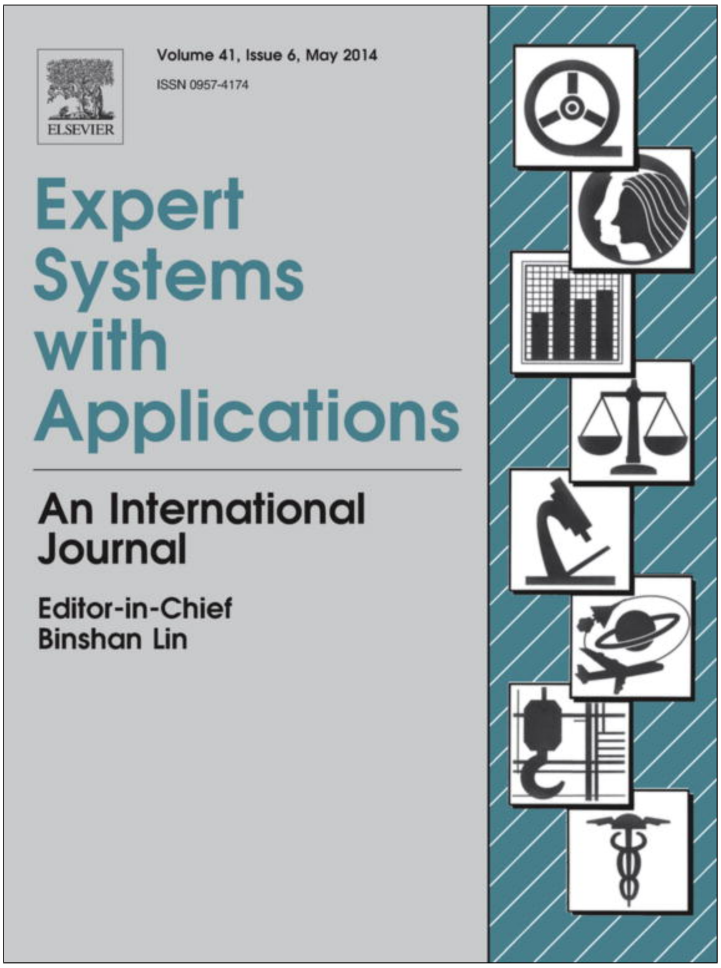
Volume 41, Issue 6, May 2014
ISSN 0957-4174
ELSEVIER
Expert Systems with Applications
An International Journal
Editor-in-Chief
Binshan Lin

Expert Systems with Applications 41 (2014) 2824–2831

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Evaluation of the provincial competitiveness of the Chinese high-tech industry using an improved TOPSIS method

Zheng-Xin Wang [a,*], Yan-Yu Wang [b]

[a] School of Economics & International Trade, Zhejiang University of Finance & Economics, Hangzhou 310018, China
[b] School of Economic and Management, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Keywords:
Competitiveness evaluation
High-tech industry
Mahalanobis distance
TOPSIS
Information overlap

ABSTRACT

Evaluation of the competitiveness of high-tech industry is a technical decision-making issue involving multiple criteria. It is also a practical path to promote a country's competitiveness. However, the competitiveness indicators in high-tech industry often act and react upon one another. Moreover, different dimensions and indicator weights also affect the evaluation results. In this paper, the Mahalanobis distance is used to improve the traditional technique for order preference by similarity to ideal solution (TOPSIS). The improved TOPSIS method has the following properties: (1) an improved relative closeness which is invariant after non-singular linear transformation, and (2) the weighted Mahalanobis distance is the same as the weighted Euclidean distance when the indicators are uncorrelated. The new method is applied to evaluate the competitiveness of the Chinese high-tech industry using data from 2011. Consideration of the correlation between indicators improves the evaluation results (in terms of sorting and closeness) to a certain extent compared to the traditional TOPSIS method. The top five provinces are: Guangdong, Jiangsu, Shanghai, Beijing, and Shandong. This finding reflects the practical linkage among provinces and softens the closeness value, consistent with reality.

1. Introduction

High-tech industries, which are based on intellectually-intensive technologies and integrate multidisciplinary technological achievements, are the strategic leading industries of the Chinese national economy. The statistical range of the Chinese high-tech industry includes five categories: aerospace manufacture, electronics and communications equipment manufacture, computer and office equipment manufacture, pharmaceuticals and medical equipment manufacture, as well as instrument and meter manufacture. High-tech industries are important because they drive the world's economic layout, political affairs, and military competition. Development of high-tech industry has become a concrete expression of the strength of a nation or region (Lu & Yu, 2010). Since the implementation of 'Torch Plan' (the national high-tech industrial development plan), the Chinese high-tech industry has made remarkable achievements. The evaluation of the competitiveness of provincial high-tech industry has become the basis for decision-making for the national high-tech industrial layout. Moreover, such an evaluation broadens our understanding of the geographical distribution and development status of the Chinese high-tech industries and provides rational suggestions for the promotion and planning of them.

Liang (2011) proposed that high-tech industries with high investment, high growth, high yield, and high risk should have the following general characteristics. They have (1) a high degree of uncertainty, (2) high-value with regard to human resources, and (3) a highly correlated value of intangible assets. Studies on the evaluation of the competitiveness of Chinese provincial high-tech industries have attracted the attention of many researchers. Chen and Sun (2011) used factor and cluster analyses to evaluate the competitiveness of Chinese provincial high-tech industry. They also provided a classification system while establishing evaluation indicators that included the level of human capital investment, the level of project organization investment, the level of capital investment, the level of industrial output, and the level of efficiency.

Wu and Li (2008) introduced the technique now traditionally used to evaluate the competitiveness of high-tech industry called the technique for order preference by similarity to ideal solution (TOPSIS) method. They applied it to 31 Chinese provincial administrative regions and built up the concepts of core competitiveness *within* the industry (industrial core technical capabilities, industrial core production capacity, and industrial core market power) and core competitiveness *outside* the industry (industrial policy environment, industry technical support environment, and industry incubator environment). They sorted the top six regions

* Corresponding author.
E-mail address: jenkin107@hotmail.com (Z.-X. Wang).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.10.015

as follows: Beijing, Guangdong, Shanghai, Zhejiang, Shandong, and Jiangsu. Later, Zheng, Shi, and He (2010) made a comprehensive competitiveness evaluation of the high-tech industry in Fujian Province. The proposed evaluation indicators included technological innovation competitiveness, economic development competitiveness, financial benefit competitiveness, industrial cluster competitiveness, and energy saving and environmental protection competitiveness. Chen (2010) considered the use of data mining methods (specifically k-means clustering) to evaluate the competitiveness of Chinese high-tech industries. Based on this method, the top six regions in terms of high-tech industry competitiveness were found to be: Guangdong, Jiangsu, Beijing, Shanghai, Tianjin, and Liaoning.

