

沿海省市海洋科技水平评价^{*}

殷克东^{1,2} 张 燕²

(1. 华中科技大学 武汉 430074; 2. 中国海洋大学 青岛 266071)

摘 要 本研究通过建立海洋科技水平评价指标体系, 利用主成分分析法和灰色关联分析法分别构建了海洋科技水平的评价模型, 对我国沿海地区海洋科技水平进行测算和评价, 研究结果为沿海地区提高海洋科技水平, 培育新的高新技术产业, 制定未来的主导产业发展规划, 提供了科学的理论依据。

关键词 海洋科技水平; 主成分分析; 灰色关联分析

近年来, 海洋经济在国民经济中的地位日益突出, 海洋科技已成为全球科技竞争的前沿和国家综合实力较量的焦点。笔者根据国内外海洋科技发展的特点, 结合海洋科技的内涵, 认为海洋科技水平是在现有海洋科技资源的基础上, 进行海洋科技投入—产出等生产、服务活动, 促进社会、经济、科技全面发展的总体能力。

1 海洋科技水平评价指标体系构建

本研究选择海洋科技发展基础水平、海洋科技投入水平、海洋科技产出水平以及海洋科技对社会

经济和技术发展的影响力等4个主要因素, 分别构建了海洋科研机构数量、海洋科学研究课题数量、海洋科技专利数量等14个二级指标。其中, 海洋科技发展基础水平、海洋科技投入水平体现了海洋科技的物质基础设施等硬件条件, 反映了科技创新能力的强弱和科技活动的活跃程度, 是海洋科技的基本条件和前提; 海洋科学技术产出水平、海洋科技对社会经济的影响力, 是海洋科技活动的结果, 也是海洋科技水平的外部表现。海洋科技水平评价指标体系见表1。

表1 海洋科技水平评价指标体系

主要影响因素 (一级指标)	二级指标
海洋科技发展基础水平 B_1	海洋科研机构数 C_1
	海洋专业技术人员 C_2
	从业人员数 C_3

* 本文系中国海洋发展研究中心资助课题: 海洋科技对海洋开发的影响与评价的阶段研究成果。

续表

主要影响因素 (一级指标)	二级指标
海洋科技投入水平 B_2	海洋科技课题数 C_4
	主要海洋科技产业产值占 R&D 经费支出 (亿元) 的比例 C_5 R&D 经费支出 (亿元) 增长率 C_6
海洋科技产出水平 B_3	发表海洋科技论文数 C_7
	海洋发明专利总数 C_8
	海洋发明专利授权数 C_9
	海洋科技专利申请受理数 C_{10}
	科技从业人员人均拥有发明专利数 C_{11} 海洋科技成果应用占海洋科技课题数的比例 C_{12}
海洋科技对社会经济、技术发展的影响力 B_4	主要海洋科技产业产值占地区生产总值的比重 C_{13}
	海洋从业人员占地区人员的比重 C_{14}

2 海洋科技水平评价模型

2.1 主成分分析——海洋科技水平测算模型

假设有 n 个待评价的样本点 e_1, e_2, \dots, e_n , 由前所述, 每个样本点均由 $p = 14$ 个指标变量描述, 则原始数据构成一个 $n \times p$ 维的矩阵

$$X = \begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_n^T \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = [x_1, x_2, \dots, x_p]$$

基于主成分分析的海洋科技水平测算模型, 通过把上述原始指标在没有信息损失的情况下转化为有代表意义的少数几个综合指标, 计算出这几个综合指标的加权和, 从而得到海洋科技水平的评价指数 P , 具体步骤如下:

首先, 对 X 进行标准化处理 Z , 即令

$$z_{ij} = \frac{x_{ij} - \bar{X}_j}{s_j}, \quad i=1, 2, \dots, n; j=1, 2, \dots, p$$

$$\text{其中 } \bar{X}_j = \frac{\sum_i x_{ij}}{n}, \quad s_j^2 = \frac{\sum_i (x_{ij} - \bar{X}_j)^2}{n-1};$$

其次, 令 $R = \frac{Z^T Z}{n-1}$, 求出 R 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$ ($m \leq n$) 和相应的方差贡献率 $Cr_1, Cr_2,$

