ming¹, XIA Wei¹,
WANG Shuo-jin², REN Jun³, WANG Xin-min⁴, HOU Xian-da⁵

(1.Agro-Environmental Protection Institute, Ministry of Agriculture, Tianjin 300191, China; 2.Beijing Research Center for Information Technology in Agriculture, Beijing 100089, China; 3.Jilin Academy of Agricultural Sciences, Changchun 130033, China; 4.Henan College of Animal Husbandry and Economy, Zhengzhou 450011, China; 5.Software Development and Service Center of Beijing Yours, Beijing 100089, China)

Abstract: The ‘channel probability theory’ of grain yield estimation is defined as follows: the historical production which belongs to recent channel is divided into 5 climatic-year types, i.e. high yield year, partial harvest year, normal year, partial less year, a bad crop year. The frequency(or probability) for the year appeared in 5 climatic types is calculated. The initial value of estimation yield is equal to the sum of the product of probability and the average yield within every channel for the forecasting year, and the value of estimation yield is equal to product of the initial estimation value and the climatic parameters. The parameters are real-time amended by experts according to the climatic conditions and crop growth. Prediction units are nation, province and county. The results of application showed that the parameters did not need to be modified on national scale, while the climatic parameters needed to be revised on province and county scales. The prediction error was within 3%. The theory would have good application effect and promising development due to its preciseness, simple method, and less parameters which were from original data and expert experience.

Keywords: grain; grain estimating; theory; methods; application

我国粮食估产研究的代表单位为中国科学院数
学研究院, 到目前为止已连续 30 多年预报全国粮食

产量^[1-2]。其所使用的方法为投入占用产出技术模型,
提前半年预报产量,平均预测误差为 2.0%左右,最大
误差为 5%~8%,在国际同类工作中处于领先水平;目
前已获得 20 多个产量预测方程,主要考虑如下 18 类
因素,即耕地面积、复种指数、粮食播种面积、受灾面
积、受灾严重程度、政策虚变量、价格因素、种粮比较
利益、化肥施用量、农家肥施用量、灌溉面积比重、机

收稿日期:2014-03-26

基金项目:中国农业科学院科技创新工程(2014-cxgc-hyl)

作者简介:米长虹(1968—),男,天津人,副研究员,环评工程师,主要
研究方向为农业环境数据挖掘。E-mail:mch68@sina.com

* 通信作者:侯彦林 E-mail:bjyours@sina.com

耕面积比重、农业机械总动力、耕畜数量、农用塑料薄膜使用量、良种推广程度、农业用电、农业劳动力数量。笔者认为以上工作存在以下不足:主要以全国粮食总产估产为主,基本未涉及中、长期生产潜力研究;模型中考虑因素太多,难以及时准确获得参数;如提前半年预测,不能反映半年后的受灾面积、受灾程度以及化肥追肥的数量等;当年预测不重视气象因素,国家级的预测因为有区域性的互补作用,所以从长时间周期来考虑气象因素是可以互相抵消的,但是对于省级及县级预测就难以实现,因为这种气象因素难以在区域范围内互相抵消;预测单一作物时可能会遇到困难,因为单一作物的互补性较小;预测具体模型不固定,随时筛选,主要形式是多项式回归,在 18 个变量中的多个变量的微小变化可以导致预测结果偏差的累加;虽有软件程序,但未公开。

笔者从 2005 年开始研究粮食生产潜力预测,并在此基础上建立了粮食估产理论、方法、算法以及编制了软件。预测中所使用的数据仅为统计年鉴中的单产、总产和播种面积数据,所建立的理论、方法和算法公开,希望将其发展成为预测平台供所有用户因地制宜地使用,为国家和各级政府、商贸机构、研究者提供系统的估产理论、方法和软件工具以及系统决策信息。软件在估产方面的功能主要是预测各行政单元的粮食或单一作物的单产和总产^[3-5]。

