

单一方法评价结论漂移性的测度研究

陈国宏, 李美娟, 陈衍泰

(福州大学管理学院, 福州 350002)

[摘要] 主要研究单一方法评价结论的漂移性, 以便为组合评价提供实验依据。把现实评价问题分为3大类分别进行仿真实验, 利用MATLAB生成待评价对象, 利用不同方法进行评价, 把评价结果进行标准化处理, 并判别评价方法的相容性, 以相容方法的平均值为参照系, 求各评价方法的漂移度。

[关键词] 综合评价方法; 计算机仿真实验; 漂移性

[中图分类号] C931 **[文献标识码]** A **[文章编号]** 1009-1742(2004)03-0058-06

1 引言

综合评价是指对以多属性体系结构描述的对象系统作出全局性、整体性的评价^[1]。目前国内外建立的综合评价方法有数百种之多, 但大多数尚处于理论研究阶段, 不十分成熟。最主要的和最常用的方法有专家评价法、经济分析法、层次分析法、主成分分析法、因子分析法、人工神经网络法、数据包络分析法、灰色关联度法、模糊综合评价法、熵值法等等。

由于应用不同的方法对同一对象进行评价所得到的结果往往存在着差异, 同时每一种方法都各有其优缺点, 无法说明哪一种方法好或不好, 仅采用一种方法进行评价无疑具有一定的片面性。为解决上述问题, 目前有不少学者提出了组合评价的思路^[2-4], 即选用多种方法进行评价, 而后将几种评价结果进行组合。但目前有关组合评价模型的依据都不够充分。因为到目前为止尚未有人对不同方法评价结论的差异(漂移)规律作过认真的研究, 因此, 很难提出让人信服的组合评价模型。

由于评价机理的不同, 不同方法对被评价对象是有偏好的, 即有些方法可能更适合于对某些问题

进行评价, 而对另一些问题可能不太适合, 或者适合的程度有所不同。要建立多方法的组合评价模型, 必须对这些规律有所了解。从实践意义上来讲, 要研究不同评价方法的漂移规律, 可以通过对大量的实例用不同的方法进行评价后加以归纳和总结。显然, 大量实例(样本)的采集, 难度相当大。因此, 通过仿真实验的方法研究单一方法评价结论的漂移性, 旨在为组合评价提供实验依据。

2 基本思路

漂移性定义为评价结论与客观实际的不一致性。漂移度是单一方法评价结论漂移性的测度。

2.1 基本假设

对单一方法评价结论漂移性的测度研究基于以下基本假设:

假设1 漂移性假设 对同一对象运用多种不同方法分别进行评价时客观上存在着结论的漂移性问题, 即评价结论与客观实际存在一定差异, 同时不同方法所得结论之间也存在差异。

假设2 相容性假设 不同的评价方法对不同的待评价问题适用程度是有差异的, 有些可能是根本不适用的。对某特定待评价问题而言, 不适用的

[收稿日期] 2003-06-25; **修回日期** 2003-07-31

[基金项目] 国家自然科学基金资助项目(70171026)

[作者简介] 陈国宏(1953-), 男, 福建福州市人, 博士, 福州大学教授, 博士生导师

方法称为不相容方法，适用的方法称为相容的方法。

假设3 待评价对象的等级差异假设 在实践中同一组内若干个被评价对象客观上存在等级差异。

假设4 待评价对象数据分布假设 某特定等级内待评价对象某属性（指标）值的样本数据服从正态分布。

2.2 基本思路

试图通过计算机仿真实验揭示不同方法评价结论的漂移规律，并对漂移性进行测度。基本思路如图1所示，具体说明如下：

Step 1 确定实验方法 选择几种有代表性的评价方法进行实验。

Step 2 利用系统仿真方法生成足够的待评价样本对象 根据实验设计要求利用 MATLAB6.1 生成接近现实的分等级的待评价样本对象。

Step 3 利用几种不同的评价方法分别对相同待评价样本对象进行评价，并将评价结果进行标准化处理。

Step 4 相容性检验包括直观判断和模糊聚类分析。

直观判断 比较各种不同方法的评价结论与预期结果（实验数据生成时人为划分的等级）是否相似，对评价结论与预期结果明显不相似的方法可认为相对于被评价对象此方法为不相容方法，应予以剔除。所有通过直观判断未被剔除的方法构成方法集 M 。

