

组合预测模型在农业经济研究中的应用——以吉林省粮食产量为例

张红芹¹, 王波², 高来斌¹ (1. 吉林农业大学信息技术学院, 吉林长春 130118; 2. 吉林大学商学院, 吉林长春 130021)

摘要 根据最优加权理论来确定权重, 将时间序列模型、指数平滑模型和 GM(1,1) 预测模型加以组合, 利用吉林省 1952~2007 年的粮食产量统计数据, 建立了吉林省粮食产量的组合预测模型; 并预测出了吉林省 2008~2015 年的粮食产量和年均增长率。

关键词 组合预测; 时间序列模型; GM(1,1) 预测模型

中图分类号 S11+9 文献标识码 A 文章编号 0517-6611(2008)22-09779-04

Application of Composition Forecasting Model in the Agricultural Economy Research

ZHANG Hong qin et al (College of Information and Technology Jilin Agricultural University, Changchun, Jilin 130118)

Abstract Based on the theory of optimum weighted composition modelling, the forecasting model which combined the time series model, the exponential smoothing model and GM(1,1) forecasting model was created with the data of the grain production during 1952 to 2007. At the same time, the future grain production and the rate was forecast from 2008 to 2015, which would offer the scientific basis for the development plan.

Key words Combination forecasting model; Time series model; GM(1,1) forecasting model

自从 20 世纪 60 年代 Bates 和 Granger 首次提出组合预测理论以来, 组合预测方法的研究和应用得到了快速发展^[1]。组合模型可以克服单一模型的局限性, 能够有效地集结更多的有用信息, 因而组合预测方法尤其适用于信息不完备的复杂经济系统^[2]。在经济系统预测中也有很多采用组合预测方法的实践工作。如李秉祥提出了一种基于模糊神经网络的企业财务危机非线性组合预测方法, 通过与其他模型比较, 认为该方法预测精度高, 学习泛化能力强, 适应性广^[3]。赵海青在电力负荷预测中使用一种神经网络优选的组合预测模型, 通过试验例证表明, 该模型具有较强的姿势性和较高的预测精度^[4]。唐小我等提出了具有较强适用性的非负最优组合权重方法^[5]。为此, 笔者利用最优组合权重的方法, 将时间序列、二次指数平滑和 GM(1,1) 模型加以组合^[6], 克服了选用回归模型对预测产生 2 次误差的缺点; 并利用 1952~2007 年的粮食产量数据, 选用最优加权组合预测模型, 来预测吉林省 2008~2015 年的粮食产量, 旨在为吉林省制定农业发展策略提供科学依据。

1 组合预测模型的基本原理^[7]

1.1 组合预测模型 假设对同一预测问题有 n 个预测模型, 这 n 个预测模型分别记为 f_1, f_2, \dots, f_n 。再根据对各单一预测模型预测结果的分析, 确定出各单一预测模型在组合预测模型中的最优权 $K_i (i=1, 2, \dots, n)$, 这样就构成了组合预测模型。

$$f_t = \sum_{i=1}^n K_i f_{it}$$

式中, f_t 为 t 时刻组合预测的预测值; f_{it} 为 t 时刻第 i 种预测方法的预测值; n 为预测方法的个数; K_i 为第 i 种预测方法的权重, $\sum_{i=1}^n k_i = 1$ 。

1.2 权重的确定 在组合预测中, 权重的选择十分重要, 合理的权重会大大提高预测精度。权重选择方法有算术平均法、标准差法、方差倒数法、均方倒数法、层次分析 (AHP) 法、德尔菲法、最优加权法等。该文选取最优加权法。

设 t 时刻实际观测值为 $y_t (t=1, 2, \dots, m)$, 则预测误差 e_{it} (第 i 种方法在 t 时刻预测误差, $i=1, 2, \dots, n$) 为:

$$e_{it} = y_t - f_{it}$$

组合预测误差 e_t (t 时刻预测误差, $t=1, 2, \dots, m$) 为:

$$e_t = y_t - f_t = \sum_{i=1}^n K_i e_{it} = [K_1, K_2, \dots, K_n] [e_{1t}, e_{2t}, \dots, e_{nt}]^T$$