1 材料与方法

从公开发表的数据中整理出全国和东北三省

1949—2010 年粮食单产数据以及 31 个省、区、直辖市 1979—2010 年粮食单产数据。使用本文提出的方法建模和分析。

2 结果与分析

2.1 粮食估产的“通道-概率”理论

影响作物当年产量的因素很多,总体可以分成 2 类综合因素。一类是具有一定周期性的气候因素,一类是持续递增的科技进步因素,后者对粮食产量的影响从量变逐渐达到质变。这样作物产量的变化过程整体就呈现波浪式上升的趋势,并由最高产量和最低产量连线形成一个具有一定振幅或宽度的通道,如图 1。当产量水平发生突变时,产量水平整体提高,产量也就进入到了另一个通道内,在新的通道内再波动上升。我们将作物产量在其间发展变化的这个通道称为产量通道。产量变化趋势随着年限的增加,产量通道也会有所变化,但每一个最新的产量通道都有 2 个产量边界,分别为高产线和低产线,其中低产线作为产量安全线,一定程度上保障着某一层次的粮食产量安全,产量一旦低于这个安全线,就需要政策或者投入的强烈干预,以保证粮食需求。

粮食估产的“通道-概率”理论和方法是:以粮食产量(单产或总产)的真实信息(一般为国家和政府的统计信息)为基础,把属于最近通道的历年来的产量划分为 5 个气候年型,分别定义为丰产年、偏丰年、平产年、偏欠年、欠产年,由于产量在通道内的波动是受气候因素影响的,那么相应的产量等级也就代表着不

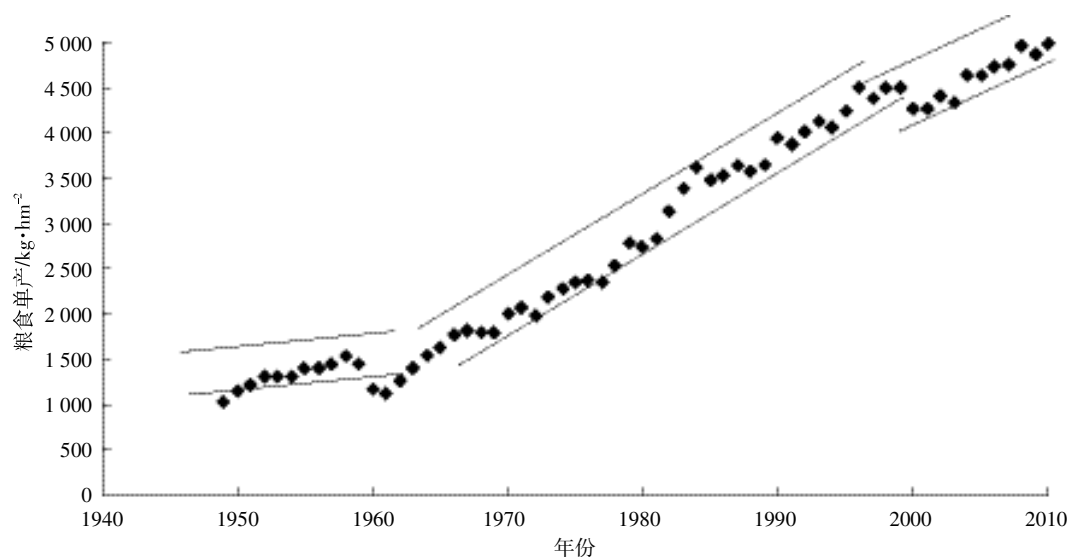


图 1 单产通道划分示意图

Figure 1 The channel division of grain yield per unit

同的气候年型;然后将历史产量点看作离散型的随机变量,计算产量出现在5个气候年型中的频率作为概率,预测产量等于预测年各通道内平均产量与概率之积的和;修正参数由当地专家根据当年实时气候条件和作物长势具体确定。预报单元为全国、省、县,预测全国和省时,则要考虑各省和各县的权重,可根据最近5年各省和各县单产或总产贡献率确定。

