African Journal of Business Management Vol.6(2), pp. 521-529,18 January, 2012
Available online at http://www.academicjournals.org/AJBM
DOI: 10.5897/AJBM11.1520
ISSN 1993-8233 ©2012 Academic Journals

Full Length Research Paper

Fuzzy analytic hierarchy process (FAHP) and balanced scorecard approach for evaluating performance of Third-Party Logistics (TPL) enterprises in Chinese context

Lin Wang[1], Hao Zhang[1] and Yu-Rong Zeng[2]*

[1]School of Management, Huazhong University of Science and Technology, Wuhan 430074, China.
[2]School of Information Management, Hubei University of Economics, Wuhan 430205, China.

Accepted 26 September, 2011

Enterprise performance evaluation is a good way to effectively and continuously improve and enhance the efficiency of logistics enterprises. To meet the mounting requirements of reliable, speedy, and flexible delivery to global customers, China's Third-Party Logistics (TPL) providers have to establish a reasonable performance evaluation system. The objective of this study is to construct an approach based on the fuzzy analytic hierarchy process (FAHP) and balanced scorecard (BSC) for evaluating a TPL enterprise in China. The BSC concept is applied to define the hierarchy with four major perspectives (that is, financial, customer, internal business process, and learning and growth), and performance criteria are selected for each perspective. Then, a FAHP approach is proposed in order to tolerate vagueness and ambiguity of information. At last, a numerical example shows the proposed model can be a useful and effective assessment tool for solving this kind of multiple-criteria decision-making problems.

Key words: Third-party logistics, performance evaluation, balanced scorecard, fuzzy analytic hierarchy process.

INTRODUCTION

Logistics industry is experiencing rapid development and plays an increasingly important role in China's global economic development. In 2010, China's total logistics expenditures were Renminbi (RMB) 7.1 trillion and had an average annual growth of 16.7%, which represented 17.8% of China's gross domestic product. In 2010, logistics contributed about RMB 2.7 trillion of China's value added (16.0% from the servicing industry) and had an annual increase of 13.1% (National Development and Reform Commission, National Bureau of Statistics and China Federation of Logistics and Purchasing, 2011). As a developing country, China's logistics costs are nearly double that of Western countries (Tian et al., 2008). At present, third-party logistics (TPL) is becoming a source for companies to gain competitive advantages by reducing logistics cost, enhancing core competency, and improving service quality (Susanne and Monica, 2003). Therefore, more and more companies continue to out-source their logistics activities to TPL providers to reduce their logistics costs. Although, the entry into the WTO in 2001 and opening up the logistics markets at the end of 2005 allowed its logistics industry to grow at an even faster rate in China, the added competitive challenges from global players and increased uncertainties of tran-sition economy have led to a more complex competitive logistics market (Wang et al., 2010). Studies have shown that enterprise performance evaluation is a good way to effectively and continuously improve and enhance the efficiency of logistics enterprises. To meet the mounting requirements of reliable, speedy and flexible delivery to

*Corresponding author. E-mail: zyrhbue@gmail.com.

global customers, China's TPL providers have to establish the reasonable evaluation criteria and an effective model to assess their performance. The metrics of measuring the performance on logistics service of TPL firms have an important role to play in setting objectives, evaluating performance, and determining future plans of actions. Unfortunately, performance measurement and metrics pertaining to TPL companies in China have not received adequate attention from researchers or practitioners. Some methods have been suggested over the years to evaluate the performance of TPL firms. However, well-known financial measures such as the return on investment (ROI), internal rate of return (IRR), net present value (NPV) and payback period have been demonstrated to be inadequate.

The balanced scorecard (BSC), a performance measurement framework that provides an integrated look at the business performance of a company by a set of both financial and non-financial measures, seems to be a good solution. There are several papers discussed the application of the BSC in logistics performance evaluation (Hu et al., 2010; Leem et al., 2007). However, the proposed indictors are not reasonable for the TPL firms of China. Moreover, conventional BSC does not rationally consolidate these performance measures, and an incorporation of the BSC and analytic hierarchy process (AHP) is an improvement (Lee et al., 2008). Since fuzziness and vagueness are common characteristics in many decision-making problems (Wang and Zeng, 2005), a fuzzy AHP (FAHP) and BSC method should be able to tolerate vagueness or ambiguity, and therefore, is proposed in this study. It is a more efficient approach in treating the fuzziness of data and analyzing the qualitative factors than other methods (Ding, 2010). The rest of this paper is organized as follows. Review of the concepts of BSC and proposes the performance evaluation criteria is provided subsequently. This is followed by the research methodology, and a numerical example. The paper concludes with a summary of the findings and directions for future research.

TPL ENTERPRISES PERFORMANCE EVALUATION CRITERIA

This section briefly reviews the underlying concepts adopted by this study, such as the definitions of performance evaluation and the BSC. Then, the performance evaluation criteria will be proposed.

