

基于漂移度的组合预测方法研究

李美娟, 陈国宏, 林志炳

(福州大学管理学院, 福建 福州 350108)

摘要: 由于不同的预测方法能够提供不同的有用信息, 其预测精度往往也存在差异, 为了分散预测的风险, 采用组合预测方法。本文首先提出相容方法集和互补模型集, 然后在对不同单一预测模型的漂移性和互补性研究的基础上提出了基于漂移度的组合预测模型, 为组合预测模型研究提供一种新的思路。最后通过实例来说明基于漂移度的组合预测模型能够提高样本期预测精度和外推预测精度及实际应用的有效性。

关键词: 组合; 预测; 漂移性; 互补性; 相容性

中图分类号: C931 **文献标识码:** A

1 引言

在预测实践中, 对于同一预测问题可以采用多种预测方法, 不同的预测方法往往能提供不同的有用信息。把不同的预测方法进行适当的组合, 形成组合预测方法。组合预测可以综合利用各单项预测方法提供的信息, 集成不同信息来源的预测结果, 从而有效提高预测精度^[1]。Bates 和 Granger (1969) 首次对组合预测方法进行系统研究^[2], 其研究成果引起预测学者的重视。20 世纪 70 年代以后, 组合预测方法的研究进一步得到了重视。1989 年, 国际预测领域的权威学术刊物 Journal of Forecasting 还出版了组合预测方法专辑, 这充分说明了组合预测方法在预测学中的重要地位^[3]。

近年来, 组合预测研究一直是国内外预测学界探讨的热点问题, 有关组合预测理论与应用研究成果较多, 例如黄岩等 (2001) 建立了基于最小二乘原理的组合预测模型, 提出了求解此组合预测模型的一种新的算法即 Frank-Wolfe 算法^[4]。卢奇等 (2003) 采用标准差法进行组合权重分配, 对我国能

源消费量进行预测^[5]。陈华友等 (2004) 建立新的基于 IOWHA 算子的组合预测模型, 并通过实例分析表明该模型能有效提高组合预测精度^[6]。刘启浩 (2008) 给出了执行风险值组合预测的具体方法, 同时对预测表现进行实证分析^[7]。王硕等 (2008) 提出基于 WNN-RAGA 的非线性组合预测方法^[1]。Kapetanios 等 (2008) 建立英国银行组合预测模型, 减少预测误差^[8]。Daz-Robles 等 (2008) 混合运用 ARIMA 和神经网络模型预测城市中的颗粒, 以智利为例^[9]。Sánchez (2008) 运用自适应组合预测方法对风能源进行预测^[10]。Guidolina 等 (2009) 利用组合预测方法对美国短期利率进行预测^[11]。曾波等 (2009) 提出了一种基于灰色关联度和 GM (1, 1) 的灰色组合预测模型, 并应用该模型对我国 GDP 进行预测^[12]。刘泽双等 (2009) 利用基于遗传算法的组合预测模型对就业需求量进行预测^[13]。陈涛 (2010) 提出基于 AGA-SVM 的非线性组合预测模型^[14]。

现有组合预测研究在模型构造、权重确定和模型比较方面还主要侧重于样本期预测精度, 对外推预测精度的考察尚少。几乎都是直接对单项预测模型进行组合, 未对几种单项预测模型是否可以组合进行判断, 建模较复杂。

本文拟在对几种单项预测模型的漂移度和可组合 (互补) 模型集研究基础上提出一种基于漂移度的组合预测方法, 来弥补现有预测方法的不足。通过样本期预测和事后 (外推) 预测, 验证基于漂移度的

收稿日期: 2010-04-14; 修订日期: 2011-03-29

基金项目: 教育部人文社会科学研究青年基金 (10YJC630120, 10YJC630139); 国家软科学研究项目 (2006GXQ3D108); 福建省软科学研究重点项目 (2009R0054, 2010R0065); 福州大学科技发展基金 (10SKQ02)

作者简介: 李美娟 (1979-), 女 (汉族), 福建古田人, 福州大学管理学院讲师, 博士, 研究方向: 评价与预测。

组合评价方法的有效性。

2 基于漂移度的组合预测模型

2.1 基本假设

基于漂移度的组合预测模型研究基于以下基本假设:

