

灌区节水改造技术经济指标的综合主成分分析

姚杰, 郭宗楼, 陆琦

(浙江大学 农业工程系, 浙江 杭州 310029)

摘要: 本文提出了用于评价灌区节水改造效益的6个指标, 应用主成分分析原理建立了灌区节水改造效益评价的综合指标-综合主成分。对样本灌区节水改造效益的综合主成分进行的正态性检验结果表明, 它服从正态分布。在此基础上提出了相应的评价标准。应用结果表明, 本文提出的灌区节水改造效益的综合主成分及其评价标准能较全面地反映灌区节水改造状况, 较好地描述灌区节水改造的效益水平。

关键词: 灌区; 节水改造效益; 综合评价指标; 主成分分析

中图分类号: S274.1 **文献标识码:** A

我国现有灌溉面积0.533亿 hm^2 , 然而, 我国的灌区正面临着水资源紧缺、水质恶化, 资金不足、老化失修, 体制不顺、机制不全, 规划落后、决策不周等问题。这些问题也是世界性的难题。目前党中央和政府已经十分关注全国大型灌区的问题, 正在着手进行节水改造规划, 以解决这一难题。灌区节水改造效益如何, 需进行正确评价, 通过其评价, 可查明项目优劣及其原因, 总结教训, 提出改进措施与建议, 提高灌区节水改造的效益。

1 灌区技术改造效益评价指标

科学的灌区节水改造效益评估方法是确定一系列技术经济指标, 进行定量评估, 为此, 本文提出以下指标: (1)公顷均投资(元/ hm^2): 灌区节水改造单位面积总投资; (2)成本水价(元/ m^3): 灌区节水改造后重新核实的成本水价; (3)水分生产效率(kg/m^3): 灌区节水改造后的单位水量的粮食总生产量; (4)公顷均效益(元/ hm^2): 灌区节水改造后单位面积的净效益; (5)公顷均节约用水量(m^3/hm^2): 灌区节水改造后与改造前比较的公顷均节约水量; (6)灌溉水利用系数: 灌区节水改造后灌入田间的水量与渠道引入总水量的比值。若以上指标均采用节水改造前后的比较值, 可以更充分地反映节水改造效益, 但由于改造前的资料往往不全, 因此大部分指标采用了节水改造后的指标值。尽管如此, 采用上述指标仍能反映灌区节水改造的总体效益, 通过灌区间的比较, 更可以反映不同灌区节水改造效益的优劣。本文将应用上述6个指标, 应用综合主成分分析方法对灌区节水改造效益进行综合评价。

2 综合主成分分析原理

主成分分析^[1]是对数据和变量结构进行分析处理的一种行之有效的多元统计分法, 它可以在不损失或尽量少损失原有指标信息的情况下, 将多个具有相关性的指标转换成少数几个互相独立的综合指标(即主成分), 为资料的后续分析提供方便。

在进行综合评价时, 先将原始指标转换成趋势性相同的若干个主成分, 并对这些主成分进行一定的线

收稿日期: 2003-06-06

基金项目: 国家高技术研究发展规划资助项目(2001AA242121)

作者简介: 姚杰(1979-), 男, 浙江鄞县人, 硕士生, 主要从事灌区节水工程和灌区地理信息系统研究。

性组合，以构造出综合主成分。然后，依据各评价对象在综合主成分上得分的大小排出其优劣次序，达到综合评价的目的。

假定有n个待评价对象(在此为灌区数)，每个对象均观察p项指标(在此p=6)，记为 X_1, X_2, \dots, X_p 。通过研究观测，获得的数据矩阵X为

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1j} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2j} & \cdots & X_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{i1} & X_{i2} & \cdots & X_{ij} & \cdots & X_{ip} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{nj} & \cdots & X_{np} \end{bmatrix}$$

为了获得综合主成分，具体步骤如下。

(1) 评价指标的同趋势化处理。为了保证求得的主成分具有相同的趋势性，需先对p个评价指标进行同趋势化处理，以保证其方向的一致性。常用的方法是将低优指标采用倒数法转化为高优指标，公式可为

$$X'_{ij} = \frac{1}{X_{ij}} \text{ 或 } X'_{ij} = -X_{ij} \quad (1)$$

式中： $i=1, 2, \dots, n$ 为评价对象序号； $j=1, 2, \dots, p$ 为指标序号(下同)； X_{ij} 表示第i个评价对象在j个指标上的原始取值； X'_{ij} 表示 X_{ij} 同趋势化处理后的值。

