

利用时间序列模型预测黑龙江省大豆的比较优势

苗森, 郑煜

(东北林业大学理学院, 黑龙江哈尔滨 150040)

摘要:以黑龙江省26年的大豆产量和种植规模作为研究对象,根据时间序列分析和作物比较优势的原理,建立大豆综合比较优势指数的时间序列模型,对优势指数做出预测。结果表明:2012~2017年黑龙江省大豆生产的比较优势指数分别为1.873、1.7323、1.7592、1.6106、1.6383,说明相对于全国大豆生产是占优势的,但是指数呈逐年下降趋势,所以建议有关部门给予足够的重视。

关键词:大豆;时间序列模型;综合比较优势;预测模型

中图分类号:F326.11

文献标识码:A

文章编号:1000-9841(2013)01-0080-04

Prediction for Soybean Comparative Advantage in Heilongjiang Province Based on Time Series Model

MIAO Sen, ZHENG Yu

(College of Science, Northeast Forestry University, Harbin 150040, Heilongjiang, China)

Abstract: According to the principle of time series and crop comparative advantage, taking the yield and planting area of soybean in Heilongjiang province from 1985 to 2010 as the research objects. We established the time series prediction model of the comprehensively comparative advantage of soybean in Heilongjiang province. The results illustrated that from 2012 to 2017 the comparative advantage for soybean production in Heilongjiang province were 1.873, 1.7323, 1.7592, 1.6106 and 1.6383, respectively. This proved the soybean production in Heilongjiang province had an advantage over national level, but the advantage showed a decreasing trend. We suggest that relevant agricultural departments should pay enough attention about this trend.

Key words: Soybean; Time series model; Comparative advantage; Prediction model

在农业经济学中,比较优势分析法是经常使用的方法之一,这个方法能够反映某地区的规模化、专业化生产程度、资源禀赋、经济发展水平状况和社会需求状况。从而能准确掌握、描述及分析预测某一区域的某种或者多种作物比较优势的 trend 和变化规律,为加快该地区生产布局的合理性、发挥比较优势、调整相应产业结构,也为促进生产、保障安全起到积极作用。

时间序列分析方法属于统计学的范畴,经过研究、分析和处理时间序列,提取出系统相关的信息,从而揭示出时间序列本身的结构和规律,认识系统的固有特性,把握系统与外界的联系,进而推断出系统在将来的变化和 behavior。因此,时间序列分析方法不仅仅是一种数据处理方法,它已经演变为一种系统分析研究的方法。随着时间序列分析方法的不断成熟,在社会生活、经济、农业生产以及国防建设中应用广泛^[1]。

近年来,受种植比较收益影响和进口大豆的冲击,黑龙江大豆种植面积部分转种玉米或水稻。根据农业相关部门的统计,农户大豆转种玉米的实际

比例要低于此前的预期值,2012年黑龙江大豆种植面积较2011年缩减两到三成。本研究利用时间序列模型的原理,对黑龙江省大豆比较优势趋势进行了预测。

1 作物比较优势

根据经济学相关原理,作物的比较优势主要表现在单产、综合和规模比较优势三方面上^[2]。

单产比较优势是描述资源、科技等相关综合生产力的一个指标,通过单产水平的高低来进行衡量农业生产效率高低。单产比较优势的指数公式为:

$$\delta_i = \frac{x_i/x}{y_i/y} \quad (1)$$

公式中, y_i 为全国第*i*种作物的单位面积产量,而 y 为全国作物的单位面积产量。 x_i 代表第*i*种作物的单位面积产量, x 为作物的单位面积产量。如果 $\delta_i > 1$,那么就说明该作物相对于全国的平均生产水平有一定的单产比较优势,如果数值越大,就代表其优势越明显;反之则不明显。

规模比较优势说明的是一个地区的某一种作

收稿日期:2012-08-16

基金项目:黑龙江省自然科学基金(G201117)。

第一作者简介:苗森(1987-),男,在读硕士,研究方向为数理统计。E-mail:546761142@qq.com。

通讯作者:郑煜(1962-),女,教授,主要研究方向为概率论与数理统计。E-mail:zhengyu62@126.com。

物的生产集中程度与生产的规模。其规模越大,则相应的指数也应越大。由于它是多种因素相互作用的结果,因此在生产上最终体现为比较优势。规模比较优势指数公式:

$$\gamma_i = \frac{u_i/u}{v_i/v} \quad (2)$$

公式中, v_i 为全国第 i 种作物的播种面积, v 为全国作物的播种面积或者总播种面积, u_i 为第 i 种作物的播种面积, u 为作物播种面积或者总播种面积。如果 $\gamma_i > 1$, 那么就表明该作物的生产拥有一定程度的专业集中度, 并且如果规模优势的指数越大, 则说明其专业化程度越高; 反之则没有优势。