Numerous studies on the evaluation of Chinese provincial high-tech industrial competitiveness have been conducted, but two issues remain unclear: (1) From the perspective of evaluation object characteristics, the differences among the provinces in economy, geography, etc. have made the development of high-tech industries unbalanced. However, the inter-provincial economic circle drives the linkage between the high-tech industries in different provinces, which is a prominent feature of provincial high-tech industry. Most studies on the competitiveness of Chinese high-tech industry have been aware of this unbalanced status quo. For example, Wang and Yu (2004) used principal component analysis to conclude that the competitiveness of Chinese high-tech industry in western regions is weaker than in the eastern regions and that the gap is gradually increasing. Wang (2007), Liang, Li, Tang, and Zhao (2007) and Sun, Xiong, and Zheng (2010) introduced empirical methods and also obtained the result that the development in high-tech industry is very unbalanced among different regions. They also found that the size of the imbalance is increasing. Although this imbalance problem is recognized, it is merely at the phenomenon level at present. Also, the majority of existing evaluation methods assume that the samples are independent and identically distributed. An imperative issue is how to take the index linkage problem into consideration through method design and thus to improve the scientific basis of the decision-making. (2) From the perspective of evaluation method, consensus has not been reached on the true evaluation of the competitiveness results for Chinese provincial high-tech industry. The differences in evaluation index systems may partly explain this situation, but more differences are found in terms of the evaluation method itself. Although the existing evaluation methods are based on the characteristics of the collected data, they all have advantages and disadvantages. TOPSIS and fuzzy methods do not consider the correlation between evaluation indicators, which often results in information overlap. Secondly, factor analysis all too easily makes the economic significance of the main components ambiguous when the factor loadings of the core variables are small. In addition, analytic hierarchy processes (AHPs) can hardly avoid deviation in subjective factors.

As a typical, uncertain multiple-criteria decision-making (MCDM) problem, evaluation of the competitiveness of the Chinese high-tech industry includes mutual interference among evaluation indicators. Developing a set of evaluation tools suitable for this kind of problem has considerable theoretical and practical significance. The TOPSIS method is an important MCDM tool. It is simple but comprehensive when applied to the evaluation of a MCDM problem. Also, the target weight is reflected in the integrated program (Boran, Genc, Kurt, & Akay, 2009; Deng, Yeh, & Willis, 2000). However, the traditional TOPSIS method, which is based on the Euclidean measure of distance to make decisions, takes the indicators as independent and do not perturb each other. This approach suffers information overlap and either overestimates or underestimates the indicators which take slack information. Considering the correlation between the competitiveness evaluation indices for the Chinese high-tech industry, we propose here an improved TOPSIS method. The method uses the concept of Mahalanobis distance to determine the distance to the ideal solution and the negative solution. The Mahalanobis distance is based on the sample covariance matrix and can solve the problem of the relevance among indicators as appropriate. This paper also provides proofs of the properties of the improved TOPSIS method and discusses its applicability through evaluation of results pertinent to the Chinese high-tech industry.

The remainder of this paper is organized as follows. Section 2 introduces the classic TOPSIS method. Section 3 uses the Mahalanobis distance to modify the traditional TOPSIS method according to the characteristics of the decision making process. It also derives the properties of the improved method. Section 4 applies the improved method to evaluate the competitiveness of the Chinese high-tech industry and analyze the evaluation effect according to the actual situation. Finally, conclusions and future work are given in the Section 5.

2. The traditional TOPSIS method

TOPSIS is an uncertain MCDM technology first proposed by Hwang and Yoon (1981). TOPSIS orders the criteria according to the distances from the object to the ideal and the negative solutions. The TOPSIS method can be summarized as follows.

Suppose there are m alternatives $A_1, A_2, \ldots, A_m$ and n decision criteria/attributes $C_1, C_2, \ldots, C_n$. Let x_{ij} denote the criteria/attribute value of A_i on C_j ($i = 1,2,\ldots,m; j = 1,2,\ldots,n$). All the values together form a decision matrix $\mathbf{X} = (x_{ij})_{m\times n}$. The decision matrix can be standardized in the form

$$R = (r_{ij})_{m\times n}, \tag{1}$$

where $r_{ij} = x_{ij}\left(\sum_{i=1}^{m} x_{ij}\right)^{-1/2}$.

The ideal solution $\mathbf{S}^+$ and the negative ideal solution $\mathbf{S}^-$ (also called the 'anti-ideal solution') are then determined:

$$\mathbf{S}^+ = \{s_1^+, s_2^+, \ldots, s_n^+\}, \quad \mathbf{S}^- = \{s_1^-, s_2^-, \ldots, s_n^-\}. \tag{2}$$

For the benefit index C_j:

$$s_j^+ = \max\{r_{ij} | 1 \leqslant i \leqslant m\}, \quad s_j^- = \min\{r_{ij} | 1 \leqslant i \leqslant m\};$$

and for the cost index C_j:

$$s_j^+ = \min\{r_{ij} | 1 \leqslant i \leqslant m\}, \quad s_j^- = \max\{r_{ij} | 1 \leqslant i \leqslant m\}.$$

We calculate the Euclidean distances of each alternative to the positive ideal and negative ideal solutions. The distance between alternative A_i and the positive ideal solution is:

$$d_i^+ = \sqrt{\sum_{j=1}^{n} (s_j^+ - r_{ij})^2}, \quad i = 1,2,\ldots,m. \tag{3}$$

The distance between alternative A_i and the negative ideal solution is:

$$d_i^- = \sqrt{\sum_{j=1}^{n} (s_j^- - r_{ij})^2}, \quad i = 1,2,\ldots,m. \tag{4}$$

Finally, we calculate the relative closeness of each alternative to the ideal solution:

$$c_i = \frac{d_j^-}{d_j^- + d_j^+}, \quad i = 1,2,\ldots,m. \tag{5}$$

The alternatives are ranked based on their relative closeness. A higher c_i value indicates that A_i is a better alternative, and *vice versa*.