2.2 粮食估产“通道-概率模型”方法

估产方法和步骤如下:

(1)确定通道:把最近趋势一致的年型作为建模样本年,并确定起、止年。

(2)平产年趋势模型建立:对通道内所有样本进行直线回归获得平产年趋势模型。

(3)将平产年趋势模型平移相交于历史最好年和最差年的点,在斜率不变的情况下把最好年和最差年的年限(X)和实际产量(Y)带入平产年趋势模型中,反求出最好年和最差年的截距,于是获得了最好年和最差年趋势模型。

(4)分别把最好年和最差年趋势线与平产年趋势线之间的距离等分3份,其中平产年趋势线两侧的2条通道合并为一个通道形成平产年通道,从而形成5条通道,如图2。

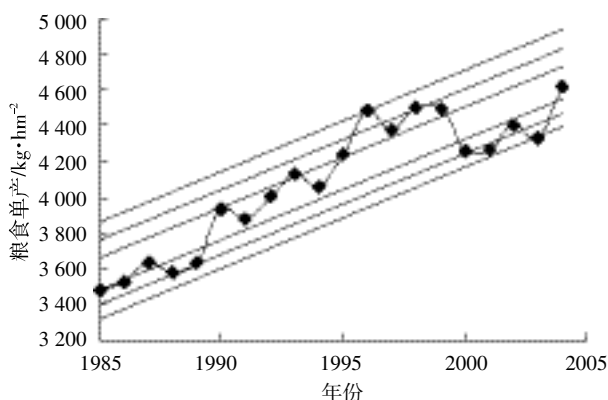


图2 全国1985—2004年粮食单产通道

Figure 2 The channel of the national grain yield per unit from 1985 to 2004

(5)计算样本年在5条通道内出现的频率,作为概率参数使用。

(6)计算5条通道的6条边界线的各年产量值,并取每条通道上下边界线的平均作为每条通道内每年的平均产量。

(7)未知年估产=预测年的每条通道内的平均产量 $\times$ 各自的概率的累加。

(8)连续逐年预测时可使用前一年预测误差作为小趋势误差修正参数,见参考文献[5],即在第2年预测时,在预测结果基础上减去前一年的误差,即为第2年修正后的预测结果。

(9)如果为了检验预测模型,可以事先预留2~3年数据来验证模型,从而计算历史年的预测误差;如果是真实年的预测,则最后以各预测单元政府公布的数据为实际产量而计算预测误差。

(10)估产所需数据:1949或1979年开始的国家、省级、县级的粮食和单一作物单产、总产、耕地面积;每个预测单元需聘请3~4位专家在生长季节实时确定当年气象条件和作物长势修正参数。

(11)省和县级参数确定方法:除以上5个气候年型趋势线外,将最好年和最差年趋势线定义为超丰年和超欠年,这样形成7个气候年型趋势线,即超丰年、丰产年、偏丰年、平产年、偏欠年、欠产年、超欠年,取7个趋势线相邻两两的平均值,并把平产年两侧2个宽度合并成平产年,从而获得丰产年、偏丰年、平产年、偏欠年、欠产年5个年型的平均产量,再加上超丰年和超欠年的产量,可以形成不同气候年型下的初始修正系数。就全国而言,一般定义为:超丰年1.18、丰产年1.12、偏丰年1.06、平产年1.00、偏欠年0.94、欠产年0.88、超欠年0.82。各行政单元可以根据具体情况修正以上参数,其中超丰年和超欠年参数只能在特殊年限使用,并要根据实际情况确定。

(12)各省研究者确定参数:根据当年到某时刻为止的气象情况来确定7个年型中的一个年型从而确定 Ka ,到某时刻为止专家根据田间作物长势来确定7个年型中的一个年型从而确定 Kb ;然后取 Ka 和 Kb 的平均值作为全省和各县粮食或单一作物当年到某时刻为止的气候年型具体修正系数,一般按省、县粮食或单一作物每旬(或每月)更新1次参数。作物播种前产量初始修正系数为1.00,动态显示时是把用户确定的 Ka 和 Kb 按旬动态平均值显示出来,它是实际使用的修正系数。如某作物从5月中旬开始播种,10月上旬收获,则5月上旬的初始值为1.00。作物估产参数数据库见表1。