由于单纯的运用单产比较优势和规模比较优势并不能综合地反映市场需求、科技等诸多方面的综合信息, 所以引入综合比较优势的这个概念, 如下定义综合比较优势的指数公式:

$$\varepsilon_i = \sqrt{\delta_i * \gamma_i} \quad (3)$$

这个模型是规模比较优势和单产比较优势的综合表述, 其更为全面的反映了某地区某种作物生产的优势度。如果 $\varepsilon_i > 1$, 那么则表明这种作物和全国平均水平相比有优势, 并且, 如果综合比较优势指数值越大, 则说明优势越明显。反之则没有比较优势。

2 时间序列模型

2.1 平稳时间序列模型

如果一个随机过程的方差和均值在时间过程上都是常数, 并且在任意二个时期之间的协方差只依赖这二个时期之间的距离, 那么就称这样的随机时间序列是平稳的。对于平稳随机时间序列可建立下列形式的时间序列模型^[3]。

2.1.1 自回归模型 (AR) 如果时间序列值 y_t 可以表示成为它的先前值 y_{t-1} 与一个冲击值 a_t 的线性函数, 那么称此模型为自回归模型。将模型 $y_t = \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + a_t$ 称为 p 阶自回归模型, 简记 AR(p)。

2.1.2 移动平均模型 (MA) 如果序列值 y_t 是现在和过去的误差或者冲击值 a_t 的线性组合, 那么称此模型为自回归滑动平均模型, 模型为 $y_t = a_t - \theta_1 y_{t-1} - \theta_2 y_{t-2} - \dots - \theta_q y_{t-q}$, 称为 q 阶滑动平均模型, 简记 MA(q)。

2.1.3 自回归平均移动模型 (ARMA) 如果时间序列值 y_t 是现在和过去的误差或者冲击值 a_t 以及之前的序列值 y_{t-1} 的线性组合, 那么称此模型为自回归滑动平均模型, 模型为 $y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \theta_1 a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$, 这里

p, q 分别表示自回归和滑动平均的阶, 简记 ARMA(p, q) 模型。实数 $(\varphi_1, \varphi_2, \dots, \varphi_p)$ 称为自回归系数, 实数 $(\theta_1, \theta_2, \dots, \theta_q)$ 称为移动平均系数。

ARMA 模型的参数估计^[4] (假设 p, q 已知)。首先对预测数据 x_1, x_2, \dots, x_N 进行零均值化处理,

如下: $y_t = x_t - \bar{x}_N, t = 1, 2, \dots, N, \bar{x}_N = \frac{1}{N} \sum_{j=1}^N x_j$ 。

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} \gamma_q & \gamma_{q-1} & \dots & \gamma_{q-p+1} \\ \gamma_{q+2} & \gamma_q & \dots & \gamma_{q-p+2} \\ \vdots & \vdots & \dots & \vdots \\ \gamma_{q+p-1} & \gamma_{q+p-2} & \dots & \gamma_q \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{q+1} \\ \gamma_{q+2} \\ \vdots \\ \gamma_{q+p} \end{bmatrix},$$

其中, $\gamma_y(k) = \frac{1}{N} \sum_{j=1}^{N-k} y_j y_{j+k}, k = 0, 1, \dots, p$

$$\begin{bmatrix} -\theta_1 \\ -\theta_2 \\ \vdots \\ -\theta_p \end{bmatrix} = \begin{bmatrix} \gamma_y(0) & \gamma_y(1) & \dots & \gamma_y(q-1) \\ \gamma_y(1) & \gamma_y(0) & \dots & \gamma_y(q-2) \\ \vdots & \vdots & \dots & \vdots \\ \gamma_y(q-1) & \gamma_y(q-2) & \dots & \gamma_y(0) \end{bmatrix}^{-1} \begin{bmatrix} \gamma_y(1) \\ \gamma_y(2) \\ \vdots \\ \gamma_y(q) \end{bmatrix}$$

其中, $\gamma_y(k) = \sum_{j=0}^p \varphi_j \varphi_{j+k} \gamma_{k+j-l}, k = 0, 1, 2, \dots, q, \varphi_0 = -1$

平均相对误差 = $\frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right|, x_i, \hat{x}_i$ 分别为数据的观测值和预测值。

