

基于包络带模型的万源市富硒茶叶产量预测*

唐海军, 李玲, 何聪

(四川文理学院 数学与财经学院, 四川 达州 635002)

【摘要】根据万源市2004~2010年茶叶产量的统计数据,建立灰色系统的包络带模型,对2011年产量进行预测。通过检验,包络带模型预测的相对误差为1.08%,小于直接采用GM(1,1)预测,模型精度达到了较高要求,能为促进万源市富硒茶叶业持续健康发展提供依据。

【关键词】包络带预测;茶叶产量;精度检验

【中图分类号】S571.1;O29 **【文献标识码】**A **【文章编号】**1673-1891(2014)03-0020-03

1 引言

万源市地处大巴山腹地,与陕、渝、鄂相邻,属于亚热带湿润季风气候区。由于万源土地含人体所须的硒元素,富硒农产品已经成为当地的农业特色产业,所产富硒茶“巴山雀舌”闻名全国。2008年4月,万源市被中国食品工业协会授予“中国富硒茶之都”称号^[1]。近几年来,当地政府非常重视富硒农产品基地建设,加大了对茶叶育苗、生产、加工、销售的投入力度,茶叶种植面积和茶叶产量快速增长。到2011年末,茶园面积达到9353公顷,茶叶产量2835吨,其中的优质名茶1576吨,名优率达到55.6%。茶叶已经成为“万源市现代农业‘千亿工程’项目的重点行业,也成为振兴革命老区的支柱产业”^[2]。茶叶易受到自然灾害、种植面积、茶园虫害、管理技术等因素影响,茶叶的产量会发生波动。基于此,利用已有统计数据,建立灰色数理模型,对万源茶叶产量进行包络带预测与评价。灰色理论最早是邓聚龙教授创立,这门理论在科学预测方面发展很迅速、应用广泛,比如工业、经济和自然现象等领域^[3]。

2 建模过程

2.1 数据引用

从万源市2004~2011年的国民经济和社会发展统计公报中获得,2004~2011年万源市富硒茶叶的年产量数据(见表1)。

表1 万源市2004~2011年茶叶产量统计表

年份	2004	2005	2006	2007	2008	2009	2010	2011
产量(吨)	2160.3	2210	2309.7	2550	2500	2725	2810	2835

下面以2004~2010年数据进行灰色模拟预测,得到2011年预测值,再进行检验,比较模型精度。

2.2 灰色包络预测

对于序列中的原始数据比较离乱,模拟的模型在难以通过精度检验的时候,可以考虑只“给出其未来变化的范围,预测出它的取值区间”^[4]。如果采集的数据上、下波动较大,就需要采用包络线灰色预测方法进行预测。对各组数据数列分别建立常规GM(1,1)模型^[5]。

定义1 设 $X^{(0)}$ 为原始序列, $x_u^{(0)}$ 是 $X^{(0)}$ 的下缘点连线所对应的序列, $x_s^{(0)}$ 是 $X^{(0)}$ 上缘点连线所对应的序列,分别采用GM(1,1)预测得到

$$\hat{x}_u^{(1)}(k+1) = \left(x_u^{(0)}(1) - \frac{b_u}{a_u} \right) \exp(-a_u k) + \frac{b_u}{a_u} \quad \text{和}$$

$$\hat{x}_s^{(1)}(k+1) = \left(x_s^{(0)}(1) - \frac{b_s}{a_s} \right) \exp(-a_s k) + \frac{b_s}{a_s}$$

分别为 $x_u^{(0)}$ 和 $x_s^{(0)}$ 对应的GM(1,1)时间响应式,则称 $S = \{(t, X(t)) | X(t) \in [\hat{x}_u^{(1)}(t), \hat{x}_s^{(1)}(t)]\}$ 为包络带。

2.2.1 作序列的上、下包络线

先根据数据在画图软件中画出折线 $X^{(0)} = (2160.3, 2210, 2309.7, 2550, 2500, 2725, 2810)$,再勾勒出折线的上包络曲线 $f_s(t)$ 与下包络曲线 $f_u(t)$ 的轮廓。构造上包络曲线轮廓,一般取序列 x 的峰点,如果峰点不是等间隔的,就采取观察、计算平均数、经验指定插补点以及函数模拟,使其变为等间隔的,同理,可以作出下包络曲线轮廓。

