

Combination Forecasting Model for Predicting the Shelf Life of Two-state Materials Based on Support Vector Machine

Li Zhiwei^{1,2}, Gao Qi¹, Liu Shenyang^{*,1} and Li Zhen¹

¹Department of Equipment Command and Management, Mechanical Engineering College, Shijiazhuang, Hebei, 050003, P.R. China

²Maintenance Depot, Shijiazhuang Flying College, Shijiazhuang, Hebei, 050003, P.R. China

Abstract: A combination forecasting model based on Support Vector Machine (SVM) whose objective is to minimize the structure risk, is proposed. The storage failure of two-state materials tends to fail immediately without any recognizable defeats prior to the failure, which increases the difficulty of forecasting, so the combination forecasting model is often used to optimize the prediction effect. The core ideas of previous combination forecasting models such as those based on forecasting error and those based on nonlinear weighted average are finding the optimal weights, but the structure of forecasting model is fixed. In this paper, three single forecasting models, Weibull distribution statistic method, BP neural network prediction method and SPFM (Sliding Polynomial Fitting Method) are chosen in which their forecast mechanisms are completely different. The results of single forecasting methods are used as training set of SVM. By using libsvm toolbox, we can get the nonlinear mapping functions that have the minimum structure risk. At last, a simulation is conducted to verify this model by using the data from Petroleum Center.

Keywords: Combination forecast, shelf life forecasting, SVM, two-state materials.

1. INTRODUCTION

Due to a variety of environmental stresses, two-states materials in storage tend to fail immediately without any recognizable defeats prior to any failure. The typical two-state materials include electronic products, electro-mechanical products and POL (Petrol-Oil and Lubricants) products. The mechanism of storage failure is complicated and non-linear owing to environmental factors, managing factors and designing factors, which can hardly be predicted. If only one forecasting method is selected, we have to take the risk of selecting mistake. Meanwhile, different models have different strong points and weak points. They are interrelated and supplement to each other. Bates and Granger [1] are the earliest people who research the combination forecasting systematically. Through the development in the last few decades, combination forecasting method has become a very important research hotspot in the current world. According to different forecast criteria, combination forecasting models can be divided into the following categories: least variance method [2], unconstrained least square method [3], constrained least square method [4], Bayes method [5], combination forecasting method based on different criteria and standard, recursive combination forecasting method. F. Geng proposed the excellent combination of forecast criteria aiming at minimum the error sum of squares without non-negative constraints, and explained that unweighted average method can perform the excellent combination forecast using the nature of absolute

error information matrix. X. W. Tang [6] studied the combination forecasting error boundary. Considering the standard deviation of predictive precision, H. Y. Chen [7] developed a combination forecasting model for better precision of prediction, and the weight coefficients of combination forecasting is calculated by linear programming. At present, different weighted means are assigned to each single forecasting method [8-20], and the common nature is that the structure of forecasting models is fixed no matter what the values of weight coefficient are. In 1963, the original SVM (support vector machine) was invented by Vladimir N. Vapnik and Alexey Ya. Chervonenkis. A SVM method constructs a hyperplane or set of hyperplanes in a high- or infinite- dimensional space, which can be used for regression or forecasting. The SVM method can reduce the structure risk of combination forecasting model. In order to improve the forecasting accuracy, this paper uses Weibull distribution statistic method, BP neural network method and SPFM to obtain the reliability of two-state materials respectively, and then regards these forecasting outcomes as training data of SVM model to calculate the 90% reliability of two-state materials.

2. SINGLE FORECASTING MODELS

The two-state materials in storage are either in good state or in failure state, and the storage failure data obtained from technical inspection annually mainly consist of three items: storage ages, sample size and sample failure numbers. Considering the shelf failure characters of two-state materials, Weibull distribution statistic method, BP neural network method and SPFM are usually used to forecast the shelf life of two-state materials and their predictive

*Address correspondence to this author at the Department of Equipment Command and Management, Mechanical Engineering College, Shijiazhuang, Hebei, 050003, P.R. China; Tel: 15127134769; E-mail: 1037279749@qq.com

mechanisms are quite different. We can use these methods to predict the shelf life of two-state materials respectively.