TOPSIS simultaneously considers information about the positive and negative ideal solutions. Moreover, the calculation is simple, making TOPSIS widely used in multi-attribute evaluation. In applications, depending on the characteristics of the survey data, the TOPSIS method modified using fuzzy theory is capable of solving many practical evaluation issues. Examples include performance evaluation (Sun, 2010), customer evaluation (Chamodrakas, Alexopoulou, & Martakos, 2009), energy plans (Kaya & Kahraman, 2011), and business competitiveness evaluation (Torlak, Sevkli, Sanal, & Zaim, 2011). However, the TOPSIS method based on Euclidean distance does not consider correlation between indices, thus causing information overlap which, in turn, affects the decision results. Thus, the reduction of indicator correlation within the application often relies on qualitative analysis, the aim being to try to increase index independence during the process of index screening. These conditions make this approach strongly subjective.

3. Improved TOPSIS method using the Mahalanobis distance

To deal with the overlapping information problem, this paper modifies the traditional TOPSIS method by incorporating the Mahalanobis distance.

3.1. Definition of Mahalanobis distance

The Mahalanobis distance is a statistical distance that was first proposed by Mahalanobis (1936). This measure represents the covariance distance between variables and gauges the similarity of an unknown sample set to a known one. Thus, the Mahalanobis distance is calculated based on correlations between variables, by which different patterns can be identified and analyzed. Compared with Euclidean distance, the Mahalanobis distance considers the correlations of the data set and is scale-invariant. In other words, Mahalanobis distance is a multivariate-effect size.

For a multivariate vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^{\mathrm{T}}$, and covariance matrix Σ, the Mahalanobis distance is defined as

$$D_M(\mathbf{x}) = \sqrt{(\mathbf{x}-\mu)^{\mathrm{T}}\Sigma^{-1}(\mathbf{x}-\mu)}.$$

A Mahalanobis distance is also defined for the degree of dissimilarity between two random vectors $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ from the same distribution with the covariance matrix Σ, as:

$$d(\vec{\mathbf{x}}, \vec{\mathbf{y}}) = \sqrt{(\vec{\mathbf{x}}-\vec{\mathbf{y}})^{\mathrm{T}}\Sigma^{-1}(\vec{\mathbf{x}}-\vec{\mathbf{y}})}.$$

If the covariance matrix is the identity matrix, the Mahalanobis distance reduces to the Euclidean distance. If the covariance matrix is diagonal, the Mahalanobis distance can also be shown to be the normalized Euclidean distance.

3.2. TOPSIS improvement using the Mahalanobis distance

A Mahalanobis distance is essentially a weighted distance. The weight is based on the variance of the indicators and the correlation degree with other indicators. The Mahalanobis distance standardizes data via the factor Σ^{-1}. Accordingly, the Mahalanobis distance not only considers the correlation between observations, but also eliminates the effect of the different dimensions of each index.

Suppose the n dimensional vector $\mathbf{r}_i = (r_{i1}, r_{i2}, \ldots, r_{in})$ is an indicator vector of A_i under the index set $C = \{C_1, C_2, \ldots, C_n\}$ (so that r_i is the data vector of A_i). Let $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ be the weight vector, where ω_j is the weight of C_j that meets constraints $\omega_j \in [0,1]$ and $\sum_{i=1}^{n}\omega_j = 1$. The ideal solution point $\mathbf{S}^{+}$ and the negative solution point $\mathbf{S}^{-}$ both come from the n dimensional population with mean $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^{\mathrm{T}}$ and covariance Σ, so that the Mahalanobis distance from A_i to the ideal solution point is:

$$d(\mathbf{r}_i, \mathbf{S}^{+}) = \sqrt{\{s_j^{+} - r_{ij}\}^{\mathrm{T}}\Omega^{\mathrm{T}}\Sigma^{-1}\Omega\{s_j^{+} - r_{ij}\}}, \quad i = 1, 2, \ldots, m. \tag{6}$$

Similarly, the Mahalanobis distance from A_i to the negative ideal solution point is:

$$d(\mathbf{r}_i, \mathbf{S}^{-}) = \sqrt{\{s_j^{-} - r_{ij}\}^{\mathrm{T}}\Omega^{\mathrm{T}}\Sigma^{-1}\Omega\{s_j^{-} - r_{ij}\}}, \quad i = 1, 2, \ldots, m. \tag{7}$$

In these expressions we have used

$$\Omega = \mathrm{diag}(\sqrt{\omega_1}, \sqrt{\omega_2}, \ldots, \sqrt{\omega_n}). \tag{8}$$

The closeness of each alternative is given by

$$c_i = \frac{d(\mathbf{r}_i, \mathbf{S}^{-})}{d(\mathbf{r}_i, \mathbf{S}^{-}) + d(\mathbf{r}_i, \mathbf{S}^{+})}, \quad i = 1, 2, \ldots, m. \tag{9}$$

We sort the results according to the value of c_i. A higher c_i suggests that A_i is a better solution. In practical application, an unknown population distribution can be replaced by a sample covariance matrix.

3.3. Properties of the improved TOPSIS method

The introduction of the Mahalanobis distance as a means to improve the traditional TOPSIS method improves the properties of the technique. Firstly, the improved TOPSIS method can overcome correlation disturbance during evaluation. Moreover, the evaluation results are not affected by the indicators' dimensions.

Property 1. The relative closeness c_i in the improved-TOPSIS method is invariant to non-singular linear transformation.

Proof. Suppose $r_i = (r_{i1}, r_{i2}, \ldots, r_{in})^{\mathrm{T}}$, $\tilde{\mathbf{r}}_i = (a_1 + b_1 r_{i1}, a_2 + b_2 r_{i2}, \ldots, a_n + b_n r_{in})^{\mathrm{T}}$, $S^{+} = (s_1^{+}, s_2^{+}, \ldots, s_n^{+})^{\mathrm{T}}$, and $\tilde{S}^{+} = (a_1 + b_1 s_1^{+}, a_2 + b_2 s_2^{+}, \ldots, a_n + b_n s_n^{+})^{\mathrm{T}}$, where the a_i and b_i are constants and $b_i \neq 0$. Let $\mathbf{A} = (a_1, a_2, \ldots, a_m)^{\mathrm{T}}$ and $\mathbf{B} = \mathrm{diag}(b_1, b_2, \ldots, b_m)$. Then $\tilde{\mathbf{r}}_i = \mathbf{A} + \mathbf{B}\mathbf{r}_i$ and $\tilde{\mathbf{S}}^{+} = \mathbf{A} + \mathbf{B}\mathbf{S}^{+}$. Given that $\tilde{\boldsymbol{\Sigma}} = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^{\mathrm{T}}$, we have $\tilde{\boldsymbol{\Sigma}}^{-1} = (\mathbf{B}^{-1})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{B}^{-1}$, so that