Logistics performance evaluation

Performance is referred to as one kind of measurement of the goals of an enterprise, while evaluation is referred to as the goal that an enterprise can effectively obtain during a specific period. Evans et al. (1996) stated that performance evaluation is an important activity of management control, used to investigate whether resources are allocated efficiently; it is applied for the purpose of operational control to achieve a goal adjustment in the short-term and for strategy management and planning in the long run. Logistics performance evaluation is a fairly important issue in logistics management. Some scholars have focused upon: 1) characteristics that measures should possess; 2) perspectives that measures should assume; 3) specific measures that firms should choose (Griffis et al., 2004). However, the traditional performance ranking of TPL companies is based on simple and consistent factors such as financial returns, returns on asset (ROA) and returns on earning (ROE). Nevertheless, performance rankings conducted in this way may not precisely illustrate institutions that embrace strategies for sustaining top performance. Non-financial criteria such as customer satisfaction, community and employee relations can be vital to TPL companies winning strategy, because using only ROA or ROE for performance ranking may not necessarily determine which institution offers the highest returns to the investors, nor does it accurately prove which one is the most profitable.

On the other hand, Kaplan and Norton (1992) described performance evaluation as a way to review the achievements of organizations of both their financial and non-financial objectives. So, a few papers discussed the logistics performance evaluation from multi-perspectives. For example, Rafele (2006) assigned logistics performance indicators into three broad areas: tangible components, ways of fulfillment and informative actions. Evaluation methods of the performance of a logistics company can be diverse. Similar research can be very useful to be adopted as a source of reference. Trappey et al. (2010) used fuzzy cognitive maps and genetic algorithms to model and evaluate the performance of RFID-enabled reverse logistic operations. Hanaoka and Kunadhamraks (2009) applied a fuzzy-AHP approach to assess Logistics Performance for Intermodal Transportation. Johnson and McGinnis (2011) used a DEA approach to evaluate the performance in the warehousing industry.

Balanced scorecard

The concept of BSC was proposed by David Norton, the CEO of Nolan Norton Institute, and Robert Kaplan, a professor at Harvard University. The BSC expands the traditional financial measures into three other dimensions to capture a balanced approach to measure performance in an organization. These additional dimensions are as follows: Customer Focus, Internal Business processes, learning and Growth. Recurring to weight the four dimensions, the enterprises can acquire an unambiguous and precise method to comprehend the strategy of the enterprises. The four perspectives are explained briefly as follows (Kaplan and Norton, 1996):

(1) Financial: This perspective typically contains the traditional financial performance measures, which are usually related to profitability. The measurement criteria are usually profit, cash flow, ROI, economic value added (EVA), and return on invested capital.
(2) Customer: Satisfying customer needs is the objective pursued by corporations. In this perspective, management determines the expected target customers and market segments and monitors the performance of operational units in these target segments.
(3) Internal business process: The objective of this perspective is to satisfy customers and shareholders by excelling at some business processes that have the greatest impact. A complete internal business-process value chain that can meet current and future needs should be constructed. A common enterprise internal value chain consists of three main business processes: innovation, operation and after-sale services.
(4) Learning and growth: This perspective stresses employee performance measurement, such as employee satisfaction, continuity, training and skills, since employee growth is an intangible asset to enterprises that will contribute to business growth. In the other three dimensions, there is often a gap between the actual and target human, system and procedure capabilities. Through learning and growth, companies can decrease this gap (Lee et al., 2008).

The BSC is a popular tool that is applied by many firms to assess their performance in diverse aspects of their organization. The BSC is used by more than 50% of the Fortune 500 companies as a performance measurement and strategic management tool. It provides insights into corporate performance not only for managers seeking ways to improve performance, but also for investors wanting to gauge the organizations' ongoing health. For TPL enterprises the benefits of using BSC are numerous: (1) can be used as a framework to assess and develop a company's strategy; (2) can be used to develop strategic objectives and performance measures to transform a company's strategy into action; (3) it provides a way to measure and monitor the performance of key performance drivers that may lead to the successful execution of a company's strategy; and (4) it is an effective tool to ensure that a TPL company continuously improves its system and process (Wu et al., 2009).

Evaluation criteria for TPL companies based on the BSC

In the early stage of implementing the BSC, it is important to collect as many ideas as possible concerning performance measurement by interviewing business managers and discussing their business vision, mission, and strategies. Venkatraman and Ramanujam (1986) identified two core dimensions of business performance for TPL companies: operational and financial performances. Huo et al. (2008) suggested that operational performance can be further classified into two major dimensions: cost and service performances. Cost performance is related to cost and price, while service performance is related to service reliability, speed, variety, and so forth. Financial performance is defined as the financial and market measures to evaluate the companies' efficiency and effectiveness, including growth rate in market share, growth in annual sales, growth in return on sales, and growth in return on assets (Wang et al., 2010). Leem et al. (2007) inherited the dimensions of the BSC, which allowed the managers to look at the business from four important perspectives to measure the performance on logistics centers. The evaluation criteria are: customer satisfaction; customer retention; new business acquisition; operating efficiency; handling efficiency; execution capability; solvency; profitability; return on investment; human resource; organization system. Hu et al. (2010) expanded the original four perspectives of the BSC based on the stakeholder theory. They also took into account the needs of other external stakeholders to measure goals such as community and government, supplier and environmental protection groups. Based on the work of Hu et al. (2010) and Leem et al. (2007), re-view of logistics performance evaluation literature (Marta et al., 2005; Harding, 1988; Dong and Chen, 2005) and interview with logistics experts, a TPL company performance evaluation hierarchy is constructed as shown in Table 1.

