International Journal of Innovative Computing, Information and Control
Volume **8**, Number **11**, November **2012** ISSN 1349-4198
pp. 7679–7696

A COLLABORATIVE FUZZY-NEURAL SYSTEM FOR GLOBAL CO2 CONCENTRATION FORECASTING

Toly Chen

Department of Industrial Engineering and Systems Management
Feng Chia University
No. 100, Wenhwa Rd., Seatwen, Taichung 40724, Taiwan
tolychen@ms37.hinet.net

Received August 2011; revised December 2011

Abstract. *The global CO2 concentration is considered to be one of the most important causes of global warming that must be closely monitored, accurately forecasted, and controlled as good as possible. To accurately forecast the global CO2 concentration, a collaborative fuzzy-neural system is developed in this study. In the collaborative fuzzy-neural system, instead of calling a number of experts in the field, a committee of virtual experts is formed. These virtual experts are asked to predict the global CO2 concentration based on their local observations, and may not share the raw data they own with each other. A collaboration mechanism is therefore established. For each virtual expert, the corresponding fuzzy back propagation network (FBPN) is constructed to predict the global CO2 concentration, based on the virtual experts' views. To facilitate the collaboration process and to derive a single representative value from these fuzzy forecasts, the concept of fuzzy group learning tree (FGLT) is proposed. Some historical data on global CO2 concentrations were used to evaluate the effectiveness of the collaborative fuzzy-neural system. According to the experimental results, the hit rate, precision, and accuracy of forecasting the global CO2 concentration were considerably improved using the fuzzy collaborative forecasting approach for virtual experts without sharing the raw data they owned.*
Keywords: Collaborative, Fuzzy back propagation network, Forecasting, Global CO2 concentration

1. **Introduction.** Surface atmospheric temperature has the greatest impact for the organisms. Data analysis and forecasting in this area is extremely important. There is more and more evidence showing that there is a widespread and long-term trend toward warmer global temperatures [1]:
(1) Heat waves and periods of unusually warm weather;
(2) Ocean warming, rise in sea-level and coastal flooding;
(3) Melting of glaciers;
(4) Warming of the Arctic and the Antarctic.

The impact of global warming also includes spreading of disease, earlier spring arrival, plant and animal range shifts and population changes, coral reef bleaching, downpours, heavy snowfalls, flooding, droughts and fires [1]. The global CO2 concentration, derived from measurements of CO2 concentration in air bubbles in the layered ice cores drilled in Antarctica and from atmospheric measurements, is considered to be one of the most important causes of global warming, and should be closely monitored, accurately forecasted, and controlled [2]. In addition, the range of the global CO2 concentration is also important, and the narrowest range should be determined so that the global CO2 concentration is neither over-estimated or under-estimated. However, forecasting the global CO2 concentration is not easy. Prentice et al. [3] forecasted the global CO2 concentration for the year 2100 after considering several scenarios, and their result ranged from 541 to 970 ppm.

The purposes of the global CO2 concentration forecasting include:

(1) To generate an accurate forecast of the global CO2 concentration, so that long-term environmental planning can be based on it.
(2) To establish a precise interval of the predicted global CO2 concentration, making it less likely for the government to raise budget unreasonably.
(3) Reducing the risk of energy shortage.

There are two viewpoints to forecasting global CO2 concentrations. The input-output relationship viewpoint, is to determine the economic, sociological, technological, and natural development factors (e.g., fossil fuel burning, deforestation, and land-use change) that influence global CO2 concentration, and then apply different approaches (e.g., multiple linear regression (MLR), and artificial neural network (ANN)) to model the relationship between the global CO2 concentration and these factors, and so forecast the future global CO2 concentration. The time-series viewpoint, is to treat the fluctuation in the global CO2 concentration as a type of time series [1,4,5]. Theoretically, there are many approaches, e.g., moving average (MA), weighted moving average (WMA), exponential smoothing (ES), MLR, ANN, auto-regressive integrated moving average (ARIMA), autoregressive moving average with exogenous variable (ARMAX) models based on genetic algorithms, and others that can be applied to forecast the global CO2 concentration. Generally, an ANN is suitable for modeling a short-term nonlinear pattern of the global CO2 concentration, while traditional approaches such as MA, WMA and ES provide good performances when the trend in the global CO2 concentration is stable. This study belongs to the second category.