模糊聚类分析 对方法集 M 进一步用模糊聚类方法检验其相容性。步骤如下：

a. 对方法集 M 中所有方法的评价结论向量作模糊聚类分析，在给定阈值 ξ 的情况下可以将其分成2类，第一类仅由第 k 种方法的评价结论 u_k 组成，第二类由除 u_k 以外的所有向量组成。

b. 求第二类评价结论的平均向量 \bar{u} ，以代表第二类评价结论总体。

c. 求 u_k 与 \bar{u} 的相关系数 r_k 。在一定显著性水平下根据 r_k 检验 u_k 与 \bar{u} 的相关性，进而判断第 k 种方法与第二类方法全体是否相容。如果检验结果认为 u_k 与 \bar{u} 密切相关，则判定集合 M 中的方法都具有相容性， M 即为相容方法集 M_1 ；否则判定第 k 种方法是 M 中的一种不相容方法并将其从 M 中剔除，然后对 M 中所剩余方法按以上步骤再进

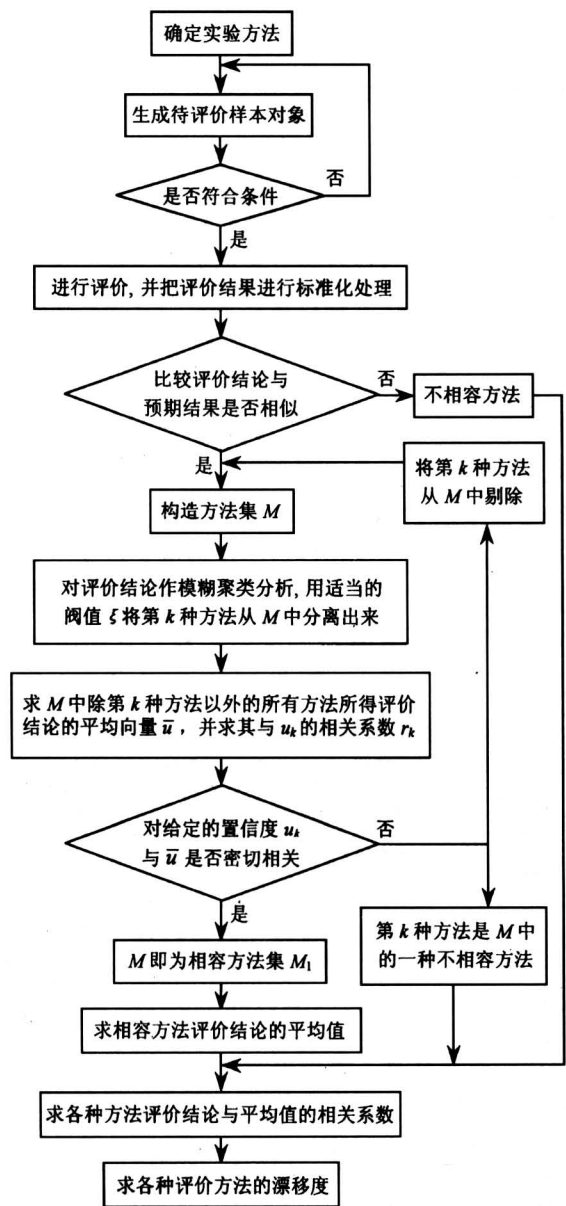


图1 单一方法评价结论漂移性测度的仿真实验基本思路

Fig.1 Basic thought of simulation to measure drifting degree of single method evaluation conclusion

行检验，直至找到相容方法集 M_1 。

Step 5 求标准化处理以后相容评价方法评价结论的平均值，并作为各种方法评价结论漂移性测度的参照系。

Step 6 求各评价方法标准化处理后的评价与所有相容方法评价结论的平均值的相关系数 r 。

Step 7 求漂移度 将漂移度定义为 $p = 1 - r$ 。

3 实验设计

3.1 确定实验方法

参考文献 [5] 中曾将已有的综合评价方法分成 12 类, 研究表明, 由于评价原理和适用对象的差异, 许多方法是不具有可比较性的, 也难以用一个统一的标准来寻找各种方法的偏移规律, 因此将方法分为 3 个层次分别讨论。