令:

$$K = [K_1, K_2, \dots, K_n]^T$$

$$E_i = [e_{i1}, e_{i2}, \dots, e_{im}]^T$$

式中, E_i 为第 i 个模型的预测误差向量。则组合预测模型的误差矩阵为:

$$e = [E_1, E_2, \dots, E_n]$$

组合预测模型的误差的平方和为:

$$E = e^T e$$

$$R = [1, 1, \dots, 1]^T$$

则组合预测模型的权系数向量由下式计算:

$$K = \frac{E^{-1} R}{R^T E^{-1} R}$$

式中, E^{-1} 为矩阵 E 的逆矩阵。

这种方法保证组合预测模型中误差平方和大的模型权系数较小, 误差平方和小的模型权系数较大, 使组合预测模型的误差平方和尽可能地小。

2 吉林省粮食产量预测模型

吉林省作为农业大省, 是我国重要的粮食主产区, 吉林省的粮食生产对我国的粮食安全具有举足轻重的作用。因此, 能够准确地预测吉林省的粮食产量对吉林省的农业发展, 乃至整个国家的农业发展具有非常重要的意义。该文利用 1952~2007 年的粮食产量数据分别建立时间序列模型、GM(1,1) 模型和二次指数平滑模型, 最后用组合预测模型对吉林省 2008~2015 年的粮食产量进行预测。

2.1 时间序列预测模型^[8] 时间序列分析方法是由美国著名统计学家博克斯 (Box) 和英国的詹金斯 (Jenkins) 于 1976 年提出的, 因此这种方法也被称为博克斯—詹金斯 (Box-Jenkins), 简称 BJ 方法, 广泛应用于各个领域。BJ 方法是一种时间序列短期预测方法, 强调“让数据自己说话”, 重点分析时间序列本身的随机性。

把原始数据 y 取对数后一次差分, 得到的序列为平稳序

基金项目 吉林省科技厅资助 (合同编号 20070610)。

作者简介 张红芹 (1971-), 女, 河北保定人, 讲师, 从事应用数学方面的研究。

收稿日期 2008-04-28

列,用EViews5.0^[9]建立时间序列模型如下:

$$\log(y) = 0.030\ 798 - 0.437\ 612\ \text{AR}(1) - 0.452\ 050\ \text{AR}(2)$$

(2.73) (- 3.55) (- 3.706)

DW=1.990 F=10.163

从估计结果看,参数估计量在1%的显著性水平下均通过t检验;模型的特征根的倒数等于-0.22+0.64i和-0.22-0.64i,模小于1,在单位圆内。

Q(10)=4.83 < 20.05(10-2-0)=15.507, DW=1.99, AC=-0.89, SC=-0.78,可以认为模型误差序列为非自相关序列,模型通过了检验,可以利用该模型进行预测。

2.2 指数平滑预测模型 当时间序列没有明显的趋势变动时,使用第t周期一次指数平滑就能直接预测第t+1期值。但当时间序列的变动出现直线趋势时,用一次指数平滑法来预测仍存在着明显的滞后偏差。因此也需要进行修正。修正的方法也是在一次指数平滑的基础上再作二次指数平滑,利用滞后偏差的规律找出曲线的发展方向和 trends,然后建立直线趋势预测模型,故称为二次指数平滑法。利用1952~2007年的粮食产量数据,选择二次指数平滑法,EViews5.0运行结果如下:

$$\hat{y}_{t+T} = 2\ 582.464 + 57.329\ 96\ T \quad T=1,2,\dots,n$$

式中,t为当前时期数;T为由当前时期数t到预测期的时期数; \hat{y}_{t+T} 为第t+T期的预测值。

2.3 GM(1,1)预测模型^[10] 灰色预测通过鉴别系统因素之间发展趋势的相异程度,即进行关联分析,并对原始数据进行生成处理来寻找系统变动的规律,生成有较强规律性的数据序列,然后建立相应的微分方程模型,从而预测事物未来发展趋势的状况。其用等时距观测到的反应预测对象特征的一系列数量值构造灰色预测模型,预测未来某一时刻的特征量,或达到某一特征量的时间。

(1) 设时间序列 $X^{(0)}$ 有 n 个观察值:

$$X^{(0)} = [X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)]$$

通过累加生成新序列:

$$X^{(1)} = [X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)]$$

其中

$$X^{(1)}(i) = \sum_{j=1}^i X^{(0)}(j) \quad i=1,2,\dots,n$$

则GM(1,1)模型相应的微分方程为:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = \mu$$

式中,a称为发展灰数; μ 称为内生控制灰数。

(2) 设 \hat{a} 为待估参数向量, $\hat{a} = \frac{a}{\mu}$,可以利用最小二乘法求解。

解得:

$$\hat{a} = (B^T B)^{-1} B^T Y_n$$

其中:

$$z^{(1)}(k) = [X^{(1)}(k) + x^{(1)}(k-1)]/2 \quad k=2,3,\dots,n$$

$$x^{(0)}(2)$$

$$x^{(0)}(3)$$

Y=

...