$$\begin{aligned} d(\tilde{\mathbf{r}}_i, \tilde{\mathbf{S}}^{+}) &= \sqrt{(\tilde{\mathbf{r}}_i - \tilde{\mathbf{S}}^{+})^{\mathrm{T}}\Omega^{\mathrm{T}}\tilde{\Sigma}^{-1}\Omega(\tilde{\mathbf{r}}_i - \tilde{\mathbf{S}}^{+})} \\ &= \sqrt{(\mathbf{A} + \mathbf{B}\mathbf{r}_i - \mathbf{A} - \mathbf{B}\mathbf{S}^{+})^{\mathrm{T}}\Omega^{\mathrm{T}}(\mathbf{B}^{-1})^{\mathrm{T}}\Sigma^{-1}\mathbf{B}^{-1}\Omega(\mathbf{A} + \mathbf{B}\mathbf{r}_i - \mathbf{A} - \mathbf{B}\mathbf{S}^{+})} \\ &= \sqrt{(\mathbf{r}_i - \mathbf{S}^{+})^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\Omega^{\mathrm{T}}(\mathbf{B}^{-1})^{\mathrm{T}}\Sigma^{-1}\mathbf{B}^{-1}\Omega\mathbf{B}(\mathbf{r}_i - \mathbf{S}^{+})} \\ &= \sqrt{(\mathbf{r}_i - \mathbf{S}^{+})^{\mathrm{T}}\Omega^{\mathrm{T}}\Sigma^{-1}\Omega(\mathbf{r}_i - \mathbf{S}^{+})} = d(\mathbf{r}_i, \mathbf{S}^{+}). \end{aligned}$$

For the same reason, $d(\tilde{\mathbf{r}}_i, \tilde{S}^{-}) = d(\mathbf{r}_i, S^{-})$. □

Thus, the closeness after non-singular linear transformation is:

$$\tilde{c}_i = \frac{d(\tilde{\mathbf{r}}_i, \tilde{\mathbf{S}}^{-})}{d(\tilde{\mathbf{r}}_i, \tilde{\mathbf{S}}^{-}) + d(\tilde{\mathbf{r}}_i, \tilde{\mathbf{S}}^{+})} = \frac{d(\mathbf{r}_i, \mathbf{S}^{-})}{d(\mathbf{r}_i, \mathbf{S}^{-}) + d(\mathbf{r}_i, \mathbf{S}^{+})} = c_i.$$

Property 1. Shows that if the standardization of the primary data is a kind of non-singular linear transformation in the decision-making process, then the standardization process will not affect the decision result.

Property 2. When the evaluation indicators $C_1, C_2, \ldots, C_n$ are unrelated, then:

$$d(\mathbf{r}_i, \mathbf{S}^{+}) = \sqrt{\sum_{j=1}^{n}\frac{\omega_j(r_{ij} - s_j^{+})^2}{\sigma_j^2}} \quad \text{and} \quad d(\mathbf{r}_i, \mathbf{S}^{-}) = \sqrt{\sum_{j=1}^{n}\frac{\omega_j(r_{ij} - s_j^{-})^2}{\sigma_j^2}}$$

Proof. We are given that both $\mathbf{r}_i$ and $\mathbf{S}^+$ come from an n dimension population with mean $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^{\mathrm{T}}$ and covariance Σ. The n decision indicators are unrelated to one another, so that $\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$. Then, $\Sigma^{-1} = \mathrm{diag}\left(\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \ldots, \frac{1}{\sigma_m^2}\right)$, and

$$d^2(\mathbf{r}_i, \mathbf{S}^+) = \{r_{ij} - s_j^+\}^{\mathrm{T}} \Omega^{\mathrm{T}} \Sigma^{-1} \Omega \{r_{ij} - s_j^+\}$$
$$= \left(\sqrt{\omega_1}(r_{i1} - s_1^+), \ldots, \sqrt{\omega_n}(r_{in} - s_n^+)\right) \times \begin{bmatrix} \frac{1}{\sigma_1^2} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n^2} \end{bmatrix} \begin{pmatrix} \sqrt{\omega_1}(r_{i1} - s_1^+) \\ \vdots \\ \sqrt{\omega_m}(r_{in} - s_n^+) \end{pmatrix} = \sum_{j=1}^{n} \frac{\omega_j (r_{ij} - s_j^+)^2}{\sigma_j^2}.$$

□

Thus, $d(\mathbf{r}_i, \mathbf{S}^+) = \sqrt{\sum_{j=1}^{n} \frac{\omega_j (r_{ij} - s_j^+)^2}{\sigma_j^2}}$. Similarly,

$$d(\mathbf{r}_i, \mathbf{S}^-) = \sqrt{\sum_{j=1}^{n} \frac{\omega_j (r_{ij} - s_j^-)^2}{\sigma_j^2}}.$$

Thus, when the evaluation indices are not related to one another, the weighted Mahalanobis distance and the weighted Euclidean distance are equivalent. The Mahalanobis distance considers index differences, the dimensions, as well as the correlation among indicators. Thus, the Mahalanobis distance can prevent the overlapping of information, making it more suitable for dealing with complex practical problems. In actual applications, the overall covariance matrix Σ is often unknown and can thus be replaced by the sample covariance matrix S.

4. Evaluation of the competitiveness of provincial high-tech industry

The improved TOPSIS method is suitable for considering the correlation among indicators and is consistent with the features of high-tech industry competitiveness evaluation. Thus, this section uses the improved TOPSIS method to evaluate the competitiveness of Chinese provincial high-tech industries.

4.1. Evaluation index system for provincial-level high-tech industrial competitiveness

Based on the principles of science and coordination, comprehensiveness and integrity, operability and ease, norms and comparability, the evaluation system for the competitiveness of high-tech industry is first established. This is based on the statistical indicators from the National Bureau of Statistics and existing study results from China and elsewhere. The system includes six aspects and a total of 17 secondary indicators (see Table 1). These aspects are: human capital input level, material resources input level, funding input level, industrial output level, industrial innovation level, and efficiency level. The weights of the index system are based on the majority of the research using AHP.

We first analyze the input indicators. Human capital input level (A) reflects the strength of the industrial development scale and refers to the potential power of high-tech industrial development. Material resource level (B) represents the degree of support by the government, institutions, and enterprises on industrial development as well as reflecting the basis for the development of high-tech industries. Funding input level (C) epitomizes the level of capital resources and financing capability. We then analyze the output indicators. Industrial output level (D) measures the operating level and performance of high-tech businesses. Industrial innovation level (E) is the source and intrinsic motivation for the sustainable development of high-tech enterprises. Efficiency level (F) reflects the update speed of the industry and the capability for equipment renewal.