1) 具有相同属性的综合评价方法 常见的单一评价方法中, 有一类方法能够实现对结构不良的系统对象进行有效的评价; 同时, 这类方法在评价时可以避开人为权重, 减小因主观性造成的偏差, 评价结论体现了评价对象的实际水平; 另外, 对于每个对象的评价结果均以一个评价值来反映, 且此数值的差异在一定意义上反映对象的差异程度。这类方法, 定义为具有相同属性的综合评价方法集合, 简称综合评价方法集合 1。

2) 特殊属性的综合评价方法 在综合评价方法中, 有一类方法由于其工作机理的特殊性, 只能适用于某些特定属性的评价对象, 而不具有一般的适用性; 同时与综合评价方法集合 1 相比, 其评价的结论没有一般的可比性。定义此类方法为具有特殊属性的综合评价方法集合, 简称综合评价方法集合 2。

3) 基于方法论层次的综合评价方法 在综合评价中, 存在另外一类评价方法, 与评价方法集合 1、评价方法集合 2 相比, 这类方法实质上是一种思想体系。它可以指导人们如何去选择合适的评价方法, 如何去有效地、合理地执行具体的评价工作; 可以很好地指导评价工作者系统地、逻辑地去思考整个评价问题, 确保不遗漏任何有可能引起重要后果的部分; 在评价中合理地考虑并处理好人际关系可以更有利于评价工作的顺利完成, 并保证评价结果的有效利用。定义这类方法为评价方法集合 3, 或方法论层次的评价方法。

由于第一层次是一些最常用的评价方法, 故实验主要对其展开讨论。它主要包括统计分析方法、人工神经网络方法、熵值法、改进的数据包络分析法、灰色关联度法、模糊综合评价法。由于人工神经网络方法必须事先有样本进行训练, 因此实验暂不考虑此方法。

熵值法、灰色关联度、模糊综合评价法、改进的 DEA 和主成分分析法都是评价中常用的方

法^[6~10]。

3.2 评价对象分类

根据现实评价问题, 把待评价对象分为以下 3 类分别进行实验:

1) 只有正向指标的评价问题。

2) 只有逆向指标的评价问题。

3) 既有正向指标又有逆向指标的评价问题, 将该类评价对象再分为 3 类进行实验: a. 投入少量增加引起产出大幅度增加的规模收益递增的投入—产出型问题; b. 投入大幅度增加而产出只有少量增加的规模收益递减的投入—产出型问题; c. 规模收益不变的投入—产出型问题。

3.3 实验数据生成

假设每个被评价对象由若干属性(指标)值来体现(其中有正向指标和逆向指标若干), 运用 Matlab 6.1 生成若干组评价样本(每组包含若干个待评价对象)。根据前面的基本假设, 在生成每组待评价对象时, 人为划分若干等级(以 1—100 的随机数表示, 视为标准化以后的样本数据), 属于同一等级的被评价对象分别在一定范围(期望值、方差)内生成服从正态分布的属性值。

3.4 评价及漂移度计算

利用前面确定的评价方法分别对以上 3 大类评价对象的实验数据进行评价, 利用 MATLAB 6.1 编程计算熵值法、灰色关联度法、模糊综合评价法和改进的 DEA 方法评价结果。利用 SPSS 计算主成分分析法评价结果。对于以上 5 种方法的评价结果都利用极差变换法进行标准化处理。然后按照以上基本思路的 Step 4 至 Step 7 进行方法的相容性检验和漂移性计算。

4 实验过程、结果及其分析

对以上 3 大类评价对象分别进行实验: 对每一类对象利用 MATLAB 6.1 生成 10 组服从正态分布的随机样本, 每组样本都包含 3 个等级, 每个等级各 20 个样本(对象、方案), 这样每组共有 60 个待评价对象。

4.1 只有正向指标的评价问题

服从正态分布的随机数按如下规则生成:

等级 7 个正向指标

一级 random (norm, 75, 9, 20, 7)

二级 random (norm, 50, 16, 20, 7)

三级 random (norm, 25, 9, 20, 7)

Random 函数说明: 第一个参数表示服从正态分布, 第

二个参数表示均值，第三个参数表示方差，第四个参数表示随机数矩阵的行数（待评价对象数），第五个参数表示随机数矩阵的列数（评价指标数）

利用熵值法、灰色关联度法、模糊综合评价法和主成分分析法分别对以上生成的待评价对象进行评价，并进行相容性检验。结果表明，此 4 种方法为相容方法，求经过标准化处理后相容方法评价结果的平均值。再分别求各种方法的评价结论与平均值的相关系数 r ，得到各单一方法评价结论的漂移度 $p = 1 - r$ ，如表 1 所示。