假设1 漂移性假设。对同一组样本运用多种不同方法分别进行建模预测时,客观上存在着预测的漂移性问题,即预测结果与客观实际值存在一定差异,同时不同方法所得预测值之间也存在差异。

假设2 相容性假设。不同的预测方法对不同的预测问题适用程度是有差异的,有些可能是根本不适用的。对某特定观察对象而言,不适用的方法我们称之为不相容方法,适用的方法称之为相容的方法。如果误差指标(例如平均相对误差 MAPE)大于预先设定的阈值,则不适用。本文将相容方法集定义为:适用于预测对象的预测方法的集合 M 。

假设3 互补性假设。对同一个观测点,有些方法可能统计观测值(实际值)比预测值大(预测误差大于0),有些方法可能统计观测值比预测值小(预测误差小于0)。这两类方法进行组合,正负预测误差抵消,得到组合预测值将更接近实际值。

2.2 定义

2.2.1 漂移度

本文将漂移性定义为预测结果与客观实际的不一致性。漂移度用来测度预测漂移性。我们将漂移度定义为

$$p_i = \frac{1}{2} \left(\left| \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_i) \right| + \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_{it}| \right) \quad (1)$$

其中, $\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_i)$ 表示预测误差均值,考查

系统偏差;平均绝对误差 MAE = $\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_{it}|$ 考查平均偏离程度。漂移度不仅要反映平均偏离程度而且要反映系统偏差,因此取二者绝对值的平均值。

2.2.2 互补模型集

设有 m 种单项预测模型, n 个样本,计算第 i 种方法第 t 个样本的预测值 \hat{y}_{it} ,第 i 种方法第 t 个样本的预测误差 e_{it} ,即统计观测值(实际值) y_t 与预测值

\hat{y}_i 的差: $e_{it} = y_t - \hat{y}_i (i = 1, 2, \dots, m; t = 1, 2, \dots, n)$ 。

根据预测误差 e_{it} 判断哪些预测模型是互补的,可以进行组合的。有以下几种情况:(1) 互补模型集。如果对于每个(绝大多数)观测点有的单项模型出现正预测误差,有的出现负预测误差,通过组合正负预测误差可以抵消一部分,它们是互补的,可以进行组合的,保留可以进行组合的互补模型。(2) 非互补模型分为3类 ①优越模型:对于每个观测点,各种模型预测误差 e_{it} 符号完全(绝大多数)相同,如果某种模型的各样本点预测误差绝对值 $|e_{it}|$ 都最小,那么该模型是优越模型,直接用该模型进行预测。②被优越模型:某种模型预测误差 e_{it} 符号与其他方法相同,并且 $|e_{it}|$ 比其它模型大,剔除该模型,再进行判断。③其它不确定情况:不存在优越模型和被优越模型的非互补模型集。例如各种模型预测误差 e_{it} 符号完全相同,但是 $|e_{it}|$ 有的模型大,有的模型小,组合预测模型和单一预测模型,不一定哪种效果好。又如各种模型预测误差 e_{it} 符号有的样本点相同,有的样本点不同,组合预测模型和单一预测模型,不一定哪种效果好。非互补模型根据实际情况进行判断是否进行组合,采用何种预测方法。

2.3 基本思路(步骤)

基于漂移度的组合预测方法的基本思路(步骤)如下:

- (1) 确定预测问题,构建初始预测方法集 M ,运用 M 中的方法建立预测模型。
- (2) 计算预测值和预测误差。
- (3) 确定相容方法集 M_1
如果误差指标(例如平均相对误差 MAPE)大于预先设定的阈值,表示该方法不适用,将其剔除;否则,该方法适用,将其保留。阈值根据实际预测精度需要设置。
- (4) 确定互补模型集 M_2
- (5) 求各种互补模型漂移度。根据公式(1)求各种互补模型漂移度。