This study is dedicated to establish a collaborative fuzzy-neural system for global CO2 concentration forecasting. The motivation of this study is explained as follows. The global CO2 concentration was highly uncertain in the past. In addition, it is difficult to fit the distribution of the global CO2 concentration during a future period, which means that a stochastic approach might not be applicable [4,6]. A fuzzy approach is therefore considered to be useful. Fuzzy approaches have been successfully applied to various fields [7-13]. However, most applications are based on a single fuzzy system. In addition, in the face of global issues, dealing with disparate data sources is becoming more and more popular. Furthermore, due to technical constraints, security or privacy considerations, the integral access to some sources is often limited. For these reasons, Pedrycz [14] proposed the concepts of collaborative computing intelligence and collaborative fuzzy modeling, and developed the so-called fuzzy collaborative system. However, most of the existing fuzzy collaboration systems are used for clustering, i.e., the so-called fuzzy collaborative clustering system [15-17]. The purpose of this study is to develop a collaborative fuzzy-neural system to predict the global CO2 concentration. Fuzzy collaborative intelligence systems for similar purposes have rarely been discussed in the literature [4]. Nevertheless, good forecasts are critical to all aspects of the business. Forecasting is a wide and complex task, both in terms of problem formulation and solution approach in different applications. The global CO2 concentration forecasting is a long-standing research task and we believe that there is still room for further development.

In the collaborative fuzzy-neural system, at first virtual experts (on behalf of automatic data collection devices) on different locations construct their own fuzzy back propagation networks (FBPNs) from various viewpoints to forecast the future global CO2 concentration. These virtual experts are asked to predict the global CO2 concentration based on their local observations, and may not share the raw data they own with each other. A collaboration mechanism is therefore needed. The future global CO2 concentration forecasted with an FBPN is a fuzzy value. Subsequently, every virtual expert conveys its views to others with the aid of a software agent. After receiving the views of others, a virtual expert may be inspired to modify its setting. To derive a single representative value from these global CO2 concentration forecasts, the fuzzy group learning tree (FGLT) approach is employed. The collaboration process is terminated if the improvement in the forecasting performance becomes negligible.

The unique features of the proposed methodology include:

(1) Fuzzy collaborative systems have considerable potential for forecasting, but have not been carefully explored. To the best of our knowledge, no literature on the application of fuzzy collaborative systems to global CO2 concentration forecasting has been published.
(2) In addition, all existing methods assume that analysts have full access to the data. However, although analysts around the world are engaging in the collection of the global CO2 concentration data, they are unable to obtain the comprehensive information.
(3) If an analyst has full access to the data, then it is possible that all actual values fall within the ranges of the fuzzy forecasts. However, such a "robust" property no longer holds under a distributed processing or decision making environment in which a decision maker might have access to only part of the data. To tackle this problem, the forecasting results by all decision makers can be communicated to each other so that they can revise their settings and generate forecasts that are more robust if all the data are taken in account.

The remainder of this paper is organized as follows. Section 2 introduces the proposed collaborative fuzzy-neural system. In Section 3, the historical data of the global CO2 concentration are used to demonstrate the application of the proposed methodology. The accuracy and precision of the proposed methodology are evaluated and compared with those of some existing approaches. Based on the results of the analysis, some points are made. Finally, the concluding remarks and some directions for future research are given in Section 4.

2. Methodology.

We can classify fuzzy collaborative forecasting systems based on three dimensions:

(1) The decentralized or centralized data access.
(2) The use of an agent (virtual expert) or real expert.
(3) The similarities and differences between the fuzzy systems used on different locations (refer to Figure 1).

In total, there are nine categories, and the proposed methodology belongs to the following category: the decentralized data access, agents, and similar fuzzy systems on different locations.