Table 1
The evaluation system for the competitiveness of high-tech industry.

First order index	Second order index	Weight
Human capital input level A	Number of employees	0.0757
	R&D activities equivalent to full-time equivalent	0.1135
Material resource input level B	New fixed assets	0.0430
	The number of high-tech enterprises	0.0553
	All completed or put into the project	0.0369
Funding input level C	R&D expenditures	0.0618
	New product development expenditures	0.0541
	Investment	0.0463
Industrial output level D	Total output value with current prices	0.0437
	Main business income	0.0509
	Export delivery value	0.0437
	Profits and taxes	0.0509
Industrial innovation level E	The output value of new products	0.0473
	New product sales	0.0541
	Number of patented inventions	0.0608
Efficiencylevel F	Application rate of fixed assets	0.0757
	The project completed and put into production rate	0.0865

4.2. Data and correlation testing

This paper gathers data on the 17 indicators pertinent to the Chinese high-tech industry in 31 provinces, autonomous regions, and municipalities. It is taken directly from *The hi-tech industry in China Statistical Yearbook 2011* which was sponsored by the National Bureau of Statistics (see Appendix A). To verify the effectiveness of the newly proposed TOPSIS method, this paper first tests the correlation of the evaluation index using Pearson analysis. The results are shown in Table 2.

Table 2 shows that: (1) the categories of the input indicators are significantly correlated to one another at a confidence level of more than 99%; (2) the correlation of the input and output indicators is also high, except for the classification of efficiency indicators (f1 and f2); and (3) the correlation among the output indicators is also high but they are not significantly correlated with the efficiency indicator. Based on such high correlations among indicators, we have to take a skeptical attitude toward the validity and reliability of the evaluation result obtained using the traditional TOPSIS method.

4.3. Evaluation of the competitiveness of provincial high-tech industry using the improved TOPSIS method

We evaluate high-tech industry competitiveness using the improved TOPSIS method. First, we determine the ideal and negative ideal solution points for each index point. Secondly, using Eqs. (6)–(8), we calculate the Mahalanobis distance between A_i and $\mathbf{S}^+$ as well as between A_i and $\mathbf{S}^-$ (Table 3, second and third columns). Thirdly, the closeness degree is calculated according to Eq. (9) (Table 3, fourth column). For comparison, the results from the improved and traditional TOPSIS methods are shown in Table 3.

The sorting results as calculated by the two different methods show significant differences (Table 3, last column). By considering the correlation among the provincial indicators, the differences are quite evident. This is especially so in provinces that are influenced

Table 2
The results of correlation test.

	a1	a2	b1	b2	b3	c1	c2	c3	d1	d2	d3	d4	e1	e2	e3	f1	f2
a1	1	0.980**	0.616**	0.943**	0.490**	0.970**	0.978**	0.651**	0.981**	0.979**	0.972**	0.984**	0.961**	0.960**	0.887**	0.083	0.070
a2	0.980**	1	0.485**	0.895**	0.385*	0.992**	0.955**	0.517**	0.941**	0.940**	0.942**	0.944**	0.974**	0.971**	0.956**	0.086	0.025
b1	0.616**	0.485**	1	0.659**	0.802**	0.453*	0.613**	0.941**	0.651**	0.647**	0.587**	0.676**	0.444*	0.448*	0.234	0.395*	0.353
b2	0.943**	0.895**	0.659**	1	0.555**	0.890**	0.946**	0.714**	0.944**	0.943**	0.911**	0.965**	0.887**	0.888**	0.759**	0.074	0.050
b3	0.490**	0.385*	0.802**	0.555**	1	0.348	0.442*	0.831**	0.498**	0.492**	0.429*	0.537**	0.327	0.332	0.201	0.189	0.575**
c1	0.970**	0.992**	0.453*	0.890**	0.348	1	0.962**	0.502**	0.945**	0.945**	0.944**	0.943**	0.985**	0.985**	0.959**	0.059	−0.013
c2	0.978**	0.955**	0.613**	0.946**	0.442*	0.962**	1	0.652**	0.991**	0.991**	0.979**	0.983**	0.962**	0.964**	0.851**	0.103	0.008
c3	0.651**	0.517**	0.941**	0.714**	0.831**	0.502**	0.652**	1	0.700**	0.698**	0.636**	0.722**	0.500**	0.507**	0.280	0.144	0.327
d1	0.981**	0.941**	0.651**	0.944**	0.498**	0.945**	0.991**	0.700**	1	1.000**	0.991**	0.986**	0.948**	0.952**	0.825**	0.051	0.039
d2	0.979**	0.940**	0.647**	0.943**	0.492**	0.945**	0.991**	0.698**	1.000**	1	0.991**	0.985**	0.950**	0.954**	0.825**	0.048	0.034
d3	0.972**	0.942**	0.587**	0.911**	0.429*	0.944**	0.979**	0.636**	0.991**	0.991**	1	0.959**	0.948**	0.951**	0.841**	0.028	0.027
d4	0.984**	0.944**	0.676**	0.965**	0.537**	0.943**	0.983**	0.722**	0.986**	0.985**	0.959**	1	0.943**	0.945**	0.825**	0.083	0.060
e1	0.961**	0.974**	0.444*	0.887**	0.327	0.985**	0.962**	0.500**	0.948**	0.950**	0.948**	0.943**	1	0.999**	0.938**	0.021	−0.039
e2	0.960**	0.971**	0.448*	0.888**	0.332	0.985**	0.964**	0.507**	0.952**	0.954**	0.951**	0.945**	0.999**	1	0.934**	0.017	−0.038
e3	0.887**	0.956**	0.234	0.759**	0.201	0.959**	0.851**	0.280	0.825**	0.825**	0.841**	0.825**	0.938**	0.934**	1	0.028	−0.008
f1	0.083	0.086	0.395*	0.074	0.189	0.059	0.103	0.144	0.051	0.048	0.028	0.083	0.021	0.017	0.028	1	0.393*
f2	0.070	0.025	0.353	0.050	0.575**	−0.013	0.008	0.327	0.039	0.034	0.027	0.060	−0.039	−0.038	−0.008	0.393*	1

* Significant at the 00.05 level (bilateral).
** Significant at the 0.01 level (bilateral).