(6) 求各种互补模型基于漂移度的权重。在漂移性测度的基础上,在组合时对于漂移度大的预测模型应赋予较小的权重,对于漂移度较小的预测模型应赋予较大的权重。设互补模型集 M_1 中有 m 种预测模型,第 i 种模型的漂移度是 p_i ,则第 i 种模型

的权重如下^[15]:

$$w_i = \frac{\min_{1 \leq j \leq m} p_j - p_i + \min_{1 \leq j \leq m} p_j}{\sum_{i=1}^m (\min_{1 \leq j \leq m} p_j - p_i + \min_{1 \leq j \leq m} p_j)} \quad (i = 1, 2, \dots, m) \quad (2)$$

(7) 建立基于漂移度的组合预测模型, 求组合预测模型预测值 \hat{y}_t 。设第 i 种方法第 t 个样本点的预测值为 \hat{y}_{it} , 基于漂移度的权重为 w_i , 则基于漂移度的组合预测模型为:

$$\hat{y}_t = \sum_{i=1}^m w_i \hat{y}_{it} \quad (i = 1, 2, \dots, m; t = 1, 2, \dots, n) \quad (3)$$

(8) 预测模型样本期预测精度评价与比较。为了反映基于漂移度的组合预测模型样本期预测精度更高, 更有效, 选择几个误差指标评价单项预测模型和组合预测模型。

$$\text{均方根误差 RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

$$\text{平均相对误差 MAPE} = \frac{1}{n} \sum_{t=1}^n |(y_t - \hat{y}_t) / y_t|$$

$$\text{平均绝对误差 MAE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

基于漂移度的组合预测模型的误差指标值和单项预测模型的误差指标值进行比较, 如果基于漂移度的组合预测模型的误差指标值小于各种单项预测模型的误差指标值, 说明基于漂移度的组合预测模型样本期预测精度更高, 预测效果更好, 更有效, 采用基于漂移度的组合预测模型进行外推预测。否则说明基于漂移度的组合预测模型预测效果不好, 重新寻找合适的预测模型。

(9) 利用基于漂移度的组合预测模型进行外推预测(事后预测), 对基于漂移度的组合预测模型的外推预测效果进行评价。如果基于漂移度的组合预测模型的外推预测精度更高(效果更好), 利用该模型进行事前预测, 如果效果不好, 重新寻找合适的预测模型。

3 实例研究

3.1 实例 1: 就业需求量预测

直接引用参考文献[13]中的就业需求量预测作为实例, 数据来自参考文献[13]中的表 1 数据, 如表

1 所示。以 1990-1999 年 10 年数据为样本建立组合预测模型, 2000-2007 年 8 年数据作为外推(事后)预测, 评价组合预测模型外推预测精度。

3.1.1 基于漂移度的组合预测模型

1) 计算模型预测误差 $e_i = y_t - \hat{y}_t (i = 1, 2; t = 1, 2, \dots, 10)$, 见表 2。

2) 经过判断, 2 种模型可以互补, 对于绝大多数样本点(除了 1993, 1995 和 1999 年), 有的单项预测模型出现正预测误差, 有的出现负预测误差, 组合以后正负预测误差可以相互抵消。互补模型集 $M_1 = \{\text{Logistic 模型; GM 模型}\}$ 。

表 1 单项预测方法的预测结果

年份	实际值	Logistic 预测	GM 预测
1990	56740	60452.7	55530.2
1991	65590	69356.1	63891.5
1992	66380	70213.3	64134.3
1993	67240	70681.2	69542.8
1994	68090	71090.6	65316.7
1995	68910	71950.1	70852.2
1996	68950	71998.6	67941.3
1997	69600	72458.1	68587.4
1998	70637	73265.9	69821.9
1999	71394	73986.3	73654.6
2000	72085	74656.1	74131.7
2001	73025	75697.8	72196.6
2002	73740	76064.5	72720.4
2003	74432	75982.6	73102.3
2004	75200	77818.7	77508.9
2005	75825	78816.4	75812.6
2006	76400	79412.5	74612.3
2007	76990	80192.4	77820.3