The collaborative fuzzy-neural system consists of several steps that will be described in the following sections:

(1) The collaborative fuzzy-neural system starts from the formation of a committee of virtual experts. These virtual experts are asked to put forward their views on certain aspects of forecasting accuracy and precision.
(2) The views are incorporated into the FBPN approach to generate fuzzy-valued global CO2 concentration forecasts.

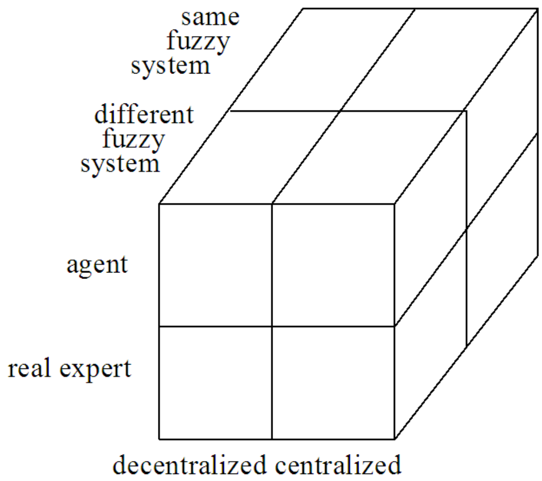

FIGURE 1. Classification of fuzzy collaborative forecasting systems

(3) Every virtual expert conveys its views to other experts with the aid of a software agent. After receiving the views of others, a virtual expert may be inspired to modify its setting.
(4) To arrive at a representative value from the fuzzy global CO2 concentration forecasts, the FGLT approach is employed.
(5) The collaboration process is terminated if the improvement in the forecasting performance becomes negligible. Otherwise, return to step (3).

The system diagram of the proposed methodology is shown in Figure 2. Before introducing the details of the fuzzy collaborative intelligence system, we first define all required parameters as follows:

(1) a_t: the actual global CO2 concentration (after normalization) at period t.
(2) $\tilde{C}_t$: the fuzzy-valued global CO2 concentration forecast at period t. $\tilde{C}_t$ is expressed with a triangular fuzzy number (TFN), i.e., $\tilde{C}_t = (C_{t1}, C_{t2}, C_{t3})$.
(3) C_t: the crisp-valued global CO2 concentration forecast at period t. C_t is derived by defuzzifying $\tilde{C}_t$, i.e., $C_t = D(\tilde{C}_t)$.
(4) $\tilde{o}(t)$: the FBPN output, which is the normalized fuzzy global CO2 concentration forecast at period t, i.e., $\tilde{o}_t = N(\tilde{C}_t)$.
(5) $\tilde{h}_l(t)$: the output from hidden-layer node l.
(6) $\tilde{w}_l^o(t)$: the weight of the connection between hidden-layer node l and the output node.
(7) $\tilde{w}_{il}^h(t)$: the weight of the connection between input node i and hidden-layer node l.
(8) $\tilde{\theta}_l^h(t)$: the threshold for screening out weak signals by hidden-layer node l.
(9) $\tilde{\theta}^o(t)$: the threshold for screening out weak signals by the output node.

2.1. **Step 1: Forming the virtual expert committee.** In the collaborative fuzzy-neural system, a committee composed of a number of virtual experts is formed. These virtual experts can be automatic data collection devices or software agents, and make different points of view on behalf of the real experts on the following aspects:

(1) The range of the fuzzy-valued global CO2 concentration forecast ($\Psi_t(g)$).
(2) The extent that the fuzzified forecast is satisfied with the actual value on the right-hand side ($s_U(g)$) (see Figure 3(a)).
(3) The extent that the fuzzified forecast is satisfied with the actual value on the left-hand side ($s_L(g)$) (see Figure 3(b)).