2.1. Weibull Distribution Statistic Method

The statistic method used in this paper was two-parameter Weibull distribution. This probability distribution was initially developed by Waloddi Weibull, professor of Applied Physics at Royal Institute of Technology in Stockholm, Sweden. It was widely applied in the field of reliability, fatigue testing, and quality control. The distribution equation is described as follows:

$$F(X) = 1 - e^{-(X/\theta)^\beta} \quad X \geq 0$$

where $F(X)$ represents the probability of shelf life. The term β and θ refer to shape and scale parameters respectively. The best method for estimating these parameters is the Maximum Likelihood (ML) method, which gives a solution using simultaneous equations. Generally, these unknown parameters are determined from ML method of the extreme value distribution. Their relationship is expressed by the following equations: $X = \exp(Y)$, $\theta = \exp(\lambda)$, and $\beta = 1/\delta$, where X, Y are random variables for Weibull and extreme value distribution. λ and δ are the parameters of extreme value distribution. The ML equation and the two partial differential equations of extreme value distribution are described in Eq.(1)~Eq.(3).

Extreme value log likelihood equation:

$$L = \sum' [-\ln(\delta) - \exp[(y - \lambda)/\delta] + [(y - \lambda)/\delta]] - \sum'' \exp[(y - \lambda)/\delta] \tag{1}$$

Constraining partial differential equations:

$$\partial L / \partial \lambda = (1/\delta) \left[\sum \exp[(y - \lambda)/\delta] - r \right] = 0 \tag{2}$$

$$\partial L / \partial \delta = (1/\delta) \left[\sum [(y - \lambda)/\delta] \exp[(y - \lambda)/\delta] - r \right] = 0 \tag{3}$$

where,

\sum =sum of failure and censored data;

\sum' =sum of failure data;

\sum'' =sum of censored data;

r =the number of failure in a sample.

One problem is that the above constraining equations Eq.(2) and Eq.(3) are not linear. Therefore, there equations must be iteratively solved by the computer. The best iterative method is the well-known Newton-Raphson method or EM (Estimation Maximization) method.

2.2. Back Propagation (BP) Network Method

Artificial Neural Network (ANN) is a non-linear computational method based on the locally sensitivity characteristic of human brain neurons, which was applied widely in nonlinear, uncertain or unknown complex engineering problems without explicit understanding of the physical mechanism, such as pattern recognition, fault diagnosis and trend prediction problems. ANN such as BP network, RBF network and Elman network are used commonly in prediction problems. The RBF network needs a large number of sample data, but the sample data obtained in practice is limited. The Elman network is a single recursive network that has a context layer as an inside self-referenced layer, which enhances its ability to predict the time-varying system. However, the shelf life of most two-state materials is quite stable. We can draw a conclusion that BP network is suitable for this task.

2.3. Sliding Poly-Nominal Fitting Method (SPFM)

The SPFM is a method based on PFM (Poly-nominal Fitting Method), which has similar principle. The essence of this method is least-squares estimate, which is also a process of matrix calculation.

Consider a set of time series data $(s_1, y_1), (s_2, y_2), \dots, (s_k, y_k)$, applying SPFM to find a $m(m < k - 1)$ order polynomial $P(s)$ to approximate the functional value of y_i at a point of s_i . The matrix equation $\mathbf{SA} = \mathbf{Y}$ is equivalent to least-squares estimate.

where,

$$\mathbf{S} = \begin{bmatrix} s_1^m & \dots & s_1 & 1 \\ s_2^m & \dots & s_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ s_k^m & \dots & s_k & 1 \end{bmatrix}_{k \times (m+1)}, \quad \mathbf{A} = \begin{bmatrix} a_m \\ a_{m-1} \\ \vdots \\ a_0 \end{bmatrix}_{(m+1) \times 1}, \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}_{(k+1) \times 1}$$

when $|\mathbf{S}^T \mathbf{S}| \neq 0$, we can get $\mathbf{A}(\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{Y}$, and then we can predict $\hat{y}_{k+1} = \sum_{j=0}^m a_j s_{k+1}^j$ by \mathbf{A} .