表 2 单项模型预测误差

年份	Logistic 预测		GM 预测	
	e_{1t}	$ e_{1t} $	e_{2t}	$ e_{2t} $
1990	-3712.7	3712.7	1209.8	1209.8
1991	-3766.1	3766.1	1698.5	1698.5
1992	-3833.3	3833.3	2245.7	2245.7
1993	-3441.2	3441.2	-2302.8	2302.8
1994	-3000.6	3000.6	2773.3	2773.3
1995	-3040.1	3040.1	-1942.2	1942.2
1996	-3048.6	3048.6	1008.7	1008.7
1997	-2858.1	2858.1	1012.6	1012.6
1998	-2628.9	2628.9	815.1	815.1
1999	-2592.3	2592.3	-2260.6	2260.6
平均值	-3192.19	3192.19	425.81	1726.93

3) 求各种互补模型漂移度

$$\text{Logistic 模型 } p^1 = (3192.19 + 3192.19) / 2 =$$

3192.19

GM 模型 $p_2 = (425.81 + 1726.93)/2 = 1076.37$

4) 求各种互补模型基于漂移度的权重

Logistic 模型 $w_1 = 0.2522$, GM 模型 $w_2 = 0.7478$ 。

5) 建立基于漂移度的组合预测模型。基于漂移度的就业需求量组合预测模型为:

$$\hat{y}_t = \sum_{i=1}^m w_i \hat{y}_{it} = 0.2522\hat{y}_{1t} + 0.7478\hat{y}_{2t} (t = 1, 2, \dots, 10)$$

计算基于漂移度的就业需求量组合预测结果如表 3 所示(1990 年—1999 年)。由于 1993, 1995 和 1999 这三年两种单项预测模型的预测误差都为负数, 不能互补, 所以 $|(y_t - \hat{y}_t)/y_t|$ 值在 3% 左右, 比其它年份(多数小于 1%)大, 预测误差 e 也较大。

表 3 基于漂移度的组合预测样本期预测结果

年份	基于漂移度的组合预测			
	\hat{y}_t	e_{it}	$ e_{it} $	$ (y_t - \hat{y}_t)/y_t $
1990	56771.47	-31.47	31.47	0.0006
1991	65269.47	320.53	320.53	0.0049
1992	65667.19	712.81	712.81	0.0107
1993	69829.86	-2589.86	2589.86	0.0385
1994	66772.66	1317.34	1317.34	0.0193
1995	71129.05	-2219.05	2219.05	0.0322
1996	68964.4	-14.40	14.40	0.0002
1997	69563.44	36.56	36.56	0.0005
1998	70690.35	-53.35	53.35	0.0008
1999	73738.24	-2344.24	2344.24	0.0328
平均值		-486.51	963.96	0.0141

6) 预测模型样本期预测精度评价与比较

计算单项及基于漂移度的组合预测模型的 MAE, RMSE, MAPE, 如表 4 所示。

表 4 单项以及组合预测方法样本期预测效果评价

指标	MAE	RMSE	MAPE
Logistic 预测	3192.19	3222.47	0.0479
GM 预测	1726.93	1843.24	0.0257
基于漂移度组合预测	963.96	1395.66	0.0141

基于漂移度的组合预测模型的 MAE, RMSE, MAPE 最小, 比 Logistic 模型和 GM 模型都小。基于漂移度的组合预测方法样本期预测效果优于单项预测方法。

7) 利用基于漂移度的组合预测模型进行外推预测。计算基于漂移度的就业需求量组合预测结果如表 5 所示(2000 年—2007 年)。由于 2000、2004 和

2007 年这三年两种单项预测模型的预测误差都为负数, 不能互补, 所以 $|(y_t - \hat{y}_t)/y_t|$ 值比其它年份(小于 1%)大, 预测误差 e 也较大。

表 5 基于漂移度的组合预测外推预测结果

年份	基于漂移度的组合预测			
	\hat{y}_t	e_{it}	$ e_{it} $	$ (y_t - \hat{y}_t)/y_t $
2000	74263.93	-2178.93	2178.93	0.0302
2001	73079.47	-54.47	54.47	0.0007
2002	73563.66	176.34	176.34	0.0024
2003	73828.60	603.40	603.40	0.0081
2004	77587.02	-2387.02	2387.02	0.0317
2005	76570.05	-745.05	745.05	0.0098
2006	75822.73	577.27	577.27	0.0076
2007	78418.45	-1428.45	1428.45	0.0186
平均值		-679.61	1018.87	0.0136