Table 3
Comparison of the results obtained using the traditional and improved TOPSIS methods.

			Improved TOPSIS		Traditional TOPSIS		
Province	d+	d−	Closeness	Order	Closeness	Order	Differ
Beijing	12.523	2.332	0.157	4	0.2172	3	1
Tianjin	12.781	1.587	0.110	9	0.1358	7	2
Hebei	12.648	1.602	0.112	8	0.0145	18	10
Shanxi	13.386	0.808	0.057	26	0.0093	21	5
Neimenggu	13.128	1.031	0.073	19	0.0017	27	8
Liaoning	13.128	1.351	0.093	14	0.0397	10	4
Jilin	13.228	1.011	0.071	20	0.0071	22	2
Heilongjiang	13.029	1.333	0.093	16	0.0118	20	4
Shanghai	12.013	2.591	0.177	3	0.1865	4	1
Jiangsu	9.962	3.967	0.285	2	0.4285	2	0
Zhejiang	12.292	2.045	0.143	6	0.1224	8	2
Anhui	12.810	1.269	0.090	18	0.0164	17	1
Fujian	12.876	1.353	0.095	12	0.1388	6	6
Jiangxi	12.997	1.335	0.093	15	0.0229	16	1
Shandong	12.468	2.285	0.155	5	0.1743	5	0
Henan	12.786	1.471	0.103	11	0.0232	15	4
Hubei	12.678	1.485	0.105	10	0.0490	9	1
Hunan	12.866	1.284	0.091	17	0.0256	13	4
Guangdong	3.703	4.783	0.564	1	0.9977	1	0
Guangxi	13.247	0.934	0.066	22	0.0059	23	1
Hainan	13.189	0.282	0.021	31	0.0010	28	3
Chongqing	13.022	0.959	0.069	21	0.0252	14	7
Sichuang	12.698	1.642	0.114	7	0.0345	11	4
Guizhou	13.173	0.469	0.034	30	0.0129	19	11
Yunnan	13.109	0.739	0.053	27	0.0045	24	3
Xizang	13.185	0.879	0.062	23	0.0003	31	8
Shanxi	12.772	1.330	0.094	13	0.0328	12	1
Gansu	13.225	0.655	0.047	28	0.0034	25	3
Qinghai	13.185	0.868	0.062	24	0.0004	30	6
Ningxia	13.190	0.578	0.042	29	0.0030	26	3
Xinjiang	13.198	0.853	0.061	25	0.0007	29	4

by their neighboring provinces or have an unbalanced development relative to their surrounding provinces, such as Hebei Province. Hebei has moderate high-tech base conditions but is located in the hinterland of Beijing and Tianjin, which are well developed high-tech centers. So the results reveal the sharp contrast in the high-tech development between Hebei and the other two municipalities (Beijing and Tianjin). In practice, high-tech resources outflow from Hebei Province to Beijing and Tianjin. The traditional TOPSIS method does not consider the interaction between provincial indicators, which seriously underestimates the high-tech competitiveness of Hebei. However, Guizhou Province shows the opposite deviation. Guizhou is encircled by many neighboring provinces, including Sichuan, Chongqing, Yunnan, Guangxi, and Hunan, which all have their own development characteristics and are superior to Guizhou in terms of high-tech development. The high-tech base of Guizhou is weaker than that of its surrounding provinces. Thus, considering the links with surrounding high-tech provinces, resources that are not representative of the high-tech competitiveness of Guizhou flow into the province. Therefore, the evaluation result obtained using the traditional TOPSIS method is inflated because this correlation was ignored. The sorting differences relative to the other provinces can also be explained in a similar manner. The improved TOPSIS method, which considers the correlation among indicators, is

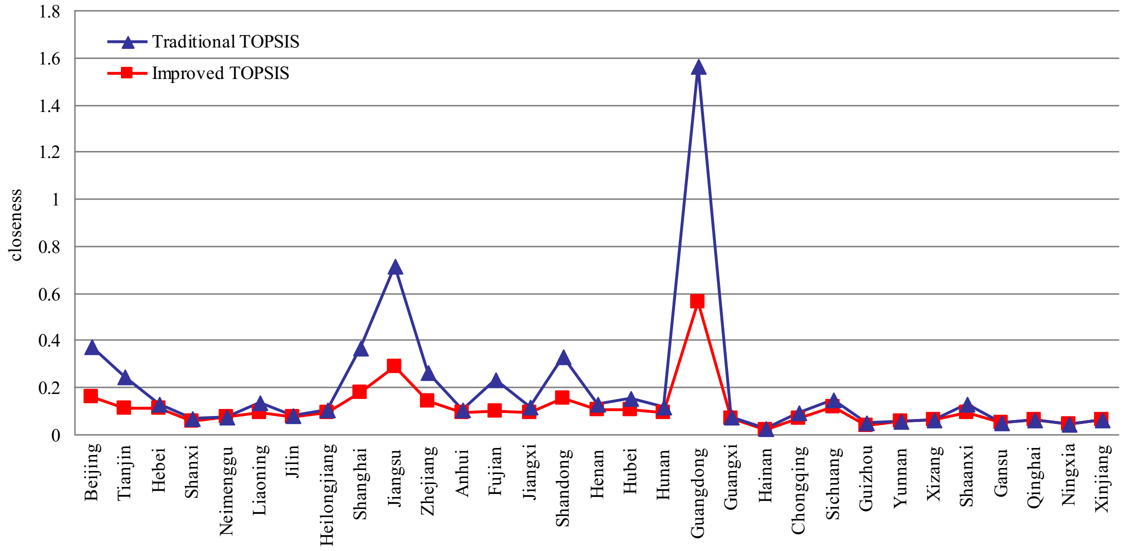

Fig. 1. Comparison of the closeness obtained using the traditional and improved TOPSIS methods.

capable of truly reflecting the characteristics of the provincial high-tech linkage to provide a firm basis for scientific decision making. A scatter plot of the closeness obtained using the two methods is given in Fig. 1.

Fig. 1 shows that in terms of the absolute value of closeness, the traditional TOPSIS method cannot effectively deal with information overlap. This overlap amplifies the closeness among provinces that have better high-tech development conditions and are surrounded by other developed provinces, such as Guangdong, Jiangsu, Beijing, Shanghai, and so on. This condition also induces a high fluctuation in the closeness. The improved TOPSIS method avoids calculating overlapping information, restores a reasonable fluctuation range, softens the closeness, and therefore reflects the real situation between independent indicators.