THE PROPOSED INTEGRATED EVALUATION METHOD

Literature review for FAHP

Many FAHP methods were proposed based on the concepts of the fuzzy set theory and hierarchical structure analysis. Some researchers have studied the FAHP which is the extension of the theory proposed by (Saaty, 1980) and also have proved that the FAHP is more effective in these kinds of decision-making processes compared to traditional AHP. Van Laarhoven and Pedrycz (1983) directly extended the AHP method with triangular fuzzy numbers (TFNs). Chang (1996) introduced a new approach of using TFNs for pairwise comparison and also supplied the key point of extent analysis method for deriving the synthetic extent values. This approach is one of the most popular approaches in the FAHP field. Mikhailov (2003) provided a good discussion of the troubles with constructing fuzzy reciprocal matrices using fuzzy comparisons and their reciprocals through the same fashion as the crisp prioritization procedures. Kahraman et al. (2004) implemented Chang's method to measure the customer satisfaction in catering firms in Turkey. Recently, Metin et al. (2009) proposed a practical decision support mechanism on ensuring multiple criteria analysis of shipping registry selection using FAHP.

As a powerful analytical procedure, FAHP is usually combined with other methods in applications. Kuo et al. (2002) integrated FAHP and artificial neural network for the location selecting of convenience store. Rostamzadeh and Sofian (2011) presented a hierarchy multiple criteria decision-making model using FAHP and TOPSIS for prioritizing effective 7Ms (Management, Manpower, Marketing, Method, Machine, Material and Money) to improve

Table 1. Evaluation criteria based on BSC.

Perspective	Goal	Criterion
Finance	Debt-paying ability	Asset-liability ratio i_1
	Operational capacity	Fixed assets turnover ratio, i_2
	Earning capacity	Net profit margin, i_3
	Development capacity	Sales growth rate, i_4
Customer	Customer satisfaction	Customer satisfaction ratio, i_5
	Market expanding	New customer acquisition, i_6
	Market share	Market share ratio, i_7
Internal business process	Quality of service	Integrated operational capacity, i_8
		Abnormal demand fulfillment rates, i_9
		Rate of timely and accurate information services, i_{10}
		Percent of perfect orders, i_{11}
	Operating costs	Unit cost of logistics services, i_{12}
	Information system	The level of information management, i_{13}
	Resources allocation	Facilities allocation level, i_{14}
Learning and growth	Business innovation	Growth rate of new logistics product development, i_{15}
		Growth rate of R&D expenses, i_{16}
	Staff learning	Proportion of staff training, i_{17}
		The number of staff recommendations were adopted or implemented, i_{18}

production systems performance. Jung (2011) proposed a FAHP-goal programming approach for integrated production-planning problem considering manufacturing partners at the background of a TFT-LCD manufacturing firm.

There has been much research regarding the FAHP for the evaluation and selection problems. However, to the best of our knowledge, the performance evaluation of TPL enterprises based on FAHP is not found in the existing literature. Especially, more comprehensive criteria have been considered by introducing balanced scorecard into our model.

Essences of fuzzy AHP

Fuzzy set theory reflects the logical behavior of human brain when faced with imprecision and triangular fuzzy number is a simplification of that.

A fuzzy number is a special fuzzy set $\tilde{M}=(x,u_{\tilde{M}}(x))$, $x\in R$, and its membership function $u_{\tilde{M}}(x)\in[0,1]$. A triangular fuzzy number is denoted as $\tilde{M}=(l,m,u)$, where $l\le m\le u$; l,u is the lower and upper limit of $\tilde{M}$, and m is the mid-value of $\tilde{M}$, $u_{\tilde{M}}(m)=1$. Therefore the membership function can be described as:

$$\mu_{\tilde{M}}(x)=\begin{cases}\dfrac{x-l}{m-l}, & l\le x\le m\\ \dfrac{u-x}{u-m}, & m\le x\le u\\ 0, & \text{otherwise}\end{cases} \quad (1)$$

If $l=m=u$, the fuzzy number turns to a crisp number. Assume that there are two triangular fuzzy numbers $\tilde{M}_1=(l_1,m_1,u_1)$ and $\tilde{M}_2=(l_2,m_2,u_2)$, the main basic operations can be expressed as follows:

$$\tilde{M}_1\oplus\tilde{M}_2=(l_1+l_2,m_1+m_2,u_1+u_2) \quad (2)$$

$$\tilde{M}_1\otimes\tilde{M}_2\cong(l_1l_2,m_1m_2,u_1u_2) \quad (3)$$

$$\tilde{M}_1^{-1}\cong(\frac{1}{u_1},\frac{1}{m_1},\frac{1}{l_1}) \quad (4)$$

In this study, nine-point fundamental scale is used for the pair-wise comparisons. In other words, the triangular fuzzy numbers, from $\tilde{1}$ to $\tilde{9}$ and their reciprocals are employed to capture the vagueness in the pair-wise comparisons. A set of numbers $\tilde{9}$ (8, 9, 9) represents "extremely important", $\tilde{5}$ (4, 5, 6) for "relatively important", $\tilde{1}$ (1, 1, 2) stands for "equally important", and (1,1,1) will be employed while one criterion is compared with itself. So the n-dimensional fuzzy pairwise comparison matrix $\tilde{A}$ can be constructed.

$$\tilde{A}=\begin{bmatrix}(1,1,1) & \tilde{r}_{12} & \cdots & \tilde{r}_{1n}\\ \tilde{r}_{21} & (1,1,1) & \cdots & \tilde{r}_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ \tilde{r}_{n1} & \tilde{r}_{n2} & \cdots & (1,1,1)\end{bmatrix}$$