8) 组合预测模型外推预测效果评价

基于漂移度的组合预测模型外推预测精度如表 6 所示。表 6 表明基于漂移度的组合预测方法的外推预测精度较高, 外推预测效果较好, MAE, RMSE, MAPE 都最小。

表 6 单项及基于漂移度的组合预测方法外推预测效果评价

指标	MAE	RMSE	MAPE
Logistic	2618.00	2662.20	0.0350
GM	1270.46	1453.42	0.0171
基于漂移度组合预测	1018.87	1312.09	0.0136

综上所述, 与 Logistic 模型和 GM 模型比较, 基于漂移度的组合预测模型具有较高的样本期预测精度和外推预测精度。可以利用基于漂移度的组合预测模型进行事前预测。

3.1.2 基于漂移度的组合预测模型与其它组合预测模型比较

基于漂移度的组合预测方法与其它几种常用的组合预测方法^[3](算术平均法、预测误差平方和倒数方法、均方误差倒数方法、简单加权平均方法和二项式系数方法)的样本期预测效果评价如表 7 所示, 外推预测效果评价如表 8 所示。

表 7 组合预测方法样本期预测效果评价

指标	MAE	RMSE	MAPE
基于漂移度组合预测	963.96	1395.66	0.01406
二项式系数方法	965.30	1396.81	0.01408
预测误差平方和倒数方法	969.15	1398.79	0.01413
算术平均法	1383.19	1622.78	0.02052
均方误差倒数方法	1031.88	1413.46	0.01515
简单加权平均方法	993.80	1393.56	0.01455

表 8 组合预测方法外推预测效果评价

指标	MAE	RMSE	MAPE
基于漂移度组合预测	1018.87	1312.09	0.01364
二项式系数方法	1019.22	1311.57	0.01365
预测误差平方和倒数方法	1019.79	1310.80	0.01366
算术平均法	1322.01	1555.19	0.01770
均方误差倒数方法	1049.83	1379.52	0.01407
简单加权平均方法	1029.29	1353.43	0.01379

表 7 和表 8 表明, 和其它几种常用的组合预测方法相比, 基于漂移度的组合预测模型具有较高的

样本期预测精度和外推预测精度。

综上所述, 本实例对有漂移规律的相容互补模型集进行组合, 基于漂移度的组合预测模型具有较高的样本期预测精度和外推预测精度, 可以利用基于漂移度的组合预测模型进行事前预测。

3.2 实例 2: 能源消费量预测

直接引用参考文献[5]中的能源消费量预测作为实例, 数据来自参考文献[5]中的表 1 数据, 如表 9 所示。

表 9 能源消费量单项预测值及预测误差

年份	能源实际消费量 y_t	灰色 \hat{y}_{1t}	神经网络 \hat{y}_{2t}	多元回归 \hat{y}_{3t}	e_{1t}	e_{2t}	e_{3t}
1985	76682	77720	77380	76651	-1038	-698	31
1986	80850	90450	81225	80377	-9600	-375	473
1987	86632	93880	87654	84890	-7248	-1022	1742
1988	92997	97440	94354	90942	-4443	-1357	2055
1989	96934	101130	97533	94556	-4196	-599	2378
1990	98703	104960	98415	97353	-6257	288	1350
1991	103783	108940	103605	101626	-5157	178	2157
1992	109170	113070	110000	107249	-3900	-830	1921
1993	115993	117350	118038	114214	-1357	-2045	1779
1994	122737	121800	125978	122118	937	-3241	619
1995	131176	126420	132110	128088	4756	-934	3088
1996	138948	131210	140721	132195	7738	-1773	6753
1997	137798	136180	140214	134875	1618	-2416	2923
1998	132214	141350	134283	136498	-9136	-2069	-4284
1999	133831	146700	124117	137937	-12869	9713.97	-4106.03

