

海洋产业发展趋势分析^{*}

——基于熵值法的组合预测

赵昕¹ 余亭²

(1. 华中科技大学管理学院 武汉 430000; 2. 中国海洋大学经济学院 青岛 266003)

摘要 为克服单项预测方法可能产生误差的问题,文章利用信息论中熵的概念,计算出组合预测加权平均系数,用灰色预测和指数预测方法对海洋产业发展趋势进行了组合预测分析,提高了预测的精度,为决策提供了可靠的依据。

关键词 海洋产业;灰色预测;指数预测;组合预测

我国是海洋大国,海洋资源可开发利用的潜力很大。加快发展海洋产业,促进海洋经济发展,对形成国民经济新的增长点,实现全面建设小康社会目标具有重要意义。2008年全国海洋生产总值29 662亿元,比上年增长11.0%,但海洋经济占国内生产总值的比例仍然很低,因此,对海洋产业未来总体发展趋势的预测意义重大。

不同的定量预测模型各有其优点和缺点,它们之间并不是相互排斥,而是相互联系且相互补充的。由于每种预测方法利用的数据不尽相同,不同的数据从不同的角度提供各方面有用的信息。Bates和Granger于1969年提出了组合预测方法的概念^[1],即综合考虑各单项预测方法的特点,将不同的单项预测方法以适当的加权平均形式进行组合,得出组合预测模型。

信息论的观点认为:熵值是系统无序程度或混乱程度的度量,系统的熵值越大,则所蕴涵的信息量越小;系统的某项指标的变异程度越小;反之,

系统的熵值越小,则所蕴涵的信息量越大,系统的某项指标的变异程度越大^[2]。本研究从这一观点出发,根据各单项预测方法预测误差序列的变异程度,利用信息熵的概念,计算出组合预测加权平均系数,然后对海洋经济的发展趋势进行误差预测,以期能获得较好的预测结果。

1 确定组合预测加权系数的熵值法

设对同一预测对象的某个指标序列为 $\{x_t, t=1, 2, \dots, N\}$,存在 m 种单项预测方法对其进行预测,设第 i 种单项预测方法在第 t 时刻的预测值为 $x_{it}, i=1, 2, \dots, m, t=1, 2, \dots, N$ 。

$$\text{令 } e_{it} = \begin{cases} 1 & |(x_t - x_{it}) / x_t| > 1 \\ |(x_t - x_{it}) / x_t| & 0 \leq |(x_t - x_{it}) / x_t| \leq 1 \end{cases} \quad (1)$$

则称 e_{it} 为第 i 种预测方法在第 t 时刻的预测相对误差, $i=1, 2, \dots, m; t=1, 2, \dots, N$ 。

显然, $0 \leq e_{it} \leq 1, \{e_{it}, i=1, 2, \dots, m; t=$

* 基金项目:国家海洋局我国近海海洋综合调查与评价(908-02-07-02)子课题海洋产业布局研究。

1, 2, ..., N} 为第 i 种预测方法在各个 t 时刻的预测相对误差序列。

用熵值法确定组合预测加权系数的步骤有以下 5 步^[3]：

1) 将各种单项预测方法预测相对误差序列归一化。即计算第 i 种单项预测方法第 t 时刻的预测相对误差的比重。

$$p_{it} = \frac{e_{it}}{\sum_{i=1}^N e_{it}}, t = 1, 2, \dots, N \quad (2)$$

可见有： $\sum_{i=1}^N p_{it} = 1, i = 1, 2, \dots, m$ 。

2) 计算第 i 种单项预测方法的预测相对误差的熵值 h_i 。

$$h_i = -k \sum_{t=1}^N p_{it} \ln p_{it}, i = 1, 2, \dots, m \quad (3)$$

$k > 0$ 为常数，这里取 $k = 1/\ln N$ ，则有 $0 \leq h_i \leq 1$ 。

3) 计算第 i 种单项预测方法的预测相对误差序列的变异程度系数，因为 $0 \leq h_i \leq 1$ ，根据系统某

项指标的熵值的大小与其变异程度相反的原则，所以定义第 i 种单项预测方法的预测相对误差序列的变异程度系数 d_i ，则 $d_i = 1 - h_i, i = 1, 2, \dots, m$ 。

$$4) \text{ 计算各种预测方法的加权系数 } l_i, l_i = \frac{1}{m-1} \left(1 - \frac{d_i}{\sum_{i=1}^m d_i}\right), i = 1, 2, \dots, m \quad (4)$$

式 (4) 体现了—个原则，即某个单项预测方法预测误差序列的变异程度越大，则其在组合预测中对应的权系数就越小。显然权系数满足 $\sum_{i=1}^m l_i = 1$ 。