3. NON-LINEAR COMBINATION FORECASTING MODEL

3.1. Introduction of Combination Forecasting Model

The combination forecasting model can take full advantage of every single forecasting method to improve the predictive precision. In the n-model forecasting combination case, the combined forecast value is \bar{y}_i , and the series of Y_i are technical inspection data of a certain two-state materials,

where $t = 1, 2, \dots, m$ and m represent the numbers of technical inspection. The aim of combination forecast is finding a mapping function φ to assure $\min \|C_t - Y_t\|, t = 1, 2, \dots, m$, where $\|\bullet\|$ represents Euclidean norm $C_t = (\varphi(\bar{y}_{it})), i = 1, 2, \dots, n$. The key point is how to find a proper mapping function φ . The usual combination forecast method is to define mapping function as weighted mean operation. In other word, define a weight vector $W_i (i = 1, 2, \dots, n)$, $C_t = \sum_{i=1}^n w_{it} \cdot \bar{y}_{it}, t = 1, 2, \dots, m$, where $\sum_{i=1}^n w_{it} = 1, w_{it} \geq 0$.

This method has a flaw. The weighted vector element w_i is a constant, but the shelf life of two-state materials is relevant to many factors, including environment factor, design factor and management factor that can hardly be represented by a constant. Thus, the mapping function φ should be defined as a non-linear function.

3.2. Combination Forecast Model of SVM

The combination forecast model of SVM means the inner product kernel function of SVM is used as the mapping function. SVM can minimize both the empirical risk and confidence interval. The basic idea of SVM is mapping the data which is inseparable in low-dimensional space to a high dimensional eigenspace. The algorithm model is as follows:

$$\begin{cases} y_i - (w \cdot x_i) - b \leq \varepsilon + \xi_i \\ (w \cdot x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{4}$$

where ε is precision error, and ξ_i, ξ_i^* note relaxation factors when considering the fitting error. Based on the theorem of optimal separating hyperplanes, the equivalence of optimal regression function is:

$$\min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \right) \tag{5}$$

where $C > 0$ represents penalty coefficient to control the penalty degree when sample data exceeds precision error ε . Transform the above equation into its dual problem by Lagrange method:

$$\sum_{i=1}^n (a_i - a_i^*) = 0, 0 \leq a_i, a_i^* \leq C \tag{6}$$

For Lagrange factors a_i and a_i^* , the maximization target function is:

$$\begin{aligned} W(a_i, a_i^*) = & \sum_{i=1}^n (a_i + a_i^*) + \sum_{i=1}^n y_i (a_i^* - a_i) \\ & - \frac{1}{2} \sum_{i,j=1}^n (a_i^* - a_i)(a_j^* - a_j)(x_i - x_j) \end{aligned} \tag{7}$$

Thus, the regression function is

$$f(x) = (w \cdot x) + b = \sum_{i=1}^n (a_i^* - a_i)(x_i - x) + b \tag{8}$$

The samples corresponding to a_i, a_i^* are called ‘‘support vectors’’. If only we can replace Eq. (4) with kernel function $K(x_i, x)$, the inner product operation can be transformed into the following non-linear fitting function:

$$f(x) = \sum_{i=1}^n a_i K(x_i \cdot x) + b \tag{9}$$

where, a_i notes coefficient, x_i notes support vector and b represents threshold. a, b can be solved by least-squares method. Considering the characteristic of storage, RBF is chosen to be the core function of $K(x_i, x)$, which is:

$$K(x_i, x) = \exp \left\{ - \frac{|x_i - x|^2}{\sigma^2} \right\} \tag{10}$$

where, σ^2 notes the width parameter of kernel function, which adjusts the radial range of RBF.

4. ANALYSIS OF A CASE

Many types of POL products are typical two-state materials. A consecutive study was conducted in cooperation with Petroleum Center which routinely conducts shelf life testing on various stored POL. The laboratory shelf life data for Silicone Brake Fluid which meets the specification MIL-B-46176 are listed in Table 1:

Table 1. Laboratory shelf life data for silicone brake fluid.

Specification	MIL-B-46176					
	2	4	6	8	10	11
storage age (year)	2	4	6	8	10	11
total number of samples (batches)	24	33	14	3	7	9
the number of failed samples (batches)	1	3	2	0	0	2

4.1. Forecast Shelf Life with Single Forecast Model

4.1.1. Weibull Statistic Model

We can substitute the shelf life data in Table 1 into Eq. (1) ~Eq. (3), and solve them by Newton-Raphson method. The results show that shape parameter $\beta = 2.2432$, scale parameter $\theta = 17.0928$, the 90 percent reliable life estimate of two-sided 90 percent confidence intervals is 6.3. Fig. (1) shows the probability shelf life diagram obtained for a silicone brake fluid product evaluated.