\dots, Cr_K ($Cr_i = \frac{\lambda_i}{\sum_j \lambda_j}$), 其中 $\lambda_1, \lambda_2, \dots, \lambda_m$ 从大到小排列, 即 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ 。

令 $K = \min_k \{k | \sum_{i=1}^k \lambda_i / \sum_{i=1}^m \lambda_i > \theta\}$ (阈值 θ 通过实验确定, 一般取为 0.85), 得到前 K 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$, 并计算出其相应的特征向量 U_1, \dots, U_K , 这些特征向量被称为主成分。

最后, 求出每个省市的原始指标 e_i 在主成分生成的子空间 $Span(U_1, \dots, U_K)$ 上的投影, 记为 S_i , 即 $S_i = (Pa_{i1}, Pa_{i2}, \dots, Pa_{iK})^T = (U_1, \dots, U_K)^T e_i$, 其中 Pa_{ij} 表示第 i 个省市在主成分 U_j 上的得分, 则第 i 个省市的海洋科技水平评价指数可以表示为:

$$P_i = S_i^T (Cr_1, Cr_2, \dots, Cr_K) = \sum_j Cr_j Pa_{ij}.$$

2.2 灰色关联分析——海洋科技水平测算模型

假设有 n 个待评价的样本点 e_1, e_2, \dots, e_n , 由前所述, 每个样本点均由 $p = 14$ 个指标变量描述, 则原始数据构成一个 $n \times p$ 维的矩阵

$$X = \begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_n^T \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = [x_1, x_2, \dots, x_p]$$

根据所选指标确定最优指标向量 $e_0 = (x_{01}, x_{02}, \dots, x_{0p})^T$, 将最优值向量与原始指标矩阵结合, 构成灰色评价的样本空间:

$$\begin{bmatrix} x_{01} & x_{02} & \dots & x_{0p} \\ x_{11} & x_{12} & \dots & x_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

由于在进行海洋科技水平测算中所选择的指标之间有一定的数量级差, 各指标的经济意义也各不相同, 需对样本进行灰色变换以保证各指标值具有同属性:

$$z_{ij} = \frac{x_{ij}}{X_j} \quad (\bar{X}_j \text{ 是第 } j \text{ 列的平均值})$$

对所得灰色变换后的样本求差序列:

$$\delta_{\max} = \max_i \max_k |z_{0k} - z_{ik}|, \delta_{\min} = \min_i \min_k |z_{0k} - z_{ik}|$$

根据下列公式计算关联系数和关联度:

$$\xi_{ik} = \frac{\delta_{\min} + \rho\delta_{\max}}{|z_{0k} - z_{ik}| + \rho\delta_{\max}} \quad (\rho \in [0, 1], \text{ 一般}$$

取为 0.5)

最后, 根据因素相对权重 W , 求出关联度 $\gamma_i = \frac{1}{n} \sum_{k=1}^n \xi_{ik} W_k$ 。若关联度越大, 说明和最优指标越接近, 即优于其他待评对象。

3 沿海省市海洋科技水平测算

3.1 数据收集

2005 年评价指数见表 2。

表 2 2005 年沿海地区海洋科技评价指标值

项目	天津 X1	河北 X2	辽宁 X3	上海 X4	江苏 X5	浙江 X6	福建 X7	山东 X8	广东 X9	广西 X10	海南 X11
C1	10	2	12	10	7	12	10	16	16	2	1
C2	2 261	59	761	1 456	1 229	881	620	2 669	1 390	57	32
C3	1 486	49	632	1 191	908	724	598	1 914	1 085	51	29
C4	318	18	94	393	920	315	309	836	404	28	14
C5	10.20	3.83	6.54	5.84	1.95	10.31	20.01	9.79	9.92	9.92	95.36
C6	0.35	0.344	0.166	0.218	0.261	0.413	0.168	0.373	0.154	0.23	-0.236
C7	276	37	117	294	536	319	289	1126	519	30	9
C8	20	2	1	8	10	10	15	163	41	0	0
C9	10	2	0	8	2	4	4	63	26	2	0
C10	27	0	0	42	2	27	17	176	20	0	0
C11	0.008 8	0.034	0.001 3	0.005 5	0.008 1	0.011	0.024	0.061	0.029	0	0
C12	0.26	0.00	0.22	0.20	0.25	0.12	0.17	0.04	0.05	0.14	1.00
C13	0.200 35	0.022 34	0.101 77	0.132 88	0.028 77	0.125 35	0.163 26	0.103 17	0.108 16	0.031 55	0.170 55
C14	0.240 71	0.021 17	0.111 56	0.107 04	0.060 20	0.025 96	0.024 56	0.080 33	0.026 05	0.015 77	0.037 05