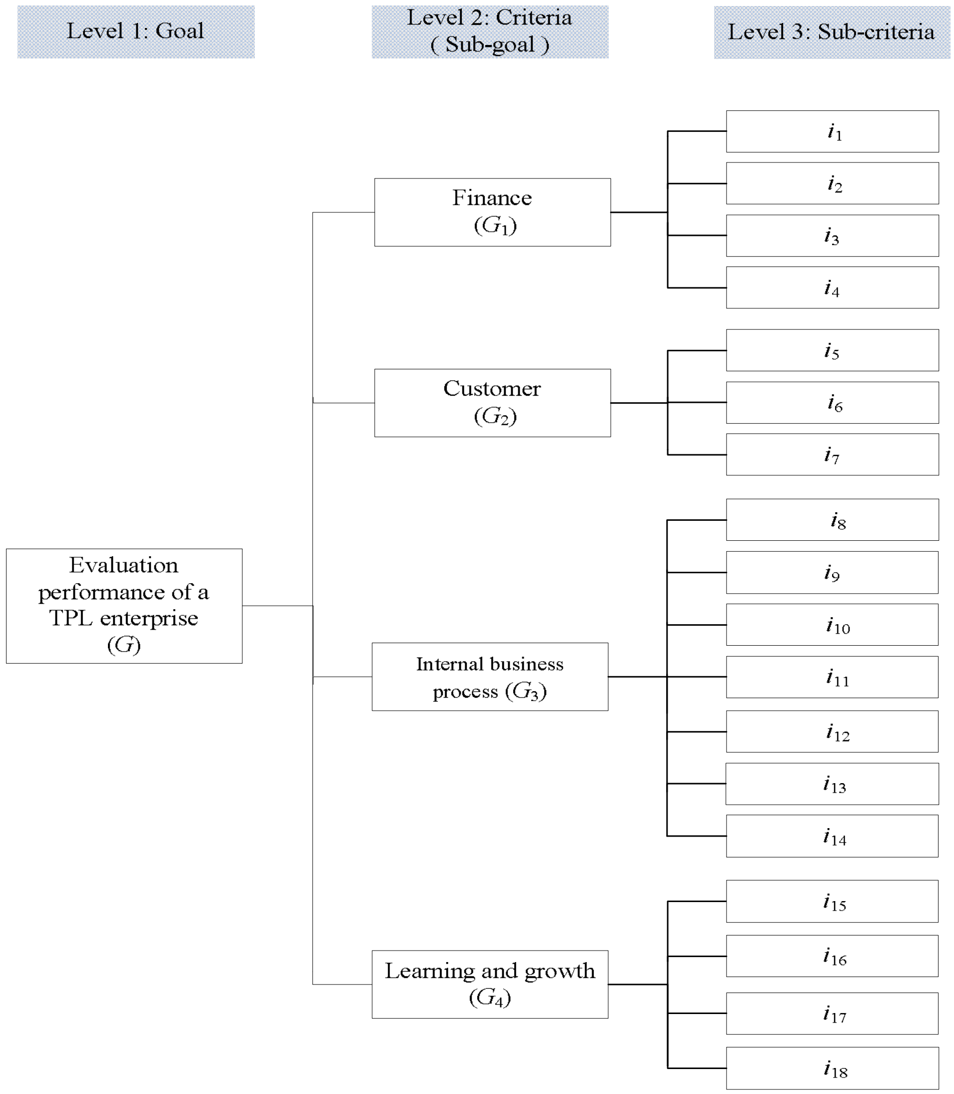

Figure 1. Hierarchical structure of evaluating a TPL enterprise performance.

where $\tilde{r}_{ij}$ is triangular fuzzy number $\tilde{r}_{ij} = (l_{ij}, m_{ij}, u_{ij})$ denoting someone's judgement between elements i and j for all $i, j \in \{1,2,...,n\}$; $\tilde{r}_{ij} = \tilde{r}_{ji}^{-1}$.

Computational procedure of fuzzy AHP

Step 1: Constructing the hierarchical framework for pair-wise comparison

As shown in Figure 1, the ultimate goal denoted by G is the performance evaluation of a TPL enterprise. There are 4 sub-goals notated by G1, G2, G3, and G4 under the overall goal. The level three represents the important factors of the sub-goals and there are 18 criteria in total. After the hierarchy architecture is established, fuzzy pairwise comparison matrices of the ultimate and sub goals can be given respectively for each evaluator.

Step 2: Consistency analysis

The fuzzy judgement matrices have been established, but it may bring a situation that is the inconsistent in logic, such as one evaluator said "criterion A is more important than criterion B, criterion B seems moderately more important when compared with criterion C, criterion A is equally important compared with criterion B". To avoid or reduce the suffering from that, we have to analyze the consistency of the evaluations.

According to the definition in Satty (1980), the consistency ration (*CR*) can be calculated with Equation 5:

$$CR = \frac{CI}{RI} \tag{5}$$

RI is the average index for randomly generated weights (as shown in Table 2), the consistency index (*CI*) can be approximated as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{6}$$

Table 2. The value of random consistency index (*RI*).

Dimension *n*	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

where $\lambda_{\max}$ is the maximum eigenvalue, and *n* is the dimension of the matrix.