为了进一步准确模拟数据的发展趋势,根据原始数据生成图像(见图1)的变化,选取 $X^{(0)}$ 中的 $(x(1), x(4)) = (2160.3, 2550)$ 作为上包络序列的已知数据;选取 $X^{(0)}$ 中的 $(x(3), x(5)) = (2309.7, 2500)$ 作为下包络序列的已知数据。在作等间隔上下包络序列时,再引入函数 $y = ax^b \ln x + c$ ($a, b, c \in R^+$),考虑到数据变化趋势是递增,且图像呈现上凸走势,可令 $b = \frac{1}{3}$,来对上下包络进行模拟。

对于上包络,通过Excel模拟得到 $y = 177.088x^{\frac{1}{3}} \ln x$

收稿日期:2014-05-25

*基金项目:2014年四川省教育厅资助课题“灰色数理模型在茶叶产业开发中的应用研究—以万源市富硒茶开发为例”(项目编号:14ZB0312)。

作者简介:唐海军(1982-),男,四川南充人,讲师,硕士,主要从事数学教育与应用数学研究。

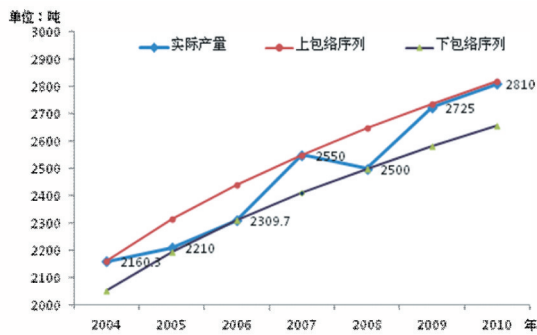


图1 茶叶产量统计图

+2160.3 ($x \in N^+$), 上包络线 $f_s(t)$ 所对应的上包络序列为 $X_s^{(0)} = (x_s^{(0)}(1), x_s^{(0)}(2), x_s^{(0)}(3), x_s^{(0)}(4), x_s^{(0)}(5), x_s^{(0)}(6), x_s^{(0)}(7)) = (2160.3, 2314.953, 2440.891, 2550, 2647.664, 2736.871, 2819.491)$

同理对于下包络,通过 Excel 模拟得到 $y = 162.98x^3 \ln x + 2051.462$ ($x \in N^+$), $f_u(t)$ 所对应的下包络序列可以构建为

$X_u^{(0)} = (x_u^{(0)}(1), x_u^{(0)}(2), x_u^{(0)}(3), x_u^{(0)}(4), x_u^{(0)}(5), x_u^{(0)}(6), x_u^{(0)}(7)) = (2051.462, 2193.795, 2309.7, 2410.117, 2500, 2582.1, 2658.138)$

2.2.2 上下包络序列建模预测

下面先对上包络序列建模:首先为了保证建模方法的可行性,需要对 $f_s(t)$ 所对应的上包络序列的数据做必要的检验处理。

(1) 其 1-AGO 序列为

$X_s^{(1)} = (x_s^{(1)}(1), x_s^{(1)}(2), x_s^{(1)}(3), x_s^{(1)}(4), x_s^{(1)}(5), x_s^{(1)}(6), x_s^{(1)}(7)) = (2160.3, 4475.253, 6916.144, 9466.144, 12113.808, 14850.679, 17670.17)$

(2) 对 $x_s^{(0)}$ 作准光滑性检验。由

$$\rho(k) = \frac{x^{(0)}(k)}{x^{(0)}(k-1)}, k = 2, 3, \dots, n$$

$\rho(2) \approx 1.07, \rho(3) \approx 0.55, \rho(4) \approx 0.37 < 0.5, \rho(5) = 0.28 < 0.5, \rho(6) \approx 0.226 < 0.5, \rho(7) \approx 0.19 < 0.5$

当 $k > 3$ 时,准光滑条件满足。

(3) 准指数律检验与判断。

由 $\sigma^{(1)}(k) = \frac{x^{(1)}(k)}{x^{(1)}(k-1)}, k = 2, 3, \dots, n$ 得

$\sigma^{(1)}(2) \approx 2071, \sigma^{(1)}(3) \approx 1.545, \sigma^{(1)}(4) \approx 1.37, \sigma^{(1)}(5) \approx 1.28, \sigma^{(1)}(6) \approx 1.226, \sigma^{(1)}(7) = 1.19.$