(2) 评价指标的标准化处理。对已作同趋势性处理的评价指标 $X_j = (X_{1j}, X_{2j}, \dots, X_{nj})^T$ 再进行标准化处理，以消除指标量纲的影响，公式为

$$Z_{ij} = \frac{X'_{ij} - \bar{X}'_j}{S_j} \quad (2)$$

式中： Z_{ij} 为 X'_{ij} 的标准化值； $\bar{X}'_j = \frac{\sum_{i=1}^n X'_{ij}}{n}$ ； $S_j = \sqrt{\frac{\sum_{i=1}^n (X'_{ij} - \bar{X}'_j)^2}{n-1}}$ 。

(3) 计算已作标准化处理的评价指标 $Z_j = (Z_{1j}, Z_{2j}, \dots, Z_{nj})^T$ 的相关系数矩阵R，其中

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix}, \text{ 其元素 } r_{jk} \text{ 表示原变量 } X_j \text{ 与 } X_k \text{ 的相关系数, 计算公式为}$$

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2} \sqrt{\sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad i, j=1, 2, \dots, p$$

(4) 计算R的特征根和特征向量，并得到相应的主成分。将R的所有特征根中，大于或近似等于1的特征根按降序排列，记为 $\lambda_1, \lambda_2, \dots, \lambda_m$ ，且与 $\lambda_k (k=1, 2, \dots, m)$ 相应的特征向量记为 $\mathbf{v}_k = (v_{k1}, v_{k2}, \dots, v_{km})^T$ 。这里与 λ_k 相应的主成分即称为第k个主成分，记为 f_k 。由于评价指标已作了高优化的同趋势性处理，因此，所求得的m个主成分也具有相同的高优化趋势性。

(5) 计算各主成分的贡献率及得分。第k个主成分 f_k 的贡献率为

$$\alpha_k = \lambda_k / p \quad (3)$$

其得分为

$$f_k = \alpha_{k1} Z_1 + \alpha_{k2} Z_2 + \dots + \alpha_{kp} Z_p \quad (4)$$

$Z_i (i=1, 2, \dots, p)$ 为标准化处理后的评价指标。

(6) 构造综合主成分。以贡献率为权数，将所求得的m个主成分进行线性组合，即可构造出所谓的综合主成分F，其计算公式为

$$F = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_n f_n \quad (5)$$

(7) 评价。计算各评价对象在综合主成分上的得分，记为 $F_i (i=1, 2, \dots, n)$ 。得分越高，相应的评价对象就越优。从而，按其综合主成分的得分大小即可将评价对象排出优劣次序，完成综合评价的目的。

另外，还可以选择n个有代表性的评价对象，应用上述方法计算出综合主成分，然后应用数理统计原理对其进行统计分析，用以推导出对评价对象的评价标准，具体方法如下：

(1) 对综合主成分 $F_i (i=1, 2, \dots, n)$ 进行正态性检验。通过样本 F_1, F_2, \dots, F_n 来检验总体分布是否服从正态分布，其检验方法较多，本文采用柯尔莫哥洛夫-斯米尔洛夫检验法，并应用权威统计软件SAS 8.0直接进行计算。若通过检验，证明样本 F_1, F_2, \dots, F_n 是服从正态分布，则说明它能反映一般规律，具有较好的代表性，并可用它来建立综合主成分的评价标准。

(2) 综合主成分的评价标准。根据正态分布计算出不同概率 $P(\lambda)$ 的综合主成分，计算公式为

$$F = \bar{F} - \lambda \sigma \quad (6)$$

式中： \bar{F} 为综合主成分的平均值； σ 为综合主成分的标准差； λ 为系数，从正态分布表（见有关概率统计的书籍）中，可以查出对应于概率 $P(\lambda)$ 值的 λ 值。

因此，可以根据给定的概率值，确定相应水平的综合主成分，以此来评价灌区节水改造效益的等级。

3 应用

根据全国大型灌区续建配套与节水改造规划资料，从400个对象灌区中选择资料较全，并具有代表性的182个灌区，进行综合主成分分析评价。

3.1 评价指标的同趋势化处理 在每公顷平均投资、成本水价、水分生产效率、每公顷平均效益、每公顷平均节约用水量、灌溉水利用系数等6个评价指标（分别用 X_1, X_2, \dots, X_6 表示）中，除每公顷平均投资、成本水价为低优指标外，其他均为高优指标。对低优指标用取负的方法作同趋势化处理，趋势化处理后的统计值见表1。

表1 趋势化处理后的统计值

项目	X_1	X_2	X_3	X_4	X_5	X_6
平均值	-8793.1963	-0.1329	1.4638	6548.3188	3824.1453	0.6246
标准差	3521.8809	0.0623	0.2101	6552.4937	2675.3501	0.0669