5) 计算组合预测值

$$\hat{x}_t = \sum_{i=1}^m l_i x_{it}, t = 1, 2, \dots, N \quad (5)$$

2 实例分析

组合预测是以单项预测为基础的，因此，本研究首先对 1999—2008 年的全国海洋生产总值分别用灰色预测和指数预测方法进行预测，得到的结果如表 1 所示。

表 1 海洋生产总值、预测值及误差百分比

年份	真实值/亿元	灰色预测/亿元	$e_{1t}/(\%)$	指数预测/亿元	$e_{2t}/(\%)$
1999	3 651.30	4 311.876 6	0.180 9	2 720	0.255 1
2000	4 133.50	4 996.780 9	0.208 9	4 180	0.011 3
2001	4 459.92	6 860.354 3	0.538 2	5 970	0.338 6
2002	9 050.29	8 060.698 7	0.109 3	8 100	0.105 0
2003	10 523.40	10 027.914 8	0.047 1	10 560	0.003 5
2004	13 704.76	12 824.868 1	0.064 2	13 370	0.024 4
2005	16 987.00	15 954.781 2	0.060 8	16 520	0.027 5
2006	18 408.00	20 544.869 4	0.116 1	20 010	0.087 0
2007	24 929.00	27 630.402 0	0.108 4	23 820	0.044 5
2008	29 662.00	31 828.281 0	0.073 0	27 990	0.056 4

从表 1 可以看出，单项预测方法存在一定缺陷，灰色预测方法总体上预测的值要比实际值偏大，而指数预测值预测的偏小，效果都不太理想。

利用组合预测进行如下分析。

1) 计算单项预测方法的预测相对误差的熵值为：

$$h_1 = \frac{-\sum_{t=1}^{10} p_{1t} \ln p_{1t}}{\ln 10} = 0.865 193,$$

$$h_2 = \frac{-\sum_{t=1}^{10} p_{2t} \ln p_{2t}}{\ln 10} = 0.764 914,$$

2) 计算各种预测方法的加权系数：

$$l_1 = 0.635 551, l_2 = 0.364 449.$$

3) 组合预测方程则为:

$$\hat{x}_{3t} = 0.635551x_{1t} + 0.364449x_{2t}$$

4) 应用上式进行计算, 可得预测结果(表2)。

表2 组合预测及其误差

年份	真实值/亿元	组合预测/亿元	$e_{3t}/(\%)$
1999	3 651.30	3 731.719	0.022 0
2000	4 133.50	4 699.106	0.136 8
2001	4 459.92	6 535.866	0.465 5
2002	9 050.29	8 075.022	0.107 8
2003	10 523.40	10 221.830	0.028 7
2004	13 704.76	13 023.540	0.049 7
2005	16 987.00	16 160.770	0.048 6
2006	18 408.00	20 349.940	0.105 5
2007	24 929.00	26 241.700	0.052 7
2008	29 662.00	30 429.420	0.025 9

5) 采用组合预测方法对海洋产业的未来发展趋势进行预测, 得到表3。

表3 海洋产业未来的发展趋势 亿元

项目	2009年	2010年	2011年	2012年
海洋生产总值	37 356.43	45 785.52	56 085.15	68 707.22

3 结束语

三种预测方法的结果比较如图1所示(以1999年为第一年), 组合预测不仅提高了预测的精度, 还提高了预测的可靠性。组合预测模型的精度要高于单一模型的预测精度。实证结果进一步表明, 即达到了预期的预测效果, 又尽可能地保留了两种预测方法所涵盖的信息。基于熵值法的组合模型为预测问题提供了一个很好的预测方法, 在实践中具有很好的应用价值。

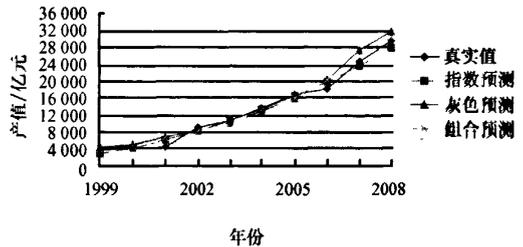


图1 海洋产业生产总值各种预测方法与真实值的比较拟合图

参考文献:

[1] BATES J M, GRANGER C W J. The Combination of Forecasts [J]. Operational Research Quarterly, 1969, 20 (4).

[2] 汪同三, 张涛. 组合预测—理论、方法及应用 [M]. 北京: 社会科学文献出版社, 2008.

[3] 陈华有. 组合预测方法有效性理论及其应用 [M]. 北京: 科学出版社, 2008.

[4] 邓聚龙. 农业系统灰色理论与方法 [M]. 济南: 山东科学技术出版社, 1988: 42-471.

[5] 何晓群. 多元统计分析 [M]. 北京: 中国人民大学出版社, 2004.

[6] 卢纹岱. Spss for windows 统计分析 [M]. 北京: 电子工业出版社, 2006.