2.3 粮食估产理论、方法的应用

2.3.1 全国粮食单产、总产估产:案例1

以1985—2007年单产数据为研究对象,其中1985—2004年用来建模,2005—2007年用来验证,分布概率使用1960—1997年的分布概率,结果见图2和表2。

表 1 作物估产参数数据库
Table 1 The database of crops estimating parameters

年份	参数	月-日										
		1-5	1-15	1-25	2-5	2-15	2-25	...	...	12-5	12-15	12-25
2011	Ka	—	—	—	—	—	—	...	...	—	—	—
	Kb	—	—	—	—	—	—	...	...	—	—	—
	$(Ka+Kb)/2$	—	—	—	—	—	—	...	...	—	—	—
2012	Ka	—	—	—	—	—	—	...	...	—	—	—
	Kb	—	—	—	—	—	—	...	...	—	—	—
	$(Ka+Kb)/2$	—	—	—	—	—	—	...	...	—	—	—
...	...	—	—	—	—	—	—	...	...	—	—	—
2018	Ka	—	—	—	—	—	—	...	...	—	—	—
	Kb	—	—	—	—	—	—	...	...	—	—	—
	$(Ka+Kb)/2$	—	—	—	—	—	—	...	...	—	—	—

注:“—”代表用户参数, Ka 为气候参数, Kb 为专家根据作物长势确定的参数。

表 2 全国 2005—2007 年粮食单产预测结果($\text{kg}\cdot\text{hm}^{-2}$)
Table 2 The forecasting result of the national grain yield per unit from 2005 to 2007 ($\text{kg}\cdot\text{hm}^{-2}$)

年份	每年各通道平均产量					预测产量	实际产量	预测误差/%
	1	2	3	4	5			
2005	4 925	4 827	4 688	4 560	4 481	4 680	4 642	0.83
2006	4 980	4 881	4 743	4 615	4 536	4 735	4 716	0.41
2007	5 035	4 936	4 798	4 670	4 591	4 790	4 748	0.88
概率	0.03	0.13	0.61	0.18	0.05			

同样,以 1985—2007 年总产数据为研究对象,其中 1985—2004 年用来建模,2005—2007 年用来验证,分布概率使用 1960—1997 年的分布概率。预测误差:2005 年为 0.72%、2006 年为-1.09%、2007 年为-0.99%。

由表 2 可得出以下结论:就全国尺度而言,总产与单产结果相似,只要是预测一个通道内的产量,其预测误差都很小;全国尺度由于地域间气象条件对产量影响的互补性,原则上可以不使用气候年型修正和小趋势误差修正;各等级出现概率计算时使用的年限可以超过建模期间的年限,以提高预测精度。

2.3.2 全国粮食单产、总产估产:案例 2

以 1979—2008 年单产和总产数据为研究对象,预测 2006 年单产时建模数据为 1979—2005 年,预测

2007 年单产时建模数据为 1979—2006 年,预测 2008 年单产时建模数据为 1979—2007 年,分布概率分别使用建模年限的所有年,结果见表 3、表 4。由此可见,全国单产和总产的预测不需要修正,都等于平产年的预测结果。对比 2.3.1 的研究结果,也说明了在总趋势不变的情况下,建模起止年对估产结果影响不大。

2.3.3 各省单产估产:案例 3

各省粮食单产估产预测误差结果见表 5。每个省气候年型的确定方法是:将实产与超丰年、丰产年、偏丰年、平产年、偏欠年、欠产年、超欠年对应的产量相比,误差最小的即为该年年型,这是未来最理想的预测结果,并非真实预测结果;将超丰年、丰产年、偏丰年、平产年、偏欠年、欠产年、超欠年分别赋予 3、2、1、