3.2 基于主成分分析的海洋科技水平测算

将表 2 中的沿海 11 省市的海洋科技水平评价指标代入基于主成分分析的海洋科技水平测算模型中, 求出相应的特征值 (表 3)。表 3 中前 4 个成

分的方差贡献率的总和已经超过了 92.257%, 包含了原始数据中的绝大部分有用信息, 所以取这 4 个成分作为主成分。各主成分在总的海洋科技水平中的权重取为各主成分的方差贡献率, 即:

$$Cr^T = (54.619, 18.749, 13.551, 5.338)$$

表 3 海洋科技评价指标方差

原始特征值数据			
成份	特征值	方差贡献率	累计贡献率
1	7.647	54.619	54.619
2	2.625	18.749	73.368
3	1.897	13.551	86.919
4	0.747	5.338	92.257
5	0.576	4.112	96.370
6	0.233	1.661	98.031
7	0.164	1.172	99.203
8	5.602E-02	0.400	99.603
9	3.786E-02	0.270	99.873
10	1.772E-02	0.127	100.000
11	1.507E-16	1.076E-15	100.000
12	6.781E-17	4.844E-16	100.000
13	-1.147E-16	-8.194E-16	100.000
14	-5.735E-16	-4.096E-15	100.000

为了更好地解释各主成分,需要计算出主成分的因子载荷矩阵。一般的,初始因子载荷矩阵并不能直观的说明各变量对主成分的影响程度,还需对初始因子载荷矩阵进行正交旋转,表 4 是旋转前后的因子载荷矩阵。由表 4 可以看到:第一个主成分主要由 C7、C8、C9、C10、C11 (表中阴影部分,下同) 决定,反映的是海洋科技产出水平,54.619% 的方差贡献率说明这一主成分对海洋科技水平具有举足轻重的作用。第二个主成分主要由 C5、C6、C12 决定,反映的主要是海洋科技投入的高低,18.749% 的方差贡献率说明这一主成分对海洋科技水平具有较大的影响。第三个主成分主要由 C1、C2、C3、C13、C14 决定,方差贡献率小于前两个主成分,对海洋科技水平的影响较弱。第四个主成分主要由 C4 决定,这个指标在 4 个主成分中方差贡献率最小,对海洋科技水平的影响已经微乎其微。

表 4 主成分的因子载荷矩阵

	正交旋转前				正交旋转后			
	1	2	3	4	1	2	3	4
C1	0.791	7.4E-02	0.218	-0.135	0.570	0.346	0.469	0.180
C2	0.911	0.231	0.315	8.8E-02	0.589	0.260	0.618	0.441
C3	0.898	0.250	0.331	8.8E-02	0.575	0.244	0.636	0.438
C4	0.766	9.2E-03	-2.2E-02	0.611	0.452	0.184	9.8E-02	0.844
C5	-0.487	0.775	-0.379	-7.8E-03	-2.8E-02	-0.970	-4.8E-02	-0.198
C6	0.592	-0.640	0.313	-0.118	0.208	0.895	0.119	0.116
C7	0.944	9.8E-02	-0.210	0.215	0.815	0.178	0.123	0.529
C8	0.887	0.193	-0.369	-8.5E-02	0.952	6.9E-02	8.5E-02	0.219
C9	0.892	0.183	-0.342	-0.126	0.952	9.5E-02	0.111	0.186
C10	0.879	0.218	-0.271	-0.118	0.911	8.5E-02	0.180	0.195
C11	0.758	-0.106	-0.480	-0.283	0.896	0.262	-0.154	-2.4E-02
C12	-0.552	0.778	-0.127	0.229	-0.270	-0.947	8.8E-02	1.7E-02
C13	5.7E-02	0.792	0.389	-0.332	0.111	-0.447	0.788	-0.244
C14	0.289	0.348	0.794	-2.4E-02	-8.8E-02	0.105	0.891	0.147