$$x^{(0)}(n)$$

$$- z^{(2)}(2) \quad 1$$

$$- z^{(1)}(3) \quad 1$$

B=

...

$$- z^{(1)}(n) \quad 1$$

求解微分方程,即可得预测模型:

$$X^{(1)}(k+1) = X^{(0)}(1) - \frac{\mu}{a} e^{-ak} + \frac{\mu}{a} \quad k=0,1,2,\dots,n$$

利用1952~2007年的粮食产量数据,用MATLAB数学软件编程计算结果如下:

$$Y = 10\ 191.122\ 7 \exp(0.035\ 84t) - 9\ 732.172\ 7, t=0,1,2,\dots,n$$

P=1, C=0.1717,按照检验标准,该模型拟合效果非常好,可以用来预测。

2.4 组合预测模型 利用以上3个模型的模拟值,运用最优加权确定权重的方法,通过MATLAB数学软件编程,计算出权重分别为: $K_1=0.080\ 8$, $K_2=0.550\ 6$, $K_3=0.368\ 6$ 。则组合预测模型为:

$$f_t = 0.080\ 8f_{1t} + 0.550\ 6f_{2t} + 0.368\ 6f_{3t}$$

式中, f_t 为t时刻组合预测的预测值; f_{1t} 为t时刻采用时间序列方法的预测值; f_{2t} 为t时刻采用指数平滑方法的预测值; f_{3t} 为t时刻采用GM(1,1)模型的预测值。各单一模型和组合模型对1952~2007年吉林省粮食产量的拟合结果见表1。

表1 1952~2007年吉林省实际粮食产量及拟合值

Table 1 Actual output and fitting value of grain in Jilin Province from 1952 to 2007

×10⁷ kg

年份 Year	实际粮食产量 Actual output of grain	时间序列模型拟合值 Fitting value of time series model	指数平滑模型拟合值 Fitting value of exponential smoothing model	GM(1,1)模型拟合值 Fitting value of GM(1,1) model	组合模型拟合值 Fitting value of combination model
1952	613.20	490.67	518.60	399.51	472.45
1953	561.45	549.73	570.84	414.09	511.35
1954	530.95	542.83	586.41	429.20	524.94
1955	556.53	549.57	584.31	444.86	530.10
1956	493.64	582.69	591.12	461.09	542.51
1957	429.35	598.64	570.80	477.91	538.81
1958	528.84	610.69	531.14	495.35	524.37
1959	526.60	633.88	540.11	513.43	537.85
1960	394.70	655.07	544.78	532.16	549.04
1961	398.55	672.97	497.45	551.58	531.58
1962	437.16	94.53	465.26	571.71	523.02