表 3 全国粮食单产估产结果($\text{kg}\cdot\text{hm}^{-2}$)
Table 3 The estimating result of the national grain yield per unit($\text{kg}\cdot\text{hm}^{-2}$)

估产年	实际产量	超丰年	丰产年	偏丰年	平产年	偏欠年	欠产年	超欠年	最小误差/%
2006	4 716	5 696	5 406	5 116	4 827	4 537	4 248	3 958	2.35
2007	4 748	5 775	5 481	5 188	4 894	4 600	4 306	4 013	3.07
2008	4 951	5 835	5 538	5 242	4 945	4 648	4 352	4 055	0.12

表 4 全国粮食总产估产结果(10⁷ kg)
Table 4 The estimating result of the national grain total production(10⁷ kg)

估产年	实际产量	超丰年	丰产年	偏丰年	平产年	偏欠年	欠产年	超欠年	最小误差/%
2006	49 748	59 995	56 944	53 893	50 843	47 792	44 742	41 691	2.20
2007	50 160	60 674	57 589	54 504	51 419	48 334	45 249	42 164	2.51
2008	52 871	61 243	58 129	55 015	51 901	48 787	45 673	42 559	1.83

表 5 各省 2006—2008 年单产估产误差
Table 5 The estimating error of grain yield per unit for each province

行政单元	2006 年		2007 年		2008 年		行政单元	2006 年		2007 年		2008 年	
	年型分值	误差/%	年型分值	误差/%	年型分值	误差/%		年型分值	误差/%	年型分值	误差/%	年型分值	误差/%
全国	0	2.37	0	3.08	0	0.12	湖北省	-1	0.41	-1	0.27	-1	3.41
北京市	2	3.11	2	1.77	3	1.50	湖南省	0	2.81	0	2.10	1	3.18
天津市	-1	1.31	0	3.58	-1	3.27	广东省	-2	2.37	-1	2.35	-2	1.35
河北省	0	0.12	1	2.80	1	2.61	广西自治区	0	1.62	1	3.08	0	2.00
山西省	1	3.09	0	1.18	0	3.06	四川省	-2	1.20	-1	0.93	0	2.28
内蒙古自治区	0	3.05	-1	0.07	1	1.68	贵州省	-1	3.51	0	3.06	1	3.28
辽宁省	0	2.32	1	2.37	1	0.44	云南省	0	3.20	0	2.73	0	2.36
吉林省	0	3.05	-2	0.13	0	2.23	西藏自治区	0	0.25	0	0.55	0	0.82
黑龙江省	-1	0.78	-3	0.13	-1	2.02	陕西省	1	3.10	1	0.91	1	0.01
上海市	-1	1.20	-1	3.64	-1	0.62	甘肃省	-1	2.95	-1	0.15	0	0.37
江苏省	0	2.50	-1	2.39	-1	2.35	青海省	-1	0.63	0	3.55	0	2.02
浙江省	0	3.22	1	1.20	1	0.54	宁夏自治区	0	0.18	-1	1.69	0	1.78
安徽省	-1	2.13	-1	3.22	0	1.72	新疆自治区	-1	2.62	0	2.78	-1	3.03
福建省	0	0.09	1	2.61	1	2.20	海南省	—	—	—	—	—	—
江西省	0	0.38	0	2.21	0	1.97	重庆市	—	—	—	—	—	—
山东省	0	0.52	0	0.74	0	0.15	平均	-0.28		-0.17		0.10	
河南省	1	0.20	1	0.30	1	2.42							

注:“—”代表数据缺失。

0、-1、-2、-3 数值,统计出平均气候年型是 2006 年为-0.28%即为平产年偏弱,2007 年为-0.17%即为平产年偏弱,2008 年为 0.10%即为平产年偏强,2006—2008 年为-0.35%即为平产年偏弱。