表 1 只有正向指标评价问题单一方法评价结论的漂移度

Table 1 Drifting degrees of evaluation objects with positive indicators only

实验组数	熵值法	灰色关联度法	模糊综合评价法	主成分分析法
1	0.001 217 000	0.005 619	0.001 817	0.000 812
2	0.000 985 424	0.005 771	0.002 636	0.000 931
3	0.001 383 530	0.006 084	0.001 689	0.001 234
4	0.001 105 721	0.006 403	0.001 891	0.000 950
5	0.001 148 000	0.005 872	0.001 905	0.001 024
6	0.001 383 530	0.006 084	0.001 689	0.001 234
7	0.001 090 378	0.006 401	0.001 887	0.000 944
8	0.001 147 921	0.005 872	0.001 905	0.001 024
9	0.001 085 043	0.006 326	0.002 783	0.000 953
10	0.001 217 361	0.005 619	0.001 817	0.000 812
平均值	0.001 176 391	0.006 005	0.002 002	0.000 992

表 1 的实验结果表明，对于此类问题，这 4 种评价方法是适合的，且漂移度很小，其漂移度大小关系为：主成分分析法 < 熵值法 < 模糊综合评价法 < 灰色关联度法。

4.2 只有逆向指标的评价问题

待评价对象按如下规则生成：

- 等级 7 个逆向指标
- 三级 random (norm, 75, 9, 20, 7)
- 二级 random (norm, 50, 16, 20, 7)
- 一级 random (norm, 25, 9, 20, 7)

利用熵值法、灰色关联度法、模糊综合评价法和主成分分析法分别对以上生成的待评价对象进行评价，并进行相容性检验。结果表明，这 4 种方法为相容方法，求经过标准化处理后相容方法评价结果的平均值。再分别求各种方法的评价结论与平均值的相关系数 r ，得到各单一方法评价结论的漂

度，如表 2 所示。

表 2 只有逆向指标评价问题单一方法评价结论的漂移度

Table 2 Drifting degrees of evaluation objects with reverse indicators only

实验组数	熵值法	灰色关联度法	模糊综合评价法	主成分分析法
1	0.059 122	0.004 483	0.022 355	0.015 149
2	0.010 674	0.003 136	0.018 768	0.007 199
3	0.091 158	0.007 901	0.028 625	0.023 610
4	0.001 148	0.005 872	0.001 905	0.001 024
5	0.010 452	0.002 864	0.019 161	0.006 993
6	0.059 122	0.004 483	0.022 355	0.015 149
7	0.021 677	0.004 568	0.024 836	0.014 975
8	0.057 980	0.004 837	0.024 796	0.014 637
9	0.015 465	0.002 860	0.023 129	0.010 016
10	0.230 610	0.010 329	0.031 281	0.028 658
平均值	0.055 741	0.005 133	0.021 721	0.013 741

表 2 的实验结果表明，对于此类问题，这 4 种评价方法是适合的，且漂移度较小，其漂移度大小关系为：灰色关联度法 < 主成分分析法 < 模糊综合评价法 < 熵值法。

4.3 既有正向指标也有逆向指标的评价问题

4.3.1 投入少量增加引起产出大幅度增加的规模收益递增的投入—产出型问题 待评价对象按如下规则生成：

- 等级 4 个输入（逆向）指标 8 个输出（正向）指标
- 一级 random (norm, 60, 9, 20, 4) random (norm, 80, 9, 20, 8)
- 二级 random (norm, 50, 9, 20, 4) random (norm, 50, 9, 20, 8)
- 三级 random (norm, 40, 9, 20, 4) random (norm, 30, 9, 20, 8)

利用熵值法、灰色关联度法、模糊综合评价法、改进的 DEA 方法和主成分分析法分别对以上生成的待评价对象进行评价，并进行相容性检验，结果表明这 5 种方法为相容方法，求经过标准化处理后相容方法评价结果的平均值。再分别求各种方法的评价结论与平均值的相关系数 r ，得到各单一方法评价结论的漂移度，如表 3 所示。