最后,计算出各主成分的特征向量 U_i ($i=1, 2, 3, 4$), 则给定某省市的指标值 e 的情况下,该省市在第 i 主成分 U_i 上的得分为 $Pa_i = U_i^T e$, 从而

该省市的海洋科技水平测算指标可以表示为:
 $P = Cr^T \cdot (Pa_1, Pa_2, Pa_3, Pa_4)^T = (19.9749, 25.7099, 25.9285, 14.8786, -6.7207, 9.3011,$

16.417 6, 13.845 7, 14.261 8, 15.546 2, 6.057 6,
3.932 2, 16.535 5, 22.451 1) · e

11个省市在主成分的得分情况及其海洋科技
水平测算结果总结见表5。

表5 各主成分的得分值和综合得分值

地区	Pa_1		Pa_2		Pa_3		Pa_4		海洋科技水平测算指标	
	得分	排名	得分	排名	得分	排名	得分	排名	得分	排名
天津	1.128	3	1.356	2	2.979	1	4.747	1	152.725	2
河北	-1.653	9	-2.453	11	-1.004	9	-1.600	9	-158.424	9
辽宁	-1.131	8	-0.220	7	1.116	3	1.778	3	-41.307	8
上海	0.355	4	0.285	4	1.168	2	1.861	2	50.499	4
江苏	0.329	5	-1.009	9	0.120	5	0.190	5	1.702	5
浙江	0.039	6	-0.796	8	0.369	4	0.587	4	-4.635	6
福建	-0.498	7	0.163	5	-0.287	6	-0.458	6	-30.486	7
山东	6.664	1	0.884	3	-1.778	11	-2.834	11	341.343	1
广东	1.469	2	-0.111	6	-0.497	7	-0.792	7	67.206	3
广西	-2.436	10	-1.693	10	-0.613	8	-0.977	8	-178.286	10
海南	-4.267	11	3.594	1	-1.571	10	-2.502	10	-200.337	11
α	54.619		18.746		13.551		5.338		92.257	

由此, 各省市海洋科技水平的排名从高到低依次
为: 山东、天津、广东、上海、江苏、浙江、福
建、辽宁、河北、广西、海南。

3.3 基于灰色关联分析法的海洋科技水平测算

按照基于灰色关联分析的海洋科技水平测算模
型, 首先将选取出最优指标集。注意到, 海南省的
海洋科技产业产值占 R&D 经费支出的比例指标是
其他各省的数十倍, 这主要是因为海南省的 R&D
经费支出过少, 导致分母偏小造成的。显然, 海南

省的该项指标值并不能代表该指标的最优值, 所以
这里我们选取折中的方法, 取中间指标值 10 作为
最优值。其他 13 个指标都是正指标, 即越大越好,
因此取各个省市中该指标的最大值作为最优值, 最
后的最优指标集为: $e_0 = (16, 1\ 914, 2\ 669,$
 $920, 10, 41.3, 1\ 126, 163, 63, 176, 0.061\ 072,$
 $0.26, 0.200\ 350, 0.240\ 709)$ 。

其次, 对最优指标集在内的所有指标向量进行
归一化处理并在此基础上求出各指标的关联系数矩
阵, 结果如表 6 所示。

表6 关联系数矩阵

项目	天津 X1	河北 X2	辽宁 X3	上海 X4	江苏 X5	浙江 X6	福建 X7	山东 X8	广东 X9	广西 X10	海南 X11
C1	0.555 6	0.348 8	0.652 2	0.555 6	0.454 5	0.652 2	0.555 6	1.000	1.000	0.348 8	0.333 3
C2	0.687 7	0.335 7	0.423 7	0.565 9	0.483 7	0.442 0	0.417 3	1.000	0.532 0	0.335 9	0.333 3
C3	0.763 7	0.335 6	0.408 6	0.520 8	0.478 0	0.424 4	0.391 5	1.000	0.507 6	0.335 5	0.333 3
C4	0.429 4	0.334 3	0.354 2	0.462 2	1.000	0.428 2	0.425 8	0.843 6	0.467 5	0.336 8	0.333 3
C5	0.995 6	0.883 3	0.931 0	0.918 2	0.853 0	0.993 3	0.823 5	0.995 6	0.998 4	0.975 1	0.353 6
C6	0.837 4	0.824 7	0.567 8	0.624 6	0.681 0	1.000	0.569 8	0.890 3	0.556 1	0.639 4	0.333 3
C7	0.396 5	0.339 0	0.356 3	0.401 7	0.486 3	0.409 0	0.400 2	1.000	0.479 2	0.337 6	0.333 3
C8	0.363 0	0.336 1	0.334 7	0.344 6	0.347 5	0.347 5	0.355 1	1.000	0.400 5	0.333 3	0.333 3