接下表

续表1

年份 Year	实际粮食产量 Actual output of grain	时间序列模型拟合值 Fitting value of time series model	指数平滑模型拟合值 Fitting value of exponential smoothing model	GM(1,1) 模型拟合值 Fitting value of GM(1,1) model	组合模型拟合值 Fitting value of combination model
1963	501.67	17.26	457.04	592.57	528.02
1964	491.80	738.99	475.53	614.19	547.93
1965	525.10	761.94	484.57	636.61	563.02
1966	597.60	786.18	503.23	659.84	583.82
1967	647.74	810.67	543.38	683.91	616.78
1968	622.15	835.87	589.95	708.87	653.65
1969	498.70	862.13	612.17	734.74	677.54
1970	738.80	889.12	580.22	761.55	672.02
1971	713.05	916.86	647.93	789.33	721.78
1972	556.99	45.56	684.75	818.14	754.99
1973	783.00	975.15	650.45	847.99	749.50
1974	858.15	1 005.63	711.01	878.93	796.71
1975	906.50	1 037.08	780.79	911.01	849.50
1976	755.50	1 069.53	846.59	944.25	900.60
1977	728.35	1 102.98	833.96	978.70	909.05
1978	914.70	1 137.47	813.29	1 014.40	913.61
1979	903.34	1 173.05	867.92	1 051.40	960.20
1980	859.60	1 209.74	900.44	1 089.80	995.23
1981	921.91	1 247.58	905.13	1 129.60	1 015.54
1982	1 000.04	1 286.60	930.45	1 170.80	1 047.82
1983	1 477.98	1 326.84	976.21	1 213.50	1 092.01
1984	1 634.46	1 368.34	1187.28	1 257.80	1 227.91
1985	1 225.26	1 411.14	1391.77	1 303.70	1 360.87
1986	1 397.71	1 455.27	1376.80	1 351.30	1 373.74
1987	1 675.81	1 500.79	1428.01	1 400.60	1 423.79
1988	1 693.25	1 547.73	1565.57	1 451.70	1 522.16
1989	1 351.29	1 596.14	1664.67	1 504.60	1 600.13
1990	2 046.52	1 646.06	1 600.61	1 559.50	1 589.13
1991	1 898.87	1 697.54	1 814.80	1 616.50	1 732.23
1992	1 840.31	7 150.64	1 904.70	1 675.40	1 807.73
1993	1 900.9	1 805.39	1 940.83	1 736.60	1 854.61
1994	2 015.7	1 861.86	1 984.40	1 799.90	1 906.49
1995	1 992.40	1 920.09	2 053.78	1 865.60	1 973.62
1996	2 326.60	1 980.15	2 089.00	1 933.70	2 022.96
1997	1 808.30	2 042.08	2 235.53	2 004.30	2 134.67
1998	2 506.00	2 105.95	2 137.39	2 077.40	2 112.74
1999	2 305.60	2 171.82	2 328.13	2 153.20	2 251.02
2000	1 638.00	2 239.75	2 381.33	2 231.80	2 314.77
2001	1 953.40	2 309.80	2 161.43	2 313.20	2 229.36
2002	2 214.80	2 382.04	2 123.07	2 397.60	2 245.19
2003	2 259.60	2 456.54	2 192.19	2 485.10	2 321.52
2004	2 510.00	2 533.38	2 254.69	2 575.80	2 395.57
2005	2 581.21	2 612.61	2 390.11	2 669.80	2 511.18
2006	2 720.00	2 694.33	2 508.40	2 767.20	2 618.82
2007	2 454.00	2 778.60	2 560.38	2 831.50	2 678.95

注:数据来源于《吉林省统计年鉴》(2007)和吉林省政府网站。

Note: Data came from the Jilin statistical Yearbook(2007) and the website of Jilin government.

3 模型的评价及预测

对于模型优劣的评价采用均方误差,以MSE表示,其定义式如下:

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2$$

使用均方误差方法对单项模型进行评价时,一般所遵循的准则是均方误差最小的方法最优^[11]。均方误差方法主要关注的是有没有大的预测误差,并倾向于否定容易产生大的预测失误的方法。3种单项预测方法以及组合预测方法的均方误差值见表2。从表2可以看出,组合预测模型的均方误差最小,可以认为组合预测模型预测效果最好。因此,可以用组合模型来预测吉林省未来几年的粮食产量。预测结果见表3。

表2 预测模型拟合误差

Table 2 Fitting error of forecast models

预测模型 Forecast model	均方误差 Mean square error
时间序列模型 Time series model	55 400.6
指数平滑模型 Exponential smoothing model	40 807.7
GM(1,1) 预测模型 GM(1,1) forecast model	44 560.4
组合预测模型 Combination forecast model	35 270.6

4 结论

(1) 建立了吉林省粮食产量的组合预测模型,预测出吉林省2008年的粮食产量为2 798.80万t,到2015年吉林省的粮食产量将达到3 377.01万t。根据这一预测结果,在

2008~2015年,吉林省的粮食产量年均增长率为2.72%,这个增长率低于2000~2007年间6.31%的年均增长率。这种增长率由快到慢的增长趋势正好符合农业生产的特点和事物发展的一般规律,为吉林省制定农业发展政策提供了科学的依据。

表3 吉林省2008~2015年的粮食产量

年份 Year	粮食产量 Grain output
2008	2 798.80
2009	2 875.43
2010	2 954.35
2011	3 034.97
2012	3 117.74
2013	3 202.31
2014	3 288.71
2015	3 377.01

(2) 基于组合预测方法构建的吉林省粮食产量预测模型,能够有效地利用各种信息,降低模型的预测误差,有效地改进了单项预测模型的误差,组合模型的拟合均方误差均低于各单项预测模型,整体表现良好。

(3) 粮食生产是相当复杂的过程,这其中影响粮食生产的因素很多,如播种面积、粮食价格、农业政策、气象因素等,

(上接第9778页)