表 3 的实验结果表明，对于此类问题，这 5 种评价方法是适合的，且漂移度较小，其漂移度大小关系为：熵值法 < 模糊综合评价法 < 主成分分析法 < 灰色关联度法 < 改进的 DEA 法。

4.3.2 投入大幅度增加而产出只有少量增加的规模收益递减的投入—产出型问题 待评价对象按如

下规则生成:

等级	5个输入指标	3个输出指标
一级	random (norm, 30, 9, 20, 5)	random (norm, 40, 9, 20, 3)
二级	random (norm, 50, 9, 20, 5)	random (norm, 50, 9, 20, 3)
三级	random (norm, 80, 9, 20, 5)	random (norm, 60, 9, 20, 3)

**表3 规模收益递增型评价问题
单一方法评价结论的漂移度**

Table 3 Drifting degrees of evaluation objects
with increasing returns to scale

实验组数	熵值法	灰色关联度法	模糊综合评价法	改进的DEA法	主成分分析法
1	0.020 713	0.027 200	0.017 580	0.127 470	0.028 489
2	0.005 035	0.039 831	0.004 002	0.087 194	0.027 249
3	0.006 267	0.039 198	0.008 945	0.142 852	0.034 578
4	0.003 379	0.025 107	0.005 508	0.105 458	0.031 006
5	0.007 600	0.033 199	0.005 002	0.099 742	0.038 056
6	0.004 078	0.036 822	0.005 207	0.108 031	0.030 190
7	0.010 453	0.022 505	0.008 166	0.079 229	0.019 989
8	0.005 882	0.031 551	0.004 175	0.146 524	0.045 690
9	0.005 266	0.036 403	0.008 576	0.125 614	0.030 039
10	0.005 278	0.032 244	0.006 924	0.124 349	0.036 230
平均值	0.007 395	0.032 406	0.007 409	0.114 646	0.032 152

利用熵值法、灰色关联度法、模糊综合评价法、改进的DEA方法和主成分分析法分别对以上生成的待评价对象进行评价,并进行相容性检验,结果表明这5种方法为相容方法,求经过标准化处理后相容方法评价结果的平均值。再分别求各种方法的评价结论与平均值的相关系数 r ,得到各单一方法评价结论的漂移度,如表4所示。

表4的实验结果表明,对于此类问题,这5种评价方法是适合的,且漂移度较小,其漂移度大小关系如下:灰色关联度法<熵值法<模糊综合评价法<主成分分析法<改进的DEA法。

4.3.3 规模收益不变的投入—产出型问题 待评价对象按如下规则生成:

4个输入指标	4个输出指标
random (norm, 25, 9, 20, 4)	random (norm, 25, 9, 20, 4)
random (norm, 50, 16, 20, 4)	random (norm, 50, 16, 20, 4)
random (norm, 75, 9, 20, 4)	random (norm, 75, 9, 20, 4)

利用熵值法、灰色关联度法、模糊综合评价法、改进的DEA方法和主成分分析法分别对以上生成的待评价对象进行评价,并进行相容性检验,结果表明主成分分析法为不相容方法;熵值法、灰色关联度法、模糊综合评价法、改进的DEA方法

为相容方法,求这4种相容方法标准化处理后的评价结果平均值。再分别求5种方法评价结果与平均值的相关系数,得到它们漂移度,如表5所示。

**表4 规模收益递减型评价问题
单一方法评价结论的漂移度**

Table 4 Drifting degrees of evaluation objects
with decreasing returns to scale

实验组数	熵值法	灰色关联度法	模糊综合评价法	改进的DEA法	主成分分析法
1	0.018 553	0.020 251	0.041 745	0.192 904	0.057 536
2	0.030 686	0.021 228	0.048 150	0.220 476	0.057 183
3	0.023 444	0.023 943	0.039 115	0.286 142	0.038 879
4	0.030 416	0.027 780	0.046 749	0.220 844	0.076 337
5	0.028 581	0.020 672	0.040 306	0.170406	0.062 919
6	0.025 473	0.020 101	0.044 249	0.185 079	0.053 195
7	0.016 156	0.028 929	0.045 208	0.216 193	0.042 450
8	0.022 675	0.024 181	0.033 174	0.190 681	0.067 395
9	0.016 262	0.020 448	0.048 076	0.142 697	0.043 037
10	0.050 042	0.028 799	0.046 132	0.178 081	0.050 583
平均值	0.026 229	0.023 633	0.043 290	0.200 350	0.054 951