3.2 标准化处理后评价指标的相关系数矩阵R 见表2

表2 相关系数矩阵R

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.0000	-0.1047	0.1515	0.2087	-0.4363	-0.4266
X_2	-0.1047	1.0000	0.0147	-0.0864	0.1526	-0.0816
X_3	0.1515	0.0147	1.0000	0.0460	-0.0463	-0.0356
X_4	0.2087	-0.0864	0.0460	1.0000	0.1369	-0.4047
X_5	-0.4363	0.1526	-0.0463	0.1369	1.0000	0.0340
X_6	-0.4266	-0.0816	-0.0356	-0.4047	0.0340	1.0000

3.3 计算R的特征根、特征向量(a_i)、相应的主成分(资料略)、各主成分的贡献率及得分、综合主成分

表3 特征根、特征向量(a_i)、各主成分的贡献率(按特征根大小排序后)

	a_1	a_2	a_3	a_4	a_5	a_6
X_1	0.6218	-0.2180	0.0930	-0.1365	-0.0959	0.7276
X_2	-0.1281	0.3425	0.7190	-0.4961	0.3139	0.0686
X_3	0.1810	-0.0689	0.6147	0.7544	-0.0410	-0.1178
X_4	0.4036	0.5108	-0.3003	0.2266	0.6588	0.0242
X_5	-0.3268	0.6579	-0.0686	0.2467	-0.4085	0.4775
X_6	-0.5427	-0.3698	-0.0403	0.2324	0.5383	0.4726
特征根	1.8029	1.3161	1.0252	0.9506	0.5100	0.3952
贡献率	0.3005	0.2193	0.1709	0.1584	0.0850	0.0659
累计贡献率	0.3005	0.5198	0.6907	0.8491	0.9341	1.0000

表4 综合主成分(按主成分大小排序后)

灌区 序号	综合 主成分	灌区 序号	综合 主成分	灌区 序号	综合 主成分	灌区 序号	综合 主成分	灌区 序号	综合 主成分
1	1.6227	38	0.3886	75	0.0942	112	-0.1945	149	-0.4368
2	1.5749	39	0.3787	76	0.0705	113	-0.2013	150	-0.4371
3	1.4454	40	0.3735	77	0.0580	114	-0.2220	151	-0.4397
4	1.4192	41	0.3733	78	0.0534	115	-0.2297	152	-0.4573
5	1.2616	42	0.3639	79	0.0485	116	-0.2568	153	-0.4677
6	1.2278	43	0.3514	80	0.0424	117	-0.2623	154	-0.4752
7	0.1221	44	0.3426	81	0.0283	118	-0.2662	155	-0.4836
8	1.1046	45	0.3393	82	0.0119	119	-0.2686	156	-0.4933
9	1.0600	46	0.3370	83	-0.0025	120	-0.2736	157	-0.5085
10	1.0489	47	0.3232	84	-0.0151	121	-0.2885	158	-0.5099
11	1.0143	48	0.2957	85	-0.0184	122	-0.3008	159	-0.5256
12	0.8635	49	0.2910	86	-0.0205	123	-0.3189	160	-0.5528
13	0.8135	50	0.2815	87	-0.0235	124	-0.3314	161	-0.5764
14	0.7972	51	0.2814	88	-0.0329	125	-0.3339	162	-0.5823
15	0.7325	52	0.2789	89	0.0331	126	-0.3436	163	-0.5864
16	0.7124	53	0.2710	90	-0.0352	127	-0.3539	164	-0.6147
17	0.6847	54	0.2679	91	-0.0357	128	-0.3629	165	-0.6162
18	0.6397	55	0.2606	92	-0.0429	129	-0.3637	166	-0.6194
19	0.6293	56	0.2587	93	-0.0442	130	-0.3643	167	-0.6349
20	0.6069	57	0.2082	94	-0.0536	131	-0.3653	168	-0.6431
21	0.5924	58	0.1805	95	-0.0556	132	-0.3667	169	-0.6855
22	0.5819	59	0.1725	96	-0.0578	133	-0.3728	170	-0.7036
23	0.5781	60	0.1717	97	-0.0772	134	-0.3758	171	-0.7128
24	0.5631	61	0.1656	98	-0.0847	135	-0.3819	172	-0.7710
25	0.5500	62	0.1577	99	-0.0849	136	-0.3852	173	-0.7923
26	0.5495	63	0.1544	100	-0.0880	137	-0.3917	174	-0.8281
27	0.5487	64	0.1455	101	-0.0923	138	-0.3938	175	-0.8975
28	0.5385	65	0.1389	102	-0.0929	139	-0.3978	176	-0.9119
29	0.5194	66	0.1364	103	-0.0979	140	-0.4045	177	-0.9131
30	0.5050	67	0.1289	104	-0.1318	141	-0.4047	178	-0.9488
31	0.4950	68	0.1250	105	-0.1342	142	-0.4062	179	-0.9724
32	0.4875	69	0.1197	106	-0.1405	143	-0.4145	180	-0.9848
33	0.4821	70	0.1186	107	-0.1506	144	-0.4265	181	-1.0660
34	0.4779	71	0.1140	108	-0.1594	145	-0.4266	182	-1.2330
35	0.4566	72	0.1051	109	-0.1654	146	-0.4272		
36	0.4414	73	0.0972	110	-0.1656	147	-0.4273		