4.1.2. BP Network

Given the sample data in Table 1, we can initialize a BP network with 6 hidden layer nodes. The training of BP network by back propagation involves three stages: the feed forward of the input training pattern, the calculation and

back propagation of error and the adjustment of the weights. The 10th year's shelf life data was chosen to verify the accuracy of BP network. After 88-step iterations, the training has finished. Fig. (2) shows the trend curve of associated error.

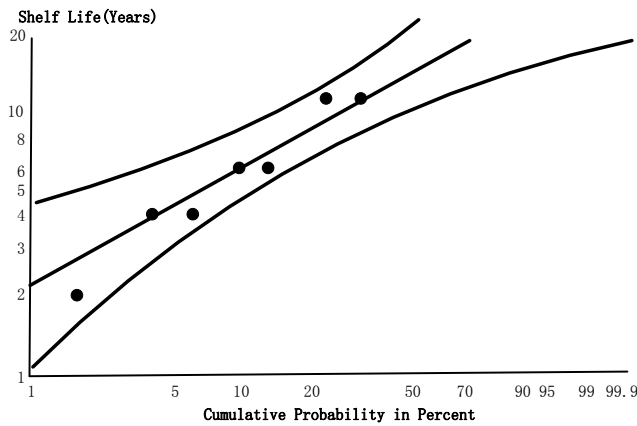


Fig. (1). Shelf Life of silicone brake fluid, MIL-B-46176.

We use the 10th year's shelf life data to verify the trained network, the result is 0.9851 whereas the expect result is nearly 1, and then the error rate is 1.49%. Bi-cubic interpolation algorithm was applied to calculate the 3rd, 5th, 7th, 9th year's shelf life reliable values. We mark these data and training data in asterisk notations, and mark the forecast reliable value in circle notations in Fig. (3).

Fig. (3) shows that prediction errors of 9th and 10th year are -0.79% and -1.51% respectively, of which the errors are larger than others. The result indicates that the 90 percent reliable life is 4.1807 years.

4.1.3. SPFM Model

A Matlab program designed to implement the SPFM mentioned in section 1.3, and Fig. (4) displays the expect

values noted in asterisk notations and the predicted values noted in cycle notations.

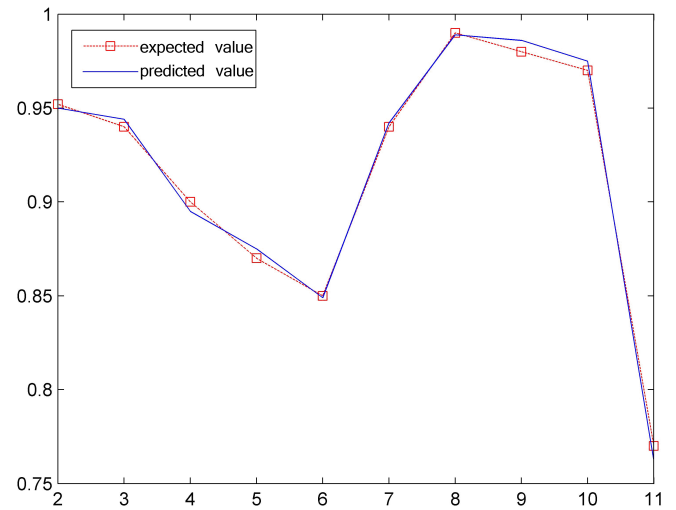


Fig. (3). Comparison chart of forecast value and expect value.

4.2. Combination Forecasts of SVM

RBF was selected for the kernel function of SVM, and a simulation has been conducted by libsvm toolbox developed by Professor Zhi-Ren Lin in Taiwan University. The parameter σ^2 of RBF has a great influence on the generalization ability of the model. Fig. (5) shows the mean absolute percent error (MAPE) when σ^2 range from 1 to 1000.

From the above figure, we can see that insufficient-fitting problem will occur when $\sigma^2 = 1 \sim 275$, and over-fitting problem will occur when $\sigma^2 = 275 \sim 300$. Then we set $\sigma^2 = 275$, hence the selection of error accuracy ϵ and