Based on the overall evaluation effect, the results obtained using the improved TOPSIS method are consistent with the outcome of Chen and Sun (2011), who used factor analysis. The result is also unanimous with the cluster analysis performed by Chen (2010). Therefore, this improved TOPSIS method, which resolves indicator correlation, has good applicability in high-tech industry evaluation and is of practical significance in scientific evaluation and decision making.

4.4. Practical implications

The modified method is more suitable for practical situations than the traditional TOPSIS method which suggests some profound implications for management practices:

(1). As seen in Fig. 1, high-tech development is unbalanced in China and there is also a gap between regions in the east and west. Guizhou, Guangxi, Xinjiang, and Ningxia are obviously weaker than the eastern provinces in terms of high-tech competitiveness. This conclusion is consistent with existing research (Sun et al., 2010; Wang, 2007, etc.). The imbalance in regional high-tech competitiveness will affect the *entire* national innovation capacity, and requires a higher level of resource allocation.

(2). From the evaluation results, we can see that some provinces are relatively high in high-tech competitiveness (such as Beijing, Tianjin, Shanghai, Jiangsu, and Guangdong). This will impact upon and motivate the neighboring provinces. The correlation among indicators is a real issue in high-tech industry competitiveness evaluation (Table 2 has been verified). Therefore, strengthening high-tech industry integration in regional economic circles, to achieve internal linking of economic circles, plays an important role in the promotion of national high-tech industrial development.

(3). The evaluation index system implies that human capital is the most important part of high-tech competition. The existing systems (Wu & Li, 2008; Zheng et al., 2010, etc.) also emphasize talent strategy in high-tech industry. Human capital is the 'soft power' to promote technological innovation, and it is the core driving force in high-tech development. Therefore, enhancing talent building in high-tech industry plays a foundational role in improving competitiveness.

5. Conclusions and future work

High-tech industry is a comprehensive national strength. Evaluation of these industries is a management issue concerning industrial layout but it is also an MCDM technical issue. The conflict and correlation among evaluation indices are common features faced by all evaluation methods. In order to avoid correlation among indicators, this paper proposes an improved TOPSIS method based on the concept of the Mahalanobis distance and the method was applied to evaluate competitiveness in Chinese high-tech industry. The improved TOPSIS method has been shown to satisfy two properties: (1) the relative closeness c_i is invariant to non-singular linear transformation; and (2) the weighted Mahalanobis distance is equal to the weighted Euclidean distance when there is no correlation among indicators. From the evaluation result, we can see that the difference between the two methods is evident through the influence that a province experiences due to surrounding provinces. Based on the closeness value, the improved TOPSIS method softens the closeness and restores the real fluctuation range in the closeness after considering the correlation. The results obtained using the improved TOPSIS method are consistent with the conclusions found for similar problems solved using different methods, and it is also consistent with practical situations.

Appendix A. Primary data

	A1	A2	B1	B2	B3	C1	C2	C3	D1	D2	D3	D4	E1	E2	E3	F1	F2
Beijing	249,889	8440	37	1103	24	368,388	503,264	137	2993	3334	1217	265	13,659,447	13,607,777	2136	27	23
Tianjin	240,022	6750	193	817	66	220,547	286,850	216	2242	2291	1115	187	8,539,382	8,481,937	872	89	50
Hebei	171,439	6632	134	438	198	90,836	84,196	214	843	883	158	123	694,240	701,062	285	63	58
Shanxi	12,0945	1098	17	157	50	13,882	14,488	48	249	234	54	20	491,332	478,459	79	35	53
Neimenggu	27,846	227	23	107	48	3672	3962	54	235	225	7	40	68,013	85,620	8	43	67
Liaoning	218,709	4047	175	987	157	258,646	187,839	440	1712	1710	516	183	2,326,777	2,185,478	271	40	49
Jilin	101,147	1644	208	436	757	19,327	20,702	275	727	642	7	80	335,604	317,110	105	76	89
Heilongjiang	72,422	4924	73	199	103	158,880	158,131	104	352	399	14	70	347,638	292,694	183	70	62
Shanghai	531,834	19,278	85	1423	78	673,565	851,462	232	6901	7020	4987	308	10,779,647	11,761,955	2509	37	37
Jiangsu	2,267,628	64,496	971	4868	893	1,351,327	1,839,490	1324	16,278	16,170	9726	1286	25,657,113	25,619,005	3604	73	65
Zhejiang	646,326	24,485	105	3339	241	524,402	580,192	153	3413	3324	1290	404	7,409,716	6,956,994	2199	69	41
Anhui	146,412	6693	147	745	372	122,141	151,688	415	682	662	45	95	688,313	841,827	347	35	53
Fujian	321,249	14,034	88	791	84	373,649	401,274	161	2621	2577	1494	242	9,075,524	8,074,863	624	55	28
Jiangxi	218,106	5418	356	555	517	104,371	123,956	453	1038	1039	147	102	1,275,836	1,175,791	203	79	70
Shandong	545,398	15,618	261	1847	384	612,385	685,158	521	5176	5149	1565	555	10,021,091	11,133,363	1268	50	50
Henan	244,892	7262	187	728	381	98,992	129,567	326	1227	1186	60	169	1,087,199	1,323,861	328	57	57
Hubei	214,977	10,461	161	798	296	198,633	285,418	266	1312	1257	378	192	2,962,179	2,781,652	1201	60	60
Hunan	157,767	4964	124	683	273	95,857	79,681	227	931	906	35	142	1,541,567	1,486,141	490	55	46
Guangdong	3,547,488	156,235	303	5774	374	3,630,850	2,710,412	496	21,050	20,953	13,479	1684	61,567,664	60,464,340	31,356	61	49
Guangxi	110,210	1115	54	338	177	15,746	20,935	93	432	384	95	69	225,311	199,155	195	58	59
Hainan	12,634	392	1	57	1	9668	14,639	8	86	77	2	20	17,900	15,401	31	15	8
Chongqing	88,616	4000	46	324	105	64,451	57,767	195	532	508	57	46	1,632,013	1,469,500	320	24	53
Sichuang	325,736	11,640	379	830	216	247,534	372,848	334	2154	2105	453	258	1,643,519	1,385,860	417	113	53
Guizhou	66,968	4932	6	150	16	98,334	134,270	20	323	266	10	39	684,536	581,453	399	30	29
Yunnan	26,672	1002	10	144	29	17,283	13,069	31	169	160	7	33	276,181	261,031	221	32	30
Xizang	1471	10	1	11	11	648		2	6	5		3				73	61
Shaanxi	198,975	12,006	286	381	99	261,870	326,309	164	858	865	69	109	1,609,766	1,586,068	466	175	45
Gansu	27,545	727	11	81	35	30,314	16,562	22	81	76	3	15	177,660	170,318	29	49	36
Qinghai	5145	22	4	28	7	722	710	5	23	21	0	2	257	215		93	58
Ningxia	6708	408	1	16	4	7680	6555	4	36	31	9	7	198,738	178,064	11	23	33
Xinjiang	7076	114	3	34	15	3701	7992	4	29	26	4	5	31,952	30,638	9	70	58