The matrix $\widetilde{A}=[\widetilde{r}_{ij}]$ is constituted with triangular fuzzy number $\widetilde{r}_{ij}=(l_{ij},m_{ij},u_{ij})$. To simplify the calculation of the *CR*, a new matrix $A=[m_{ij}]$, formed by the crisp value m_{ij} instead of $\widetilde{A}$, will be used. According to the research of Csutora and Buckey (2001), if *A* is consistent, then $\widetilde{A}$ is consistent. In general, if the *CR* is less than 0.1, the comparisons are acceptable. Otherwise, the original values in the matrix must be revised by the evaluator, or it could affect the overall results negatively.

Step 3: Aggregation of group decisions

Assume that there are *K* experts, in Step 1 each of experts supplied us a fuzzy judgement matrix about the goal, $\widetilde{A}^1=[\widetilde{r}_{ij}^1],...,\widetilde{A}^k=[\widetilde{r}_{ij}^k]$, the ultimate fuzzy pair-wise comparison matrix $\widetilde{A}=[\widetilde{r}_{ij}]$ could be combined with $\widetilde{A}^1,...,\widetilde{A}^k$.

$$l_{ij}=\left(\prod_{k=1}^{K} l_{ij}^k\right)^{1/K},\quad m_{ij}=\left(\prod_{k=1}^{K} m_{ij}^k\right)^{1/K},\quad u_{ij}=\left(\prod_{k=1}^{K} u_{ij}^k\right)^{1/K} \tag{7}$$

Step 4: Determining fuzzy priorities of criteria

The priorities will be estimated based on the synthetic extent analysis in (Chang, 1996) which requires the consistent of evaluations. Assume that $\widetilde{M}_1=(l_1,m_1,u_1)$ and $\widetilde{M}_2=(l_2,m_2,u_2)$ are triangular fuzzy numbers. $V(\widetilde{M}_2\geq\widetilde{M}_1)$ denotes the degree of possibility of $\widetilde{M}_2\geq\widetilde{M}_1$ and it can be formulated as following:

$$V(\widetilde{M}_2\geq\widetilde{M}_1)=hgt(\widetilde{M}_1\cap\widetilde{M}_2)=\begin{cases}1 & ,\ if\ m_2\geq m_1\\ 0 & ,\ if\ \ l_1\geq u_2\\ \dfrac{l_1-u_2}{(m_2-u_2)-(m_1-l_1)}, & otherwise\end{cases} \tag{8}$$

Let $X=\{x_1,x_2,...,x_n\}$ be an object set, and $U=\{u_1,u_2,...,u_m\}$ be a goal set.

According to the extent analysis of (Chang, 1996), each object is taken and extent analysis for each goal is performed respectively. Therefore the *m* extent analysis values for each object are obtained with the following signs:

$$\widetilde{M}_{g_i}^1,\widetilde{M}_{g_i}^2,...,\widetilde{M}_{g_i}^m \quad i=1,2,...,n$$

With respect to the ith object, the value of fuzzy synthetic extent is defined as:

$$\widetilde{S}_i=\sum_{j=1}^{m}\widetilde{M}_{g_i}^j\otimes\left[\sum_{i=1}^{n}\sum_{j=1}^{m}\widetilde{M}_{g_i}^j\right]^{-1} \tag{9}$$

$$W=(\min V(\widetilde{S}_1\geq\widetilde{S}_k),\min V(\widetilde{S}_2\geq\widetilde{S}_k),...,\min V(\widetilde{S}_n\geq\widetilde{S}_k))^T \quad k\in\{1,2,...,n\} \tag{10}$$

Now, $\widetilde{S}_i$ is still a triangular fuzzy number. But the non-fuzzy weight vector *W* can be estimated by equation 10. After the normalization of *W*, the normalized weight vector W_0 can be obtained.

Step 5: Calculating criteria weights and ranking the alternatives

According to the normalized weight vectors, the weight of hierarchy model could be obtained, that means global weights of all sub-criteria $W_T=(w_1,w_{2,...,}w_q)^{\mathrm{T}}$ can be gained, where *q* is the number of sub-criteria. In tradition AHP, the alternative score can be obtained from the pair-wise comparison, but as a matter of fact that is not a sensible choice especially when the criteria can be measured in some way or plenty of alternatives needs to be ranked. They are the just situation we have to face. In our study, most of the criteria can be calculated by the definition and the remaining can be acquired through an expert assessment method. For the purpose of obtaining standardized scores regarding to each criteria, all the scores will be transformed into a common scale between 0 and 1 by Equations 11 and 12:

$$S_{pq}^*=\frac{S_{pq}-\min_p(S_{pq})}{\max_p(S_{pq})-\min_p(S_{pq})} \tag{11}$$

$$S_{pq}^*=\frac{\max_p(S_{pq})-S_{pq}}{\max_p(S_{pq})-\min_p(S_{pq})} \tag{12}$$

Where S_{pq} is the score of enterprise *p* in regard to criterion *q*,

Table 3. Fuzzy judgement matrix with respect to *G* of expert 1.