以 1985-1994 年 10 年数据为样本建立组合预测模型, 单项以及组合预测方法样本期预测效果评价见表 10, 基于漂移度的组合预测方法的样本期预测效果很好, 基于漂移度的组合预测方法的 MAE, RMSE, MAPE 最小, MAPE = 0.47%, 小于 1%。基于漂移度的组合预测方法的样本期预测效果最好的原因在于 10 个样本点, 对于每个样本点, 有的单项模型出现正预测误差, 有的出现负预测误差, 组合以后正负预测误差可以相互抵消, 三种单项预测模型互补, 而且具有一定的漂移规律(见表 9, 1985-1994 年预测误差)。除了 1986 年 $|(y_t - \hat{y}_t)/y_t| = 1.28%$, 其它年份都小于 1%, 在本实例中基于漂移度的组合预测方法的样本期预测效果很好。

1995-1999 年 5 年数据作为外推(事后)预测, 评价组合预测模型外推预测精度。单项以及组合预测方法外推效果评价见表 11。基于漂移度的组合预测方法外推预测精度较高。组合预测模型外推预测精度比单项预测模型高, 1995-1999 年三种单项预测模型预测误差基本能够互补(见表 9, 1995-1999 预测误差)。

表 10 单项以及组合预测方法样本期预测效果评价

指标	MAE	RMSE	MAPE
灰色	4413.30	5165.19	0.0475
神经网络	1063.30	1391.48	0.0103
多元回归	1450.50	1638.61	0.0145
基于漂移度组合预测	448.39	576.51	0.0047
二项式系数方法	573.21	730.74	0.0057
预测误差平方和倒数方法	491.08	626.52	0.0048
算术平均法	1248.50	1484.51	0.0136
均方误差倒数方法	499.81	630.04	0.0053
简单加权平均方法	705.87	848.06	0.0075

表 11 单项以及组合预测方法外推效果评价

指标	MAE	RMSE	MAPE
灰色	7223.40	8175.47	0.0538
神经网络	3381.20	4658.23	0.0251
多元回归	4230.80	4447.35	0.0313
基于漂移度组合预测	1943.39	2301.83	0.0144
二项式系数方法	1969.29	2410.09	0.0147
预测误差平方和倒数方法	1909.74	2283.10	0.0142
算术平均法	2967.85	3355.21	0.0221
均方误差倒数方法	1930.18	2312.04	0.0144
简单加权平均方法	1874.76	2305.15	0.0140

3.3 实例 3:非互补模型

引用参考文献[14]表 3.1 三种单一模型的预测结果如表 12 所示,表 12 中总共有 12 组样本点,9 组样本点预测误差 e_{it} 符号相同,7 组样本点神经网络预测模型预测误差较小,不能互补。

单项以及组合预测方法样本期预测效果评价见表 13,对于非互补模型,几种组合预测方法的预测效果不如神经网络这种单项预测模型,比灰色模型和多元回归模型好。

表 12 三种单一模型的预测结果和预测误差

观测值	指数自回归	ARMA	灰色模型			
y_t	模型 \hat{y}_{1t}	\hat{y}_2	\hat{y}_{3t}	e_{1t}	e_{2t}	e_{3t}
4 56	4 29	4 25	4 16	0 27	0 31	0 4
4 4	4 24	4 23	4 33	0 16	0 17	0 07
4 22	4 08	4 13	4 41	0 14	0 09	- 0 19
4 09	3 89	4	4 14	0 2	0 09	- 0 05
3 89	3 82	3 92	4 02	0 07	- 0 03	- 0 13
4 18	3 9	3 96	3 85	0 28	0 22	0 33
4 29	3 97	4 08	4 03	0 32	0 21	0 26
4 44	4 17	4 22	4 22	0 27	0 22	0 22
4 26	4 06	4 24	4 13	0 2	0 02	0 13
4 02	3 91	3 93	3 86	0 11	0 09	0 16
3 85	3 86	3 91	3 94	- 0 01	- 0 06	- 0 09
4 31	4 11	4 08	4 16	0 2	0 23	0 15