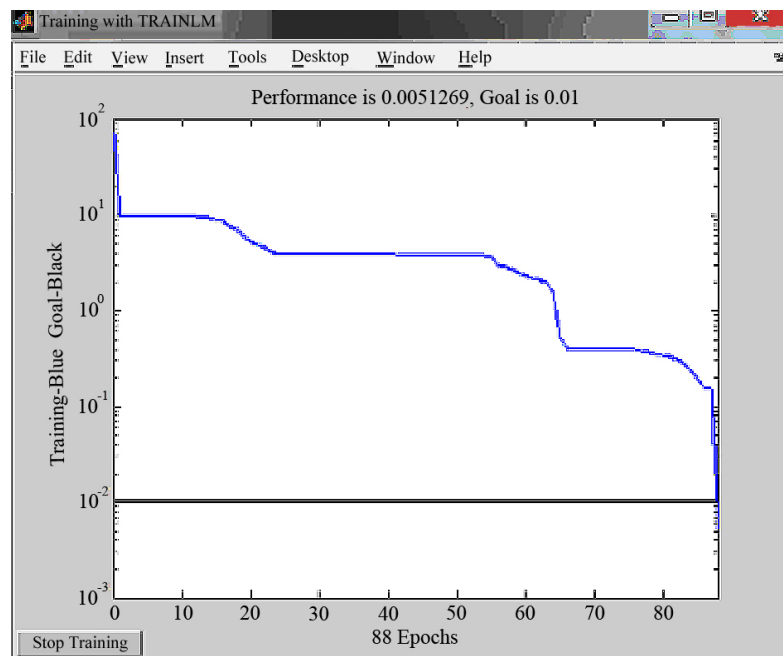


Fig. (2). Tendency chart of BP network training associated error.

penalty function C are based on RBF (275). The $C-MAPE$ and $\epsilon-MAPE$ Curve, see Fig. (6).

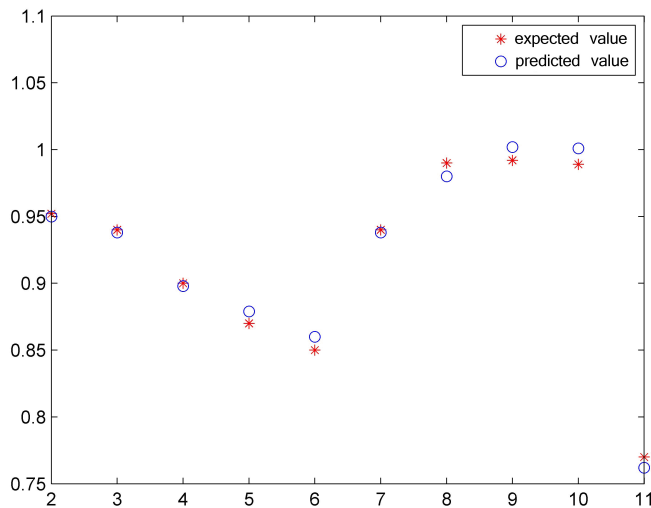


Fig. (4). Forecast results in SPFM.

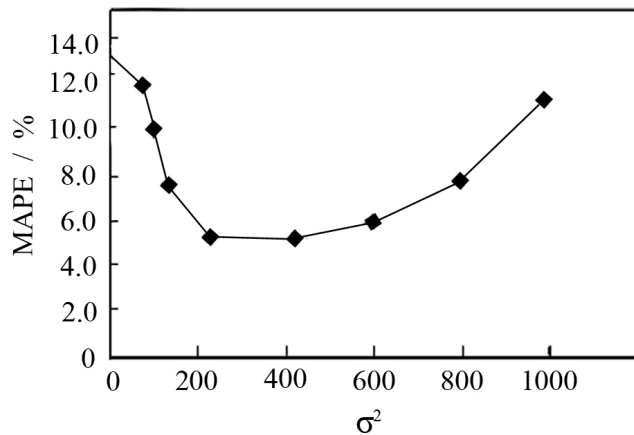


Fig. (5). RBF kernel function $\sigma^2 - MAPE$ curve.

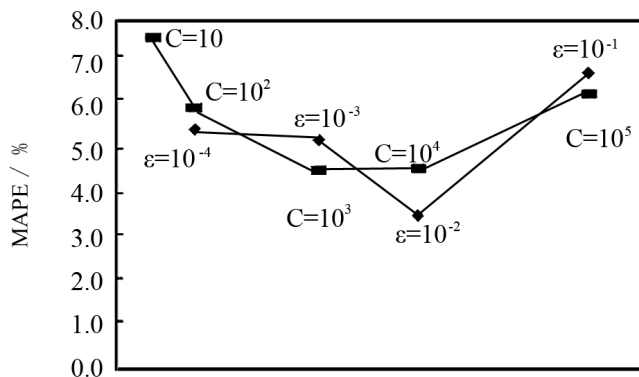


Fig. (6). SVM parameter $C - MAPE$ and $\epsilon - MAPE$ curve.