G	*G*1	*G*2	*G*3	*G*4
G_1	(1, 1, 1)	(0.43, 0.53, 0.61)	(0.30, 0.36, 0.50)	(0.46, 0.70, 0.84)
G_2	(1.63, 1.87, 2.31)	(1, 1, 1)	(0.52, 0.61, 0.97)	(0.90, 1.07, 1.38)
G_3	(2.01, 2.76, 3.36)	(1.03, 1.65, 1.94)	(1, 1, 1)	(1.65, 2.32, 2.99)
G_4	(1.19, 1.42, 2.17)	(0.72, 0.93, 1.11)	(0.33, 0.43, 0.61)	(1, 1, 1)

$n = 4$; $\lambda\ max = 4.0101$; $CI = 0.0034$; $RI = 0.90$; $CR = 0.0037 < 0.1$.

$\min_p(S_{pq})$, $\max_p(S_{pq})$ is the minimum and maximum score about criterion *q*. Equation 11 is used for the positive indicators and Equation 12 is utilized for the negative indicators.

Then, the final evaluation of enterprise *p* is given by the Equation 13:

$$C_p = \sum_{q=1}^{Q} w_q * S_{pq}^* \quad (13)$$

At last, the ranking of the enterprises could be obtained via the comparison of the C_p.

NUMERICAL EXAMPLE

The problem discussed here is concerned with evaluating the performance of TPL enterprises. A decision-making group with five experts who come from the senior management department is formed. The expert selection criteria are: active career in related business for at least 10 years with rich experiences of the enterprise management; a global vision beyond local and temporary concerns; and accessibility and willingness to engage in intellectual dialogue.

Based on the hierarchy model as shown in Figure 1, the fuzzy judgement matrices of each expert can be obtained by the pairwise comparison. As for Gi, here have five fuzzy judgement matrices. Tables 3 and 4 are the two of them. Consistency of each fuzzy comparison matrix also has been checked.

In the similar way, the fuzzy comparison matrices with respect to *G* of other experts have been constructed. Then, the ultimate fuzzy pair-wise comparison matrix can be obtained according to Equation 7 as shown in Table 5. Based on the aforementioned method, $\tilde{S}_i$ can be obtained by Equation 9 as follows:

$$\tilde{S}_1 = (0.0818, 0.1376, 0.2448),$$
$$\tilde{S}_2 = (0.1503, 0.2516, 0.4207),$$
$$\tilde{S}_3 = (0.2051, 0.3930, 0.7009),$$
$$\tilde{S}_4 = (0.1362, 0.2177, 0.3780).$$

The degree of possibility of $\tilde{S}_i \geq \tilde{S}_j (i \neq j)$ can be determined by Equation 8.

$$V(\tilde{S}_1 \geq \tilde{S}_2) = 0.4532,\ V(\tilde{S}_1 \geq \tilde{S}_3) = 0.1345,\ V(\tilde{S}_1 \geq \tilde{S}_4) = 0.5755,$$
$$V(\tilde{S}_2 \geq \tilde{S}_1) = 1.0000,\ V(\tilde{S}_2 \geq \tilde{S}_3) = 0.6040,\ V(\tilde{S}_2 \geq \tilde{S}_4) = 1.0000,$$
$$V(\tilde{S}_3 \geq \tilde{S}_1) = 1.0000,\ V(\tilde{S}_3 \geq \tilde{S}_2) = 1.0000,\ V(\tilde{S}_3 \geq \tilde{S}_4) = 1.0000,$$
$$V(\tilde{S}_4 \geq \tilde{S}_1) = 1.0000,\ V(\tilde{S}_4 \geq \tilde{S}_2) = 0.8705,\ V(\tilde{S}_4 \geq \tilde{S}_3) = 0.4966.$$

According to Equation 10, the weight vector is given as $W = (0.1345, 0.6040, 1, 0.4966)^{\mathrm{T}}$ and after the normalization process, the weight vector W_0 with respect to *G* can be presented as:

$$W_0 = (0.0602, 0.2702, 0.4474, 0.2222)^{\mathrm{T}}.$$

Then, the weights of criteria respect to each sub-goal G_1, G_2, G_3, G_4 are calculated similarly as follows:

$$W_0^{G_1} = (0.1274, 0.1333, 0.5127, 0.2266)^{\mathrm{T}}, W_0^{G_2} = (0.2819, 0.1993, 0.5187)^{\mathrm{T}}$$
$$W_0^{G_3} = (0.1238, 0.1356, 0.1541, 0.1745, 0.1996, 0.1151, 0.0973)^{\mathrm{T}}, W_0^{G_4} = (0.1795, 0.2706, 0.3458, 0.2042)^{\mathrm{T}}.$$

So, the global weights of all sub-criteria W_{T} can be obtained as:

$$W_{\mathrm{T}} = (0.0077, 0.0080, 0.0309, 0.0136, 0.0762, 0.0539, 0.1402, 0.0554, 0.0607, 0.0689,$$
$$0.0781, 0.0893, 0.0515, 0.0435, 0.0399, 0.0601, 0.0768, 0.0454)^{\mathrm{T}}.$$

Table 4. Fuzzy judgement matrix with respect to *G* of expert 2.

G	G_1	G_2	G_3	G_4
G_1	(1, 1, 1)	(0.39, 0.57, 0.85)	(0.24, 0.32, 0.65)	(0.38, 0.57, 0.83)
G_2	(1.18, 1.75, 2.58)	(1, 1, 1)	(0.53, 0.67, 0.93)	(0.94, 1.27, 1.42)
G_3	(1.54, 3.10, 4.17)	(1.07, 1.49, 1.89)	(1, 1, 1)	(1.21, 1.77, 2.53)
G_4	(1.21, 1.77, 2.60)	(0.70 ,0.79, 1.06)	(0.40, 0.57, 0.83)	(1, 1, 1)

$n = 4$; $\lambda\ max = 4.0061$; $CI = 0.0020$; $RI = 0.90$; $CR = 0.0023 < 0.1$

Table 5. Fuzzy judgement matrix with respect to G of all experts.