3.4 综合主成分的正态性检验 对表4的综合主成分(均值为0,标准差为0.5341),采用柯尔莫哥洛夫-斯米尔洛夫检验,求得检验统计量 $D_{182} = 0.0567$ 。取显著性水平 $\alpha = 0.05$,查得 $D_{182, \alpha} = 0.0657$,因为

0.0567 < 0.0657, 故在水平 = 0.05下, 认为综合主成分服从正态分布 $N(0, 0.5341^2)$ 。实际上, 显著性水平 $> 15\%$ 。

3.5 综合主成分的评价标准 根据正态分布, 确定概率P () 的综合主成分为: $F = - = 0 - 0.5341$, 计算的不同概率P () 的综合主成分列于表5。

表5 不同概率P () 的综合主成分

概率(%)	综合主成分	概率(%)	综合主成分	概率(%)	综合主成分	概率(%)	综合主成分
5	0.8786	30	0.2801	55	-0.0671	80	-0.4495
10	0.6845	35	0.2058	60	-0.1353	85	-0.5536
15	0.5536	40	0.1353	65	-0.2058	90	-0.6845
20	0.4495	45	0.0671	70	-0.2801	95	-0.8786
25	0.3603	50	0.0000	75	-0.3603		

从上表可知, 灌区节水改造效益为平均水平 (取 $p=50\%$) 的综合主成分为0.00; 灌区节水改造效益为平均先进水平 (取 $p=25\%$) 的综合主成分为0.3603; 灌区节水改造效益为先进水平 (取 $p=10\%$) 的综合主成分为0.6845。

3.6 应用 选用10个大型灌区的实际资料 (见表6), 趋势化处理后应用表1的资料进行标准化, 然后应用表3资料计算综合主成分 (见表7), 根据表5资料给出评价等级。

表6 实际灌区资料

灌区编号	每公顷投资 /(元/hm ²)	成本水价 /(元/m ³)	水分生产效率 /(kg/m ³)	每公顷效益 /(元/hm ²)	每公顷节约用水量 /(m ³ /hm ²)	灌溉水利用系数
01	8895	0.0950	1.42	18000	10005	0.5200
02	4126.5	0.1400	1.65	1572	2200.5	0.6500
03	7219.5	0.1630	1.64	2655	2760	0.6000
04	6765	0.1430	1.66	2670	2820	0.6000
05	6339	0.1200	1.51	2448	2563.5	0.6400
06	6360	0.1030	1.40	21195	7410	0.5200
07	4905	0.0500	1.80	9750	4702.5	0.6670
08	6765	0.1180	1.60	10770	1695	0.5500
09	6255	0.1650	1.58	20715	2220	0.6800
10	7507.5	0.1800	1.36	18495	3573.75	0.6000

表7 综合评价结果

灌区编号	综合主成分	评价	排名	灌区编号	综合主成分	评价	排名
06	1.0143	先进	1	10	0.3942	平均先进	6
01	0.8635	先进	2	04	0.1805	平均	7
07	0.8135	先进	3	02	0.1389	平均	8
09	0.6397	平均先进	4	03	0.0959	平均	9
08	0.5631	平均先进	5	05	-0.0357	低于平均	10

4 结论

(1) 在主成分分析的基础上提出的灌区节水改造效益评价的综合指标-综合主成分,是在不损失或尽量少损失原有指标信息的情况下,将多个具有相关性的指标转换成一个综合指标,能较好的描述灌区节水改造效益水平。(2)综合主成分的正态性检验结果表明,灌区节水改造效益的综合主成分指标较强地服从正态分布,说明它能反映一般规律,具有较好的代表性。提出的综合主成分的评价标准具有实用性和可操作性。(3)应用综合主成分分析方法,把多个指标综合为一个综合指标,还可以具此对多个评价对象进行排队,评价其优劣。

参考文献:

[1] 袁志发,周静芋.多元统计分析[M].北京:科学出版社,2002.

Comprehensive principal components analysis on technical and economic index of water saving reform for irrigation area

YAO Jie, GUO Zong-lou, LU Qi
(Zhejiang University, Hangzhou 310029, China)

Abstract: Six indexes for benefit assessment of water saving reform for irrigation area are suggested. On this basis, the comprehensive principal analysis method for assessing the benefit of water saving reform is established. The analysis on sampling shows that it obeys the law of normal distribution. The corresponding standards for assessment are deduced accordingly. The application shows that the proposed method can comprehensively reflect the result of reform.

Key words: irrigation area; benefit of water saving reform; comprehensive assessment indexes; principal component analysis