**表5 规模收益不变型评价问题
单一方法评价结论的漂移度**

Table 5 Drifting degrees of evaluation objects
with constant returns to scale

实验组数	熵值法	灰色关联度法	模糊综合评价法	改进的DEA法	主成分分析法
1	0.295 706	0.139 068	0.162 640	0.368 636	1.045 135
2	0.346 771	0.161 545	0.249 379	0.323 425	1.316 514
3	0.494 285	0.136 897	0.160 525	0.316 399	0.991 857
4	0.211 996	0.174 708	0.255 238	0.301 345	0.766 821
5	0.105 338	0.094 765	0.188 223	0.303 039	1.157 558
6	0.198 759	0.139 405	0.167 139	0.295 580	1.117 831
7	0.328 100	0.139 641	0.193 115	0.357 292	1.205 989
8	0.128 509	0.104 290	0.216 847	0.252 908	1.316 560
9	0.117 520	0.136 192	0.170 332	0.327 708	1.065 489
10	0.196 129	0.106 219	0.174 739	0.313 298	1.138 674
平均值	0.242 311	0.133 273	0.193 818	0.315 963	1.112 243

实验中,对于规模收益不变的投入—产出型问题,主成分分析法为不相容方法,漂移度很大,不适合评价该类问题;熵值法、灰色关联度法、模糊综合评价法和改进的DEA法为相容方法。这5种方法的漂移度大小关系为:灰色关联度法<模糊综

合评价法<熵值法<改进的DEA法<主成分分析法。

5 结论

通过仿真实验研究单一评价方法的漂移性，为综合评价方法的研究提供了很好的思路。研究结果可为建立组合评价模型提供实验依据。一般认为，用于组合评价的方法集应定义为所有相容方法的集合，在组合时对于漂移度大的评价方法应赋予较小的权重，对于漂移度较小的评价方法应赋予较大的权重。

参考文献

- [1] 王宗军. 综合评价的方法、问题及其研究趋势 [J]. 管理科学学报, 1998, 1(1): 73~79
- [2] 郭显光. 一种新的综合评价方法——组合评价法 [J]. 统计研究, 1995, (5): 56~59
- [3] 彭勇行. 国际投资环境的组合评价研究 [J]. 系统工程理论与实践, 1997, (11): 13~17
- [4] 毛定祥. 一种最小二乘意义下主客观评价一致的组
合评价方法 [J]. 中国管理科学, 2000, 10(5): 95
~97
- [5] Chen Guohong, Chen Yantai. The research progress
& development trend of comprehensive evaluation
methods [A]. Proceedings of 2002' International
Conference on Management Science & Engineering
[C]. Harbin: Harbin Institute of Technology (HIT)
Press, 2002. 462~470
- [6] 王伟. 基于熵的财税政策相对优异性评价 [J].
数量经济技术经济研究, 2000, (3): 45~48
- [7] 彭勇行. 管理决策分析 [M]. 北京: 科学出版社,
2000. 147~285
- [8] 楼顺天, 胡昌华, 张 伟. 基于 MATLAB 的系统分
析与设计——模糊系统 [M]. 西安: 西安电子科技
大学出版社, 2001. 1~22
- [9] 李 果, 沈晓勇, 王应明. 对决策单元进行排序的
一种方法 [J]. 预测, 2002, (4): 51~53.
- [10] 何晓群. 现代统计分析方法与应用 [M]. 北京: 中
国人民出版社, 1998. 281~315

The Research on Measurement of Drift in the Evaluation Conclusion of Single Method

Chen Guohong, Li Meijuan, Chen Yantai

(Management School, Fuzhou University, Fuzhou 350002, China)

[Abstract] In this paper drifting nature in the evaluation conclusion of single method is investigated in order to offer the experiment basis for combination evaluation. First, real evaluation problems are divided into 3 classes and emulation experiments are carried on them separately. Then, objects to be evaluated are generated by Matlab 6.1, and are evaluated by using different methods. Third, the compatibility of the comprehensive evaluation methods is discriminated. Finally, the average of the compatible methods is adopted as reference frame in order to measure drifting degree of each method quantitatively.

[Key words] comprehensive evaluation methods; computer simulation experiment; measurement of drift