G	G_1	G_2	G_3	G_4
G_1	(1, 1, 1)	(0.32,0.55,0.81)	(0.21, 0.37, 0.69)	(0.44, 0.64, 0.88)
G_2	(1.24, 1.81, 2.6)	(1,1,1)	(0.52, 0.73, 0.87)	(0.86, 1.14, 1.34)
G_3	(1.46, 2.88, 4.11)	(1.13, 1.49, 1.94)	(1, 1, 1)	(1.35, 1.94, 2.63)
G_4	(1.15,1.59,2.32)	(0.75, 0.89, 1.16)	(0.38, 0.57, 0.74)	(1, 1, 1)

$n = 4$; $\lambda\ max = 4.0722$; $CI = 0.0241$; $RI = 0.90$; $CR = 0.0267 < 0.1$.

Table 6. The detailed date of eight enterprises.

i	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
i_1	1	0.734	0.4232	0.8536	0.3765	0	0.5439	0.2637
i_2	1	0.3723	0.5365	0.7435	0	0.4543	0.8273	0.6238
i_3	0.2452	1	0.4843	0.8528	0.5928	0.6459	0.3902	0
i_4	0.8628	0.5792	1	0.6762	0.4634	0.3245	0.5283	0
i_5	0.6237	0	0.5834	0.7324	1	0.2674	0.7329	0.9021
i_6	0.782	0.6293	0	0.4249	1	0.2857	0.623	0.7342
i_7	0.2719	0.4373	0.5382	1	0.3348	0.7145	0.7542	0
i_8	0.7526	0.6457	0.4723	0.8340	1	0	0.2564	0.5289
i_9	0.6425	0	0.5203	0.7256	0.8352	1	0.4873	0.2754
i_{10}	0	0.4183	0.6428	1	0.7634	0.7294	0.5332	0.8245
i_{11}	1	0.4098	0	0.7288	0.8346	0.8203	0.3757	0.6238
i_{12}	0.4396	0.3292	0.7659	0	1	0.8527	0.7293	0.3723
i_{13}	0.3253	0.638	0.6925	0.8526	0.692	0	1	0.7361
i_{14}	0.2578	0.4259	0.6347	0	0.7319	0.6283	1	0.5273
i_{15}	1	0.2653	0.4371	0.9025	0.7526	0.3829	0.3872	0
i_{16}	1	0.2274	0	0.5073	0.8239	0.5283	0.4983	0.6283
i_{17}	0	0.3982	0.6492	0.8024	1	0.5746	0.4729	0.8367
i_{18}	0.8743	0.3769	0.5728	0.4793	0	0.6284	1	0.7284

Now, there are eight TPL enterprises will be evaluated and the date in 2010 is shown in Table 6. The data has already been normalized by Equations 11 and 12.

According to Step 5, the final scores of these TPL enterprises can be calculated by using Equation 13 as follows:

$$C = (C_1, C_2, C_3 ..., C_8)^{\mathrm{T}} = (0.5285,\ 0.3920,\ 0.4830,\ 0.6768,\ 0.7416,\ 0.5621,\ 0.6223,\ 0.4961)^{\mathrm{T}}.$$

CONCLUSIONS AND FUTURE RESEARCH

Performance measurement to TPL companies in China, a complex multi-attribute decision-making problem, has not received adequate attention from researchers or practitioners. This paper proposes a valuable approach based on the FAHP and BSC for evaluating the performance of TPL enterprises. The analytic hierarchy is structured by the four major perspectives of the BSC including financial, customer, internal business process, and learning and growth, followed by performance indicators. Because the human decision-making process usually contains fuzziness and vagueness, the FAHP is adopted to solve the problem.

The main contributions are two fold. First, based on literature review and interview with experts in the logistics field, 18 most important performance indicators are finalized for the performance assessment of TPL enterprises. These indicators can be a reference for TPL enterprises in performance evaluation. Second, a systematic performance evaluation model is designed based on the fuzzy set theory and AHP to provide reasonable guidance to managers regarding performance evaluation and strategies for improving performance. The illustrative example has demonstrated the thoughtfulness, flexibility, and efficiency of the proposed model to directly tap the subjectivity and preferences of the decision makers. Future studies can adopt additional fuzzy multi-attribute approaches (such as fuzzy TOPSIS and fuzzy outranking methods) to estimate the relative weights of the influences on a TPL's performance. Moreover, a knowledge-based or expert system can be integrated to help decision-makers make the calculations more concisely and interpret the results in each step.

ACKNOWLEDGEMENTS

The authors gratefully thank the editor and anonymous referees for their valuable comments. This research is partially supported by the National Natural Science Foundation of China (No.70801030), Fundamental Research Funds for the Central Universities (HUST: 2010MS133); Educational Commission of Hubei Province of China (D20112201); and Humanities and Social Sciences Foundation of Chinese Ministry of Education (11YJC630275).

REFERENCES

Chang DY (1996). Applications of the extent analysis method on fuzzy AHP. Eur. J. Oper. Res., 95(3):649-655.