表 13 单项以及组合预测方法预测效果评价

指标	MAE	RMSE	MAPE
灰色	0 1858	0 2059	0 0434
神经网络	0 1450	0 1703	0 0337
多元回归	0 1817	0 2078	0 0427
基于漂移度组合预测	0 1518	0 1820	0 0354
二项式系数方法	0 1539	0 1783	0 0358
预测误差平方和倒数方法	0 1525	0 1807	0 0355
算术平均法	0 1536	0 1831	0 0358
均方误差倒数方法	0 1531	0 1819	0 0356
简单加权平均方法	0 1542	0 1799	0 0359

受篇幅所限,其它实例不列出。

4 结语

本文在对各个单项预测模型的相容性、漂移性和互补性(对单项预测模型是否可以组合进行判断)研究的基础上,提出基于漂移度的组合预测模型,对互补模型集中的预测模型进行组合,对预测模型的样本期预测精度和外推预测精度进行评价与比较。通过实例表明对于有一定漂移规律的相容互补模型集,基于漂移度的组合预测模型样本期预测精度和外推预测精度都较高,在提高预测精度,减少预测风

险方面具有优势。基于漂移度的组合预测模型适用于相容互补模型集,对于非互补模型,组合预测效果不一定比单项预测模型好。本研究为组合预测研究提供一种新的思路。可以进一步利用基于漂移度的组合预测方法进行实证研究。

参考文献:

- [1] 王硕,杨善林,马溪骏.基于 WNN-RAGA 的非线性组合预测方法[J].系统工程理论与实践,2008,(12):160-165.
- [2] Bates, J. M., Granger, C. W. J.. The Combination of Forecasts[J]. Operations Research Quarterly, 1969, 20(4):451-468.
- [3] 陈华友.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] 黄岩,张国春,王其藩,朱道立.一种新的计算组合预测权重的方法[J].管理工程学报,2001,15(2):44-46.
- [5] 卢奇,顾培亮,邱世明.组合预测模型在我国能源消费系统中的建构及应用[J].系统工程理论与实践,2003,(3):24-30.
- [6] 陈华友,刘春林,盛昭瀚. IOWHA 算子及其在组合预测中的应用[J].中国管理科学,2004,12(5):35-40.
- [7] 刘启浩. VaR 的组合预测方法[J].预测,2008,27(3):75-80.
- [8] Kapetanios, G., Labhard, V., Price, S.. Forecast combination and the Bank of England's suite of statistical forecasting models[J]. Economic Modelling, 2008, (25):772-792.
- [9] Daz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., Moncada-Herrera, J. A.. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile[J]. Atmospheric Environment, 2008, 42(35):8331-8340.
- [10] Sanchez, I.. Adaptive combination of forecasts with application to wind energy[J]. International Journal of Forecasting, 2008,(24):679-693.
- [11] Guidolina, M., Timmermann, A.. Forecasts of US short-term interest rates: A flexible forecast combination approach[J]. Journal of Econometrics, 2009,(150):297-311.
- [12] 曾波,刘思峰,方志耕,谢乃明.灰色组合预测模型及其应用[J].中国管理科学,2009,17,(5):150-155.
- [13] 刘泽双,闫付强.基于遗传算法的就业需求量组合预测模型[J].系统工程,2009,27(8):62-68.
- [14] 陈涛.基于 AGA-SVM 的非线性组合预测模型[J].数学的实践与认识,2010,40(4):57-62.
- [15] 陈国宏,李美娟,陈衍泰.组合评价及其计算机集成系统研究[M].北京:清华大学出版社,2007.

A Combination Forecasting Model based on Drift

LI Mei-juan, CHEN Guo-hong, LIN Zhi-bing

(School of Management, Fuzhou University, Fuzhou 350108, China)

Abstract: As different methods diverse in providing information and forecast precision, the combined forecasting model is employed. Combination forecasting model based on drift is proposed on basis of the study on drift and complementarity of different single forecasting models. A new thought of combination forecasting model is given. Finally some practical examples are given to show that combination forecasting model based on drift can improve forecasting precision and is effective in practice.

Key words: combination; forecasting; drift; complementarity; compatibility