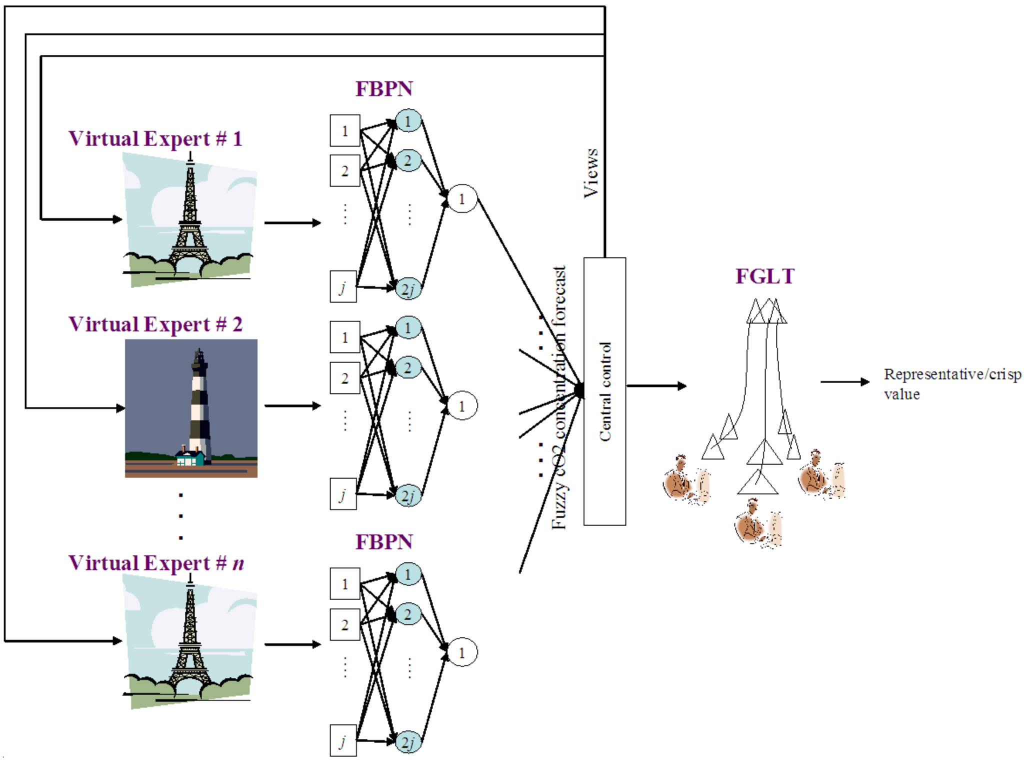

FIGURE 2. The system diagram of the proposed methodology

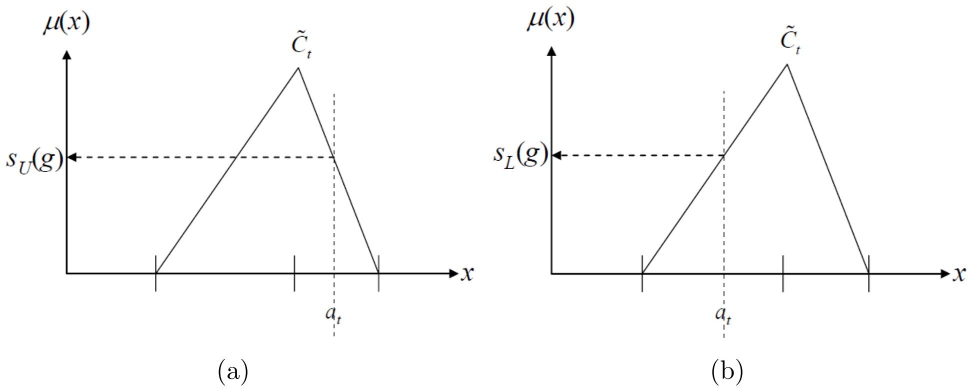

FIGURE 3. The satisfaction levels

These views are incorporated into the FBPN approach to generate fuzzy-valued global CO2 concentration forecasts.

2.2. **Step 2: Forecasting global CO2 concentration with the FBPN approach.** In the collaborative fuzzy-neural system, every virtual expert uses the FBPN approach to predict the future global CO2 concentration with fuzzy values. The theoretical background of the FBPN approach is explained as follows.