3 讨论

关于估产精度的理论分析:首先我们可以确定的是根据气候条件和作物长势确定的修正系数,在优于或劣于平产年型的判断上不会有错,即当地专家根据气象条件和作物长势,不会把明显的减产年判断为增产年和平产年,同样也不会把明显的增产年判断为减产年和平产年。固定参数系统 K 的取值为超丰年 1.18、丰产年 1.12、偏丰年 1.06、平产年 1.00、偏欠年 0.94、欠产年 0.88、超欠年 0.82。如果确定作物长势优于平产年则 K 可取 1.06,明显优于平产年则 K 可取 1.12,特别优于平产年则 K 可取 1.18;如果作物长势

不好 K 可取 0.94,明显长势不好 K 可取 0.88,特别长势不好 K 可取 0.82;如果作物长势和一般年一样 K 可取 1.00。我们需要确定的就是 7 个参数之一,如果这 7 个等级确定对了,理论上的预测误差在 3%左右。以上粮食估产方法预测精度之高的主要原因在于每年科技进步增加的产量已经被排除在波动之外了,它作为一种趋势的增加量已经被计算在未来年的产量之中了。

4 结论

关于粮食估产理论、方法和应用研究得出以下 4 点结论:本研究提出的理论严谨,方法简单,参数少,参数来自原始数据本身,易于推广;使用了气候年型的概念,但是没有用到具体的气候参数,这样较适合国家和省级尺度粮食的估产;预测国家尺度时,平均预测误差在 2%以下,原则上不需要修正;预测省级尺

度时,需要气候年型参数修正,修正后的平均误差一般低于 3%。关于县级长期和根据当年气象数据确定估产修正系数的方法将另文报道。

参考文献:

- [1] 陈锡康, 郭菊娥. 中国粮食生产发展预测及其保证程度分析[J]. 自然资源学报, 1996, 11(3): 197-202.
CHEN Xi-kang, GUO Ju-e. Forecast of China's grain production development and its guarantee degree analysis[J]. *Journal of Natural Resources*, 1996, 11(3):197-202.(in Chinese)
- [2] 陈锡康. 全国粮食产量预测研究[J]. 中国科学院院刊, 1992(4): 330-333.
CHEN Xi-kang. Forcast study on the whole China[J]. *Grain Yield Bulletin of Chinese Academy of Sciences*, 1992(4): 330-333.(in Chinese)
- [3] 侯彦林, 郑宏艳, 刘书田, 等. 粮食产量预测理论、方法与应用 I. 科技进步增产理论、模型及其应用[J]. 农业资源与环境学报, 2014, 31(3):205-211.
HOU Yan-lin, ZHENG Hong-yan, LIU Shu-tian, et al. The theory, method and its application of the grain yield forecast I. The theory, model and its application of scientific and technological progress in increasing grain yield [J]. *Journal of Agricultural Resources and Environment*, 2014, 31(3):205-211.(in Chinese)
- [4] 郑宏艳, 刘书田, 王铎今, 等. 粮食产量预测理论、方法与应用 II. 粮食生产潜力短期预测理论、模型及其应用[J]. 农业资源与环境学报, 2014, 31(3):212-219.
ZHENG Hong-yan, LIU Shu-tian, WANG Shuo-jin, et al. The theory, method and its application of the grain yield forecast II. The theory, method and its application of short-term forecast of the grain yield potential [J]. *Journal of Agricultural Resources and Environment*, 2014, 31(3):212-219.(in Chinese)
- [5] 刘书田, 王铎今, 米长虹, 等. 粮食产量预测理论、方法与应用 III. 粮食生产潜力中、长期预测理论、模型及其应用[J]. 农业资源与环境学报, 2014, 31(3):220-226.
LIU Shu-tian, WANG Shuo-jin, MI Chang-hong, et al. The theory, method and its application of the grain yield forecast III. The theory, method and its application of medium and long-term forecast of the grain yield potential [J]. *Journal of Agricultural Resources and Environment*, 2014, 31(3):220-226.(in Chinese)