当 $k > 2$ 时, $\sigma^{(1)}(k)$ 都落在可容覆盖 $[1, 1 + \delta], \delta = 0.5$ 内,故可用序列 $x^{(1)}$ 作较为满意的 GM(1,1) 模型进行灰色预测。

其次,建立 GM(1,1) 模型.通过对 $x_s^{(1)}$, 构造数据矩阵 B 及数据向量 Y (见文献 4), 借助 matlab 软件计算得到: $\hat{u} = [a, b]^T = [-0.0386, 2213.8]^T$, 建立模型:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b$$

及时间响应式: $\hat{x}_s^{(1)}(k) = (x_s^{(0)}(1) - \frac{b}{a})e^{-a(k-1)} + \frac{b}{a} =$

$$59512.6316 e^{0.0386(k-1)} - 57352.3316$$

由此可得 $x_s^{(0)}$ 的累减还原式:

$$\hat{x}_s^{(0)}(k) = \hat{x}_s^{(1)}(k) - \hat{x}_s^{(1)}(k-1) = 2253.4168 e^{0.0386(k-1)}$$

从而有 2004 ~ 2010 年模拟值

$$\hat{X}_s^{(0)} = (\hat{x}_s^{(0)}(1), \hat{x}_s^{(0)}(2), \hat{x}_s^{(0)}(3), \hat{x}_s^{(0)}(4), \hat{x}_s^{(0)}(5), \hat{x}_s^{(0)}(6), \hat{x}_s^{(0)}(7)) = (2160.3, 2342, 2434.1, 2529.8, 2629.3, 2732.7, 2840.2)$$

以及 2011 最高预测值 $\hat{x}_s^{(0)}(8) = 2951.9$

同理根据 $f_u(t)$ 所对应的下包络序列为:

$$X_u^{(0)} = (x_u^{(0)}(1), x_u^{(0)}(2), x_u^{(0)}(3), x_u^{(0)}(4), x_u^{(0)}(5), x_u^{(0)}(6), x_u^{(0)}(7)) = (2051.462, 2193.795, 2309.7, 2410.117, 2500, 2582.1, 2658.138)$$

经过检验,作出 $x_u^{(0)}$ 的 GM(1,1) 时间响应式,从而有 2004-2010 年模拟值

$$\hat{X}_u^{(0)} = (\hat{x}_u^{(0)}(1), \hat{x}_u^{(0)}(2), \hat{x}_u^{(0)}(3), \hat{x}_u^{(0)}(4), \hat{x}_u^{(0)}(5), \hat{x}_u^{(0)}(6), \hat{x}_u^{(0)}(7), \hat{x}_u^{(0)}(8)) = (2051.5, 2218.5, 2303.4, 2391.7, 2483.3, 2578.4, 2677.1)$$

以及 2011 最低预测值 $\hat{x}_u^{(0)}(8) = 2779.6$

对上包络序列的预测值是预测区间的上限,对下包络序列的预测值是预测区间的下限^[5]。取两者平均数作为基本模拟预测值:

$$\hat{X}^{(0)} = (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \hat{x}^{(0)}(4), \hat{x}^{(0)}(5), \hat{x}^{(0)}(6), \hat{x}^{(0)}(7), \hat{x}^{(0)}(8)) = (2105.9, 2280.25, 2368.75, 2460.75, 2556.3, 2655.55, 2758.65, 2865.75)$$

3 模型的精度检验

模型的精度检验主要采用残差检验、平均相对误差。

$$\text{残差} = \varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$$

$$(-70.25, -59.05, 89.25, -56.3, 69.45, 51.35) \quad (k = 2, 3, \dots, 7)$$

$$\text{相对误差} = \Delta(k) = \frac{|x^{(0)}(k) - \hat{x}^{(0)}(k)|}{x^{(0)}(k)}$$

$$(3.17\%, 2.55\%, 3.5\%, 2.25\%, 2.54\%, 1.82\%) \quad (k = 2, 3, \dots, 7)$$

2005 ~ 2010 年的平均相对误差为 2.64% < 0.05, 符号建模精度等级要求^[4]。

对原始数据 $x^{(0)}$ 直接通过 GM(1,1) 模型预测得到 2011 年产量为 2962.2 吨,通过灰色包络带模型预测得到该年的产量为 2865.75 吨,该年实际富硒茶叶