Ding JF (2010).Critical factors influencing customer value for global shipping carrier-based logistics service providers using Fuzzy AHP approach. Afr. J. Bus. Manage., 4(7): 1299-1307.

Dong M, Chen F (2005). Performance modeling and analysis of integrated logistic chains: An analytic framework. Eur. J. Oper. Res., 162(1):83-98.

Evans H, Ashworth G, Chellew M, Davidson A, Towers D (1996). Exploiting activity-based information: Easy as ABC. Manage. Account., 74(7):24-29.

Griffis SE, Cooper M, Goldsby TJ, Closs DJ (2004). Performance measurement: measure selection based upon firm goals and information reporting needs. J. Bus. Logist., 25(2): 95-118.

Hanaoka S, Kunadhamraks P (2009). Multiple criteria and fuzzy based evaluation of logistics performance for intermodal transportation. J. Adv. Transp., 43(2):123-153.

Harding FE (1988). Logistics service provider quality: private measurement, evaluation, and improvement. J. Bus. Logist., 19(1):103-120.

Hu MM, Wang X, Li S, Deng YL, Yue QQ, Tang QF(2010). Research on performance evaluation of logistics enterprises based on the balanced scorecard. In: Proc. Int. Conf. Intell. Comp. Technol. Automation. IEEE CS Press: 65-68.

Huo B, Selen W, Yeung JHY, Zhao X. (2008).Understanding drivers of performance in the 3PL industry in Hong Kong. Int. J. Oper. Prod. Manage., 28(8): 722-800.

Johnson A, McGinnis L (2011). Performance measurement in the warehousing industry. IIE Trans., 43(3): 220-230.

Jung H (2011). A fuzzy AHP-GP approach for integrated production-planning considering manufacturing partners. Expert Syst. Appl., 38(5): 5833-5840.

Kahraman C, Cebeci U, Ruan D. (2004).Multi-attribute comparison of catering service companies using fuzzy AHP: the case of Turkey. Int. J. Prod. Econ., 87(2):171-184.

Kaplan RS, Norton DP (1992). The balanced scorecard measures that drive performance. Harv. Bus. Rev., 70(1):71-79.

Kaplan RS, Norton DP (1996). The balanced scorecard: translating strategy into action. Harvard Business School Press, Boston .

Kuo RJ, Chi SC, Kao SS (2002).A decision support system for selecting convenience store location through integration of fuzzy AHP and artificial neural network. Comput. Ind., 47(2):199-214.

Lee AHI, Chen WC, Chang CJ (2008). A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan. Expert Syst. Appl., 34: 96-107.

Leem B, Hong M, Kang J, Yim BJ (2007). Modeling the metrics for measuring the performance on logistics centers by BSC and ANP in Korean context. In Proc of Int. Conf. Manage. Eng. Technol., Portland State University Press: 2412-2417.

Marta O, Lorenzo D, Raffaele P, Walter U (2005).Multiple-attribute decision support system based on fuzzy logic for performance assessment. Eur. J. Oper. Res., 160 (3):710-725.

Metin C, Er ID, Ozok AF (2009). Application of fuzzy extended AHP methodology on shipping registry selection: The case of Turkish maritime industry. Expert Syst. Appl., 36(1):190-198.

Mikhailov L (2003). Deriving priorities from fuzzy pairwise comparison judgements. Fuzzy Sets Syst., 134(3):365-385.

National Development and Reform Commission, National Bureau of Statistics and China Federation of Logistics and Purchasing (2011). Briefing of China's Logistics Operations in 2010, National Development and Reform Commission, Beijing.

Rafele C (2006). Logistics service measurement: a reference framework. J. Manuf. Technol. Manage., 15(3):280-290.

Rostamzadeh R, Sofian S (2011). Prioritizing effective 7Ms to improve production systems performance using fuzzy AHP and fuzzy TOPSIS (case study). Expert Syst. Appl., 38(5):5166-5177.

Saaty TL (1980). The Analytic Hierarchy Process. McGraw-Hill, NY.

Susanne H, Monica A. (2003). Strategic development of third party logistics providers. Ind. Mark. Manage., 32(2):139-149

Tian Y, Lai F, Daniel F (2008).An examination of the nature of trust in logistics outsourcing relationship: empirical evidence from China. Ind. Manage. Data Syst., 108(3):346-67.

Trappey AJC, Trappey CV, Wu CR (2010). Genetic algorithm dynamic performance evaluation for RFID reverses logistic management. Expert Syst. Appl., 37 (11):7329-7335.

Van Laarhoven PJM, Pedrycz W (1983). A fuzzy extension of Saaty's priority theory. Fuzzy Sets Syst., 11(1-3):199-227.

Venkatraman N, Ramanujam V (1986). Measurement of business performance in strategy research: a comparison of approaches. Acad. Manage. Rev., 11(4):801-14.

Wang L, Zeng YR (2005). The risk identification and assessment in E-Business development. Lect. Notes Artif. Int., 3614: 1142-1149.

Wang Q, Huo BF, Lai FJ, Chu ZF (2010). Understanding performance drivers of third-party logistics providers in mainland China: A replicated and comparative study. Ind. Manage. Data Syst., 110(9):1273-1296.

Wu HY, Tzeng GH, Chen YH (2009). A fuzzy MCDM approach for evaluating banking performance based on balanced scorecard. Expert Syst. Appl., 36: 10135-10147.