2.2 非平稳时间序列模型

如果一个随机过程是齐次非平稳时间序列的时候, 首先运用差分法将随即序列平稳化, 然后再运用平稳时间序列的方法去完成。这种对非平稳时间序列先进行平稳化, 然后再用 ARMA 模型去拟合的模型成为 ARMA(p, d, q) 模型, d 为非平稳时间序列差分的次数。这是一种精度较高的时间序列短期预测模型。

2.3 模型识别和模型检验

模型识别就是通过确定 p, q 的值, 再根据时间序列的自相关函数 p_k 和偏相关函数 φ_k 的统计特征, 运用相关分析方法判断时间序列 $\{y_t\}$ 适合的自回归和滑动平均的阶数 p, q 。参数估计是确定 φ_i, θ_i 等值, 通过软件所提供的参数值模型检验的方法有若干种, 一旦建立了 ARMA 模型并且给出预测值后, 就可以根据以下三个原则进行综合判断模型的优劣取舍: 第一是信息准则, 即 AIC 越小越好; 第二是模型的拟合度 C 和相关系数 R 应相对大, 平均相对误差应相对小; 第三是残差序列 $\{\varepsilon_t\}$ 应为白噪声。

3 黑龙江省大豆综合比较优势趋势的预测

根据《中国统计年鉴》和《黑龙江省统计年鉴》有关全国大豆生产和黑龙江省大豆种植面积和单

产的数据(表1),利用公式(1)、(2)、(3)计算得到黑龙江省大豆比较优势的时间序列(表2)。

表1 黑龙江省及全国大豆播种面积与单产

Table 1 Sowing area and yield of soybean in Heilongjiang province and all over the nation

年份 Year	黑龙江 Heilongjiang province		全国 China	
	面积 Planting area/ $\times 10^4$ hm ²	单产 Yield/kg·hm ⁻²	面积 Planting area/ $\times 10^4$ hm ²	单产 Yield/kg·hm ⁻²
1985	179.5	1618.4	771.8	1251.0
1986	216.7	1447.6	829.5	1265.8
1987	219.7	1720.5	844.5	1375.3
1988	240.0	1597.2	812.0	1500.5
1989	242.9	1582.5	805.7	1445.3
1990	226.4	1288.9	756.0	1352.8
1991	209.4	1515.8	916.3	1361.0
1992	216.0	1638.9	898.3	1393.8
1993	297.9	1696.2	1237.7	1575.8
1994	279.6	1905.6	1273.6	1645.4
1995	251.3	1783.5	1123.2	1591.4
1996	216.1	2015.7	1054.3	1698.1
1997	239.4	2459.1	1116.4	1680.0
1998	246.0	1864.2	1167.1	1714.2
1999	215.3	2202.0	1119.0	1692.6
2000	286.8	1707.1	1266.0	1587.7
2001	287.4	1870.2	1326.8	1547.2
2002	263.1	2321.1	1254.3	1786.8
2003	324.2	1900.4	1289.9	1649.4
2004	340.1	2139.1	1279.9	1744.0
2005	421.5	1899.6	1290.1	1672.5
2006	424.6	1623.4	1214.9	1649.3
2007	380.9	1384.3	1178.0	1460.2
2008	397.2	1679.3	1211.8	1686.2
2009	486.3	1271.8	1194.9	1615.4
2010	447.9	1343.8	1127.6	1681.9

表2 黑龙江省大豆比较优势分析时间序列

Table 2 Analysis comparative advantage of soybean in Heilongjiang by time series

年份 Year	规模优势 Scale advantage	效率优势 Efficiency advantage	综合优势 Comparative advantage
1985	3.51	1.99	2.64
1986	5.07	1.59	2.84
1987	3.91	1.84	2.68
1988	4.73	1.54	2.70
1989	4.66	1.58	2.71
1990	4.58	1.52	2.64
1991	3.46	1.48	2.26
1992	3.62	1.46	2.30
1993	3.52	1.41	2.22
1994	3.21	1.37	2.09
1995	3.28	1.37	2.12
1996	2.96	1.36	2.01
1997	3.01	1.65	2.24
1998	2.97	1.32	1.98
1999	2.69	1.54	2.03
2000	3.13	1.41	2.10
2001	2.89	1.55	2.11
2002	2.78	1.52	2.06
2003	3.18	1.56	2.23
2004	3.29	1.49	2.21
2005	3.45	1.45	2.23
2006	3.48	1.30	2.13
2007	3.16	1.23	1.97
2008	3.19	1.28	2.02
2009	3.38	1.16	1.98
2010	3.22	1.07	1.86

3.1 大豆综合比较优势时间序列模型的建立

以1985~2010年26年的综合比较优势的数据建立模型,其中前23年的数据进行预测,后3年的数据进行检验。利用DPS v7.05软件,根据ARMA模型阶次的选择原理,选用模型,并得到表3。