续表

项目	天津 X1	河北 X2	辽宁 X3	上海 X4	江苏 X5	浙江 X6	福建 X7	山东 X8	广东 X9	广西 X10	海南 X11
C9	0.372 8	0.340 5	0.333 3	0.364 2	0.340 5	0.348 1	0.348 1	1.000	0.459 9	0.340 5	0.333 3
C10	0.371 3	0.333 3	0.333 3	0.392 9	0.335 9	0.371 3	0.356 3	1.000	0.360 7	0.333 3	0.333 3
C11	0.369 0	0.529 1	0.338 2	0.354 6	0.365 8	0.380 5	0.453 0	1.000	0.491 6	0.333 3	0.333 3
C12	0.403 2	0.333 3	0.390 6	0.384 6	0.400 0	0.362 3	0.375 9	0.342 5	0.344 8	0.367 6	1.000
C13	1.000	0.333 3	0.474 5	0.568 8	0.3416	0.542 7	0.705 9	0.478 0	0.491 2	0.345 2	0.749 2
C14	1.000	0.338 8	0.465 5	0.456 9	0.383 9	0.343 7	0.342 2	0.412 2	0.343 8	0.333 3	0.355 8

最后, 计算关联度。关联系数的权重 W 取为主成分分析法中的指标权重, 即 $W = (0.252 2, 0.364 5, 0.383 0, -0.025 9, -0.438 5, 0.362 1, 0.243, 0.426 9, 0.395 7, 0.313 6, -0.555 4, -0.146 9, 0.450 1, 0.918 7)$ 。注意到, 当中有负值, 由于其代表的是对主成分的影响程度, 因此应该把各权重取为绝对值, 最后归一化到在 0~1 范

围内得到新的权重向量 $W = (0.094 4, 0.121 5, 0.122 6, 0.070 3, 0.031 8, 0.044 0, 0.077 6, 0.065 4, 0.067 4, 0.073 5, 0.028 6, 0.018 6, 0.078 2, 0.106 1)$ 。将权重向量的转置矩阵与上表灰色关联度矩阵相乘, 可得出各个省市的灰色加权关联度 (表 7)。

表 7 沿海地区海洋科技水平关联度

	天津	河北	辽宁	上海	江苏	浙江	福建	山东	广东	广西	海南
关联度	0.637 5	0.381 8	0.442 5	0.491 6	0.481 1	0.474 7	0.474 7	0.868 6	0.526 6	0.371 8	0.381 3

因此, 各省市海洋科技水平的排名从高到低依次为: 山东、天津、广东、上海、江苏、浙江、福建、辽宁、河北、海南、广西。

3.4 测评结果分析

通过两种测算模型得到的结果基本一致。山东

的评测指数在两个测算模型中均要远高于其他省市, 山东省在海洋科技上的投入水平和科技产出水平高于大部分省市。海南和广西在两个测算模型中的排名虽然不一样, 但这两个地区在两个测算模型中的评测指数都非常接近。表 4 中, 海南和广西在各个主成分上的得分都接近最后。

参考文献

- [1] 何晓群. 现代统计分析方法与应用 [M]. 北京: 中国人民大学出版社, 1998: 281-343.
- [2] 伍业锋, 施平. 中国沿海地区海洋科技竞争力分析与排名 [J]. 上海经济研究, 2006 (2): 26-33.
- [3] 陈光潮, 张辉, 韩建安. 基于灰色系统理论的区域科技竞争力比较 [J]. 暨南大学学报, 2004 (2): 19-25.
- [4] 李桂香. 海洋科学技术在海洋可持续发展中的地位和作用 [J]. 海洋信息, 1998 (12): 27-28.
- [5] 韩伟, 李刚. 主成分分析在地区科技竞争力评测中的应用 [J]. 数理统计与管理, 2006, 25 (5): 512-517.