A1: number of employees; A2: R&D activities equivalent to full-time equivalent.
B1: New fixed assets (100 million); B2: the number of high-tech enterprises; B3: all completed or put into the project.
C1: R&D expenditures (10 thousand); C2: new product development expenditures (10 thousand); C3: investment (100 million).
D1: total output value with current prices (100 million); D2: main business income (100 million); D3: export delivery value (100 million); D4: Profits and taxes (100 million).
E1: the output value of new products (10 thousand); E2: new product sales (10 thousand); E3: number of patented inventions.
F1: application rate of fixed assets; F2: project completion and put into production rate.

This study makes the following important contributions: (1) from a methodological perspective, the improved TOPSIS method provides a more accurate tool for solving MCDM problems. The correlation among indicators is a widely recognized characteristic in many MCDM issues. The existing evaluation methods, which concern the relationships between indicator weights or focuses on the application of a single index of information, find it difficult to take both relevance and integrity into consideration (Li & Zang, 2012). This paper considers the correlation among indicators under the general framework of the TOPSIS method. On the one hand, this avoids mutual interference between indicators, and, on the other, it obtains the evaluation results from the overall layout. It also takes on the role of evaluating the value judgments to provide a scientific basis for decision-making. (2) From a practical point of view, Chinese high-tech industrial competitiveness is an important part of the national strategy. Evaluation results which take into account the correlation among indicators are strongly supported by practical evidence. Therefore, the improved TOPSIS method will improve the accuracy of the evaluation and better reflect the actual situation.

In this paper, an improved TOPSIS method based on the weighted Mahalanobis distance is proposed which focuses on eliminating the linear correlation among indicators. This, to some extent, improves the accuracy of the evaluation results. However, compared to linear correlation, nonlinear correlation is even more pervasive in actual systems. Obviously, our improved TOPSIS method cannot effectively solve nonlinear-related issues. Therefore, in future work, the problem of how to eliminate nonlinear correlations in the relationships among the indicators needs to be taken into account. However, as multi-attribute evaluations and associated decision making problems are widely encountered in real life situations (and correlation among the different attributes is also a very common phenomenon), future work may also focus on using this improved TOPSIS method for other important systems. Potential applications include management information systems, financial systems, macroeconomic systems, etc., which will all provide good tests of the validity of the new method.

Acknowledgements

The authors are grateful to the editors and the anonymous reviewers for their insightful comments and suggestions. This research is supported by the National Natural Science Foundation of China (Grant No. 71101132) and the Soft Science Research Foundation of Zhejiang Province, China (Grant No. 2012C35014).

References

Boran, F. E., Genc, S., Kurt, M., & Akay, D. (2009). A multi-criteria intuitionistic fuzzy group decision making for supplier selection with TOPSIS method. *Expert Systems with Applications, 36*(8), 11363–11368.

Chamodrakas, I., Alexopoulou, N., & Martakos, D. (2009). Customer evaluation for order acceptance using a novel class of fuzzy method based on TOPSIS. *Expert Systems with Applications, 36*(4), 7409–7415.

Chen, C. H. (2010). An empirical study of competitiveness evaluation in high-tech industry. *Soft Science, 24*(8), 21–29.

Chen, L., & Sun, H. M. (2011). The competitiveness evaluation of Chinese provincial high-tech industry. *Science and Technology Management Research, 5*(2), 76–79.

Deng, H. P., Yeh, C. H., & Willis, R. J. (2000). Inter-company comparison using modified TOPSIS with object weights. *Computers & Operations Research, 27*(10), 963–973.

Hwang, C. L., & Yoon, K. (1981). *Multiple attribute decision making-methods and applications*. Berlin: Heidelberg, Springer-Verlag.

Kaya, T., & Kahraman, C. (2011). Multicriteria decision making in energy planning using a modified fuzzy TOPSIS methodology. *Expert Systems with Applications, 38*(6), 6577–6585.

Li, X. W., & Zang, S. W. (2012). Study on the regional development countermeasures of Chinese high-tech industries. *Science & Technology Progress and Policy, 29*(24), 1–5.

Liang, B. M. (2011). Evaluation of high-tech enterprises. *Canadian Social Science, 7*(6), 245–248.

Liang, X. Y., Li, Z. G., Tang, S. K., & Zhao, L. J. (2007). A study on the spatial distribution of Chinese Hi-tech industries––spatial econometrics analysis based on province-level industrial output value. *Studies in Science of Science, 25*(3), 453–460.

Lu, Y. D., & Yu, F. (2010). The evaluation of the innovation capability of China's high-tech industries. *International Business Research, 3*(2), 87–91.

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India, 2*(1), 49–55.

Sun, C. C. (2010). A performance evaluation model by integrating fuzzy AHP and Fuzzy TOPSIS methods. *Expert Systems with Applications, 37*(12), 7745–7754.

Sun, Y., Xiong, W., & Zheng, W. (2010). Research on the policies of High-Tech industry spatial transferring in China. Science of Science and Management of S.&T. 31(11), 163–168.

Torlak, G., Sevkli, M., Sanal, M., & Zaim, S. (2011). Analyzing business competition by using fuzzy TOPSIS method: An example of Turkish domestic airline industry. *Expert Systems with Applications, 38*(4), 3396–3406.

Wang, X. J. (2007). The factor analysis of comprehensive evaluation on the development level of high-tech industry in different provinces in China. *Mathematics in Practice and Theory, 37*(8), 17–28.

Wang, J. G., & Yu, Y. C. (2004). A compare research on Chinese high-tech industry competition. *Industrial Technology & Economy, 23*(1), 92–95.

Wu, M., & Li, D. (2008). The core competitiveness analysis of high-tech industry based on TOPSIS. *Soft Science, 22*(8), 105–108.

Zheng, Z. Y., Shi, S. X., & He, S. W. (2010). Competiveness evaluation study of high-tech industry in Fujian Province. *Southeast Academic Research, 32*(5), 91–97.