营。采取多种形式,通过发展“订单农业”、合作经济组织和各类行业协会等方式,加速各类经济利益共同体的形成,提高农业生产的组织化程度。

3.6 切实转变政府职能,搞好宏观调控 农业产业化经营不是孤立的个别行为,它涉及到不同产业、不同所有制、不同地区、不同部门的利益关系。这就要求各级政府必须从全局出发,转变政府职能,加强对农业产业化经营的宏观调控:一是搞好引导。结合国家规划从衡水市的实际需要出发,制定农业产业化经营发展规划并具体实施,通过制定并落实有关政策、法规,为农业产业化经营提供政策指导和法律依据;二是搞好协调。通过协调农业产业化经营在地区之间的产业布局,合理调节地区之间的利益关系,通过协调农业产业化经营单位内部各方之间存在的矛盾和问题,合理调节相互之间的利益关系;三是搞好服务。通过健全信息网络提供信息服务和通过健全技术咨询和推广机构提供技术服务以及通过金融体制改革提供政策性的贷款服务。此外,还要在搞好基础设施建设、健全市场体系和规范市场秩序等方面,为农业产业化经营的健康发展创造良好的外部条件。

3.7 整体提高劳动者的素质 人力资源是一个国家或地区发展的重要保障,技能型人才和高素质劳动者是人力资源构成的主体,也是一个国家和地区核心竞争力的重要方

都会直接影响粮食产量。并且,在不同时期粮食生产的影响因素也不相同。因此,根据过去的粮食产量来预测未来的产量,肯定会存在偏差,这可能是所有预测模型的不足之处。对粮食产量的预测要根据诸多因素的变化适时调整模型,以减小预测误差,实现更准确的预测。

参考文献

- [1] MAHMOUD E. Accuracy in forecasting: A survey[J]. Journal of Forecasting, 1984, 3:139-150.
- [2] REDDICK C G. An empirical examination of revenue forecasting technique in local governments[J]. Municipal Finance Journal, 2004, 24:25-48.
- [3] 李秉祥. 上市公司财务危机动态预测模型研究[J]. 数学的实践与认识, 2006(2):60-65.
- [4] 赵海青. 负荷预测的交叉自适应优选组合预测模型[J]. 华北电力大学学报, 2000(4):13-17.
- [5] 唐小我, 马永开, 曾永, 等. 现代组合预测和组合投资决策方法及应用[M]. 北京: 科学出版社, 2003:18-25.
- [6] 丁晨芳. 组合预测方法在我国粮食产量预测中的应用[J]. 农业现代化研究, 2007(1):101-103.
- [7] 樊胜根, 莫塞迪塔·索姆比拉, 刘庆华. 中国未来粮食供求预测的差别[J]. 中国农村观察, 1997(3):17-23, 35.
- [8] 孙元, 吕宁. 地方财政一般预算收入预测模型及实证分析[J]. 数量经济技术经济研究, 2007(1):38-43.
- [9] 易丹辉. 数据分析与EViews应用[M]. 北京: 中国统计出版社, 2002:43-45.
- [10] 毕守东, 王东平. 安徽省粮食产量的最优加权组合预测[J]. 预测, 2000(3):70-72.
- [11] 李宝仁. 经济预测理论、方法及应用[M]. 北京: 经济管理出版社, 2005:9-11.

面,要把提高劳动者素质放在更加突出的位置。因此,推进衡水市农业产业化经营,必须整体提高劳动者的素质。政府要紧密结合该区经济社会发展的需要,统筹该地区职业教育的发展,满足人民群众不同层次和终身学习需求,把农民培养成有文化、懂技术、会经营的新型农民。

4 小结

要想加快贯彻落实好科学发展观,推动经济社会全面、协调、可持续发展,实现社会主义物质文明、政治文明和精神文明共同进步,必须紧紧抓住推进农业产业化这一重大战略举措,坚持理论和实际相结合,因地制宜、因时制宜地把科学发展观的要求贯穿于农业产业化各方面的工作,提高我国农业产业化的经营水平。同时要充分考虑地区之间、部门之间的发展差异和不同情况,坚持一切从实际出发,根据实际条件和发展需要有重点、有步骤地采取措施,特别注重解决当地农业产业化发展中存在的突出矛盾和问题,更快更好地推动农业产业化健康发展。

参考文献

- [1] 江泽民. 中国共产党第十六次全国代表大会文件汇编[M]. 北京: 人民出版社, 2002.
- [2] 胡锦涛. 高举中国特色社会主义伟大旗帜为夺取全面建设小康社会新胜利而奋斗[M]. 北京: 人民出版社, 2007.
- [3] 衡水市农业产业化办公室. 衡水市农业产业化十一·五发展规划[Z]. 2006.
- [4] 国务院. 农业科技发展纲要(2001—2010)[N]. 人民日报, 2001-05-24(05).