3.2 模型检验

以2008~2010年3年的实际数据同4个模型的预测值进行计算,得表4,对表3的相对误差进行比较。模型ARMA(211)的拟合度和相关系数较好,且相对误差绝对值为2.9308%,在误差范围允许(5%)之内^[5],而另外3组,其中ARMA(101)和ARMA(201)的拟合度虽然为100%,但是他们的相对误差较大;ARMA(111)相对误差虽然较小,但是

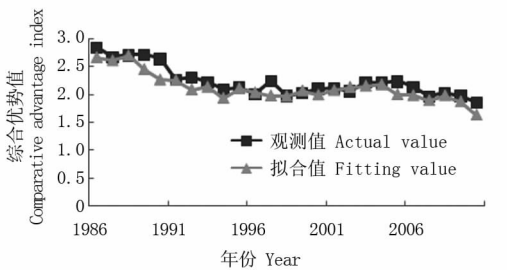


图1 比较优势指数

Fig. 1 Comparative advantage index

表 3 ARMA 模型的参数估计
Table 3 Parameter estimation for ARMA model

模型 Model	AICAIC	相关系数 Correlation coefficient	拟合度 Fitting degree/%
ARMA(101)	-1013.60	1.00	100.00
ARMA(111)	-88.36	0.41	16.74
ARMA(201)	-965.73	1.00	100.00
ARMA(211)	-112.34	0.91	83.69

表 4 ARMA 模型的相对误差
Table 4 Relative error of ARMA model

年份 Year	实际值 Actual value	ARMA(101) 预测值 Predictive value of ARMA (101)	ARMA(111) 预测值 P-value of ARMA (111)	ARMA(201) 预测值 P-value of ARMA (201)	ARMA(211) 预测值 P-value of ARMA (211)
2008	2.02	2.22	1.92	2.40	2.01
2009	1.98	2.19	1.90	2.21	1.83
2010	1.86	2.17	1.87	2.35	1.87
相对误差 Relative error		0.12	0.03	0.19	0.03

其拟合度过低,故上述 3 种模型不符合,所以选定 ARMA(211)为大豆综合比较优势的预测模型。模型拟合曲线为图 1,其模型为 $Z(T+L)=-0.5840-0.2262574 Z(T+L-1)+\exp(T+L)-2.6387530 \exp(T+L-1)$,且模型为白噪声。

3.3 运用模型进行预测

建立的模型 ARMA(211)通过了误差检验和稳定性检验,所以运用 ARMA(211)模型进行预测是可行的。利用 ARMA(211)模型对黑龙江省 2012~2017 年大豆综合比较优势进行预测,预测结果为1.88、1.73、1.76、1.61、1.64。

4 结束语

由模型的预测结果可知,2012~2017 年黑龙江省种植大豆的综合比较优势优于全国的平均水平,但是比较优势指数呈下降趋势,原因可能由于进口转基因大豆带来的冲击,导致部分农民对大豆的种植热情降低,致使黑龙江省的大豆种植面积有一定的减少。因此建议国家有关部门对黑龙江大豆科研和生产给予足够的重视,加大政策支持和科技投入力度,在保证

一定种植面积的基础上,进一步提高大豆的单产和品质,进而保持黑龙江省的大豆种植优势。

参考文献

- [1] 樊敏,顾兆林. 时间序列分析在大气环境中的应用[J]. 资源环境与发展,2010(1):19-22. (Fan M,Gu Z L. The application of time series analyze for atmosphere environment[J]. Sources Environment and Development,2010(1):19-22.)
- [2] 尹昌斌,陈印军,毕于运. 红黄壤地区粮食生产的区域比较优势测度[J]. 农业技术经济,1998(5):42-45. (Yin C B,Chen Y J,Bi Y Y. Comparative advantage measurement degree of food production in red and yellow area[J]. Journal of Agrotechnical Economics,1995(5):42-45.)
- [3] 李庆华. 计量经济学[M]. 北京:中国经济出版社,2005. (Li Q H. Econometrics [M]. Beijing: China Economic Publishing House,2005.)
- [4] 白万平. 经济时间序列模型:方法与应用[M]. 北京:中国商务出版社,2005. (Bai W P. Economic time series model:method and application[M]. Beijing:China Commerce and Trade Press,2005.)
- [5] 卢二坡. 我国能源需求预测模型研究[J]. 统计与决策,2005(10):29-31. (Lu E P. Research on national energy need prediction model[J]. Statistics and Decision,2005(10):29-31.)