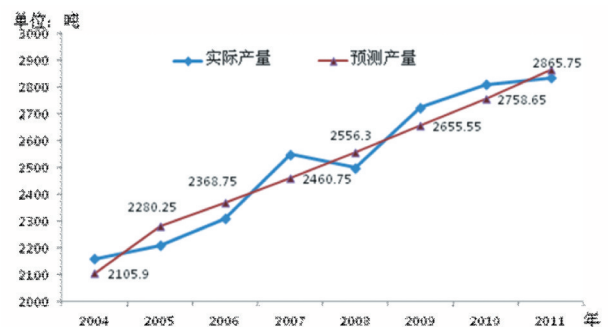


图2 茶叶实际产量与预测产量对比图

的产量为2835吨。可见灰色包络带模型预测的相对误差为1.08%，小于直接灰色预测的相对误差4.49%。如果 $\Delta(k) < 0.2$ ，则可认为建模的模型精度达到了较高要求，可进行预测和预报。

通过观察图1、图2，从2004~2011年万源市茶叶产量呈整体上升趋势，实际产量在2007，2009~2010年期间上升趋势明显，并且较程度的高于预测值。说明自2007年以来，该市茶叶发展得到政府的大力扶持，2010年3月，万源富硒茶获农业部批准使用“中国地理标志”，2011年“强力推进茶叶万亩核心示范区建设，发展现代农业产业基地，打造茶产业强县”战略^[7]等利好因素不断增加，促进了万源市种茶业的发展，结论与事实相符，直观的证明了模型的合理性。

4 结论

通过灰色包络带模型分析万源市近十年来茶叶产量的发展变化趋势，可以得到如下结论：

(1)万源市富硒茶叶产量一直处于增长阶段，2009年后由于“中国富硒茶都—四川万源天然富硒茶暨生态旅游推介会”、“中国富硒食品展示展销会”、“一乡一业”、“一村一品”等措施的促进，产量增加数明显高于以前的预测数据。面对产量的迅速增加，当地政府已经制定出合理的富硒茶叶发展战略规划，加大了富硒茶的生产、加工、销售等配套设施建设投资力度，致力种茶叶的健康持续发展，为提升川东北革命老区的农民的经济收入作出贡献。

(2)预测的产量增长趋势是建立在精确度可靠的数学模型基础上，结论可以为当地政府部门产业规划与发展决策，以及特色优势农业资源的开发与规模化发展提供科学依据。利用灰色GM(1,1)模型与包络带模型对万源茶叶发展状况进行的定量分析，预测2014年即以后该市茶叶产量仍将保持良好的增长势头，这有利于以特色富硒农产品为支撑产业的山区农民脱贫致富。

注释及参考文献：

- [1]李传松,张德涛,陈显铭,等.关于把万源打造成中国硒茶之都的调查与思考[J].达州新论,2008(S1):103-107.
- [2]唐海军,张瑞,何聪.万源市茶叶产量的发展带预测[J].西华师范大学学报,2013,34(4):418-421.
- [3]S.Liu,Y.Dang,Z.Fang.Grey Systems Theory and its Applications,third ed,Science Press,Beijing,2004.
- [4]刘思峰,党耀国,方志耕,等.灰色系统理论及其应用[M].北京:科学出版社,2010:231-232,226-227.
- [5]邓美兰,张珍.奥运会男子十项全能前三名成绩的灰色预测模型的研究[J].数学的实践与认识,2009,39(18):44-49.
- [6]张亮,赵云.图书馆采购文献的灰色预测模型构建[J].情报杂志,2007(6):131-133.
- [7]四川省统计局.关于万源市茶叶产业发展情况的调研报告[EB/OL].http://govinfo.nlc.gov.cn2012-09-14.

The Forecast of Wanyuan's Selenium-rich Tea Production Based on Model of Envelope

TANG Hai-jun, LI Ling, HE Cong

(College of Mathematics and Finance-Economics, Sichuan University of arts and science, Dazhou, Sichuan 635002)

Abstract: Setting up a model of grey envelope and forecasting selenium-rich tea production of Wanyuan in 2011, the research is made according to the statistics about output of tea from 2004 to 2010. By examination, the conclusion is that relative error to model of grey envelope is 1.08%, which inferior to forecast by GM(1,1). Model precision qualifies as high requirement, it provides bases for promoting selenium-rich tea of Wanyuan development sustainably and healthily.

Key words: envelope forecast; tea's output; accuracy test