4.3. Analysis

The prediction results and error analysis of all the three single forecast methods and the combination forecast method are listed in Table 2.

Table 2. Prediction Results and Error Analysis.

Model	Shelf Life	Error/%
Weibull statistic method	6.3	13%
BP network	4.18	1.51%
SPFM	4.2	2.21%
Combination forecasting Model	4.3	1.48%

Comparative analysis suggests that:

1. The result of combination forecast has higher accuracy than other three methods, which proved the feasibility of combination forecast method.
2. When the samples are insufficient, the statistic method will lead to an obvious error.
3. The forecast accuracies of BP network and SPFM are in the same level.

CONCLUSION

The combination forecast model of SVM is indeed an effective measure to forecast the shelf life of two-state materials if the parameters and kernel function are chosen properly. Although the result is better than any other result of single forecast model, the combination forecast model has some deficiency. The accuracy of combination forecast model relies on the accuracy of single forecast models, but the training process of BP network tends to be non-convergent when the samples have a series of sequential 0 values and the predicted results of SPFM allow values that are greater than 100 percent because SPFM does not consider the physical meanings. These problems will be solved in the next study.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] J M Bates, C W J Granger. The combination of forecasts. Oper Res Quart 1969; 20(4): 451-68.
- [2] R T Clemen, R L Winkler. combining economic forecasts. J Busi Econ Stat 1986; 4(1): 39-46.
- [3] C W J Granger, R Ramanathan. Improved methods of combining forecasts. Inter J Fore 1989; 5(4): 559-83.
- [4] G Trenkler, E P Liski. Note: linear constrains and the efficiency of combined forecasts. J Fore 1986; 5(3): 197-202.
- [5] D W Bunn. A Bayesian approach to the linear combination of forecasts. Opera Res Quart 1975; 26(2).
- [6] X W Tang, Z F Zhou, Y Shi. The error bounds of combined forecasting. Math Comp Model 2002; 36: 997-1005.
- [7] H Y Chen, D P Hou. Combination forecasting model based on forecasting effective measure with standard deviate. J Syst Eng 2003; 18(3): 203-10.
- [8] D W Bunn, Combining forecasts. Europ J Oper Res 1988; 33(3): 223-9.

- [9] C W J Granger, Combining forecasts—twenty years later. *J Forec* 1989; 8(3): 167-173.
- [10] K Holder, D A Peel. Unbiasedness, efficiency and the combination of economic forecasts. *J Forec* 1989; 8(3): 175-88.
- [11] J Hallman, M. Kamstra. Combining algorithms based on robust estimation techniques and co-integrating restrictions. *J Forec* 1989; 8(3): 189-98.
- [12] G Anandalingam, L. Chen. Linear Combination of forecasts: a general bayesian model. *J Forec* 1989; 8 (3): 199-214,
- [13] J B Guerard, R T Clemen, Collinearity and the use of latent root regression for combining GNP forecasts. *J Forecasting* 1989; 8(3): 231-8.
- [14] R T Clemen. Combing forecasts. *Europ J Opera Rese.* 1988; 33(3): 223-9.
- [15] K F Lam, H W.Mui, H K Yuen. A note on minimizing absolute percentage error in combined forecasts. *Comp Oper Rese* 2001; 28(11): 1147.
- [16] R R Yager. Families of OWA operators. *Fussy Sets Syst* 1993; 59: 125-48.
- [17] R R Yager. Induced aggregation operators. *Fussy Sets Syst* 2003; 137: 59-69.
- [18] F Herrera, E Herrera-Viedma, F Chiclana. Multiperson decision-making based on multiplicative preference relations. *Europ J Oper Rese* 2001; 129: 372-85.
- [19] A.B.Ji, Y H Qiao, H J Qiu. Combination forecasting of fuzzy forecast. In: *Proceeding of 6th International Conference on Fuzzy Systems and Knowledge Discovery*, Tianjin, 2009, pp. 262-266.
- [20] S Gao. Generalization of linear combining forecasting. *J Cent S Univ.* 2013; 44(9): 178-82.

Received: February 23, 2015

Revised: June 8, 2015

Accepted: June 8, 2015

© Zhiwei *et al.*; Licensee *Bentham Open*

This is an open access article licensed under the terms of the (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.