First, for each virtual expert, an FBPN is constructed with the following configuration to predict the future global CO2 concentration. Although there have been some more advanced artificial neural networks, such as compositional pattern-producing network, cascading neural network, and dynamic neural network, a well-trained FBPN with an optimized structure can still produce very good results. That is why it is selected in this study:

(1) K inputs, corresponding to the global CO2 concentration K periods ago. To facilitate the search for solutions, it is strongly recommended to normalize the inputs to a range narrower than [0 1] [18].
(2) The FBPN has only one hidden layer. Two or more hidden layers slow down the convergence speed, and may not lead to any better solution. The number of nodes in the hidden layer is chosen from 1 to $2K$ after trying each of them.
(3) The output from the FBPN is the normalized fuzzy global CO2 concentration forecast.
(4) The activation function used for the hidden layer is the hyperbolic tangent sigmoid function, and for the others is the linear activation function.
(5) 70000 epochs will be run each time. The start conditions will be randomized to reduce the possibility of being stuck on local optima.

The training of the FBPN is decomposed into three subtasks: determining the core value, upper, and lower bounds of the parameters. First, to determine the core of each parameter (such as $w_{il2}^h(t)$, $\theta_{l2}^h(t)$, $w_{l2}^o(t)$ and $\theta_2^o(t)$), some advanced algorithms are applicable. For details see Eraslan [19]. In the fuzzy collaborative intelligence system, we choose the CGF algorithm. At every new step, the search direction will be conjugated to all previous ones, which eliminates the need to calculate the Hessian. An example is given in Figure 4. Obviously, an FBPN provides a very good fit of the data.

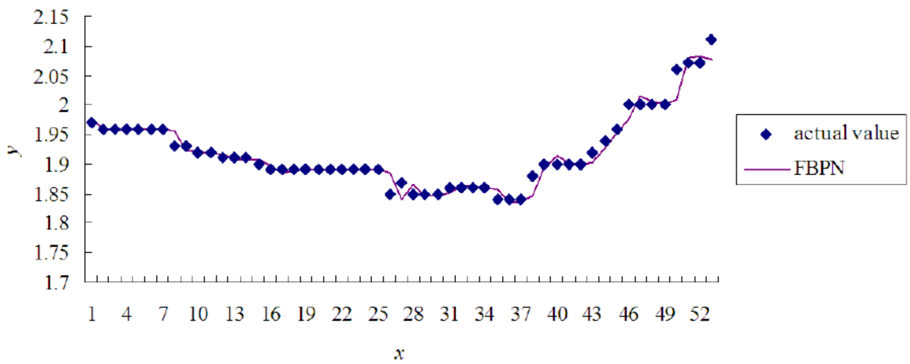

Figure 4. The good fit by using the FBPN approach

Subsequently, the following goal programming (GP) problem is solved to determine the upper bound of each parameter (e.g., $w_{il3}^h(t)$, $\theta_{l3}^h(t)$, $w_{l3}^o(t)$ and $\theta_3^o(t)$) [20].

(GP I)

$$\text{Min} \quad \sum_{\text{all } t} \Psi_t(g) \tag{1}$$

s.t.

$$\ln\left(\frac{1}{C_{t3}} - 1\right) = \theta_3^o(t) - \sum_{\text{all } l} w_{l3}^o(t) h'_{l3}(t), \tag{2}$$

$$\sum_{\text{all } l} w_{l3}^o(t) h'_{l3}(t) - \theta_3^o(t) \leq -\ln(1/\Psi_t(g) - 1), \tag{3}$$

$$\sum_{\text{all } l} w_{l3}^o(t) h'_{l3}(t) \leq \theta_3^o(t) - \ln\left(\frac{1 - s_U(g)}{a_t - s_U(g) C_{t2}} - 1\right), \tag{4}$$

$$\sum_{\text{all } l} w_{l3}^o(t)h_{l3}'(t) \geq \theta_3^o(t) - \ln\left(\frac{1}{a_t} - 1\right), \tag{5}$$

$$\sum_{\text{all } i} w_{il3}^h(t)x_i(t) - \theta_{l3}^h(t) \geq -\ln(1/h_{l3}'(t) - 1), \tag{6}$$

$$\sum_{\text{all } i} w_{il3}^h(t)x_i(t) - \theta_{l3}^h(t) \leq -\ln(1/h_{l3}'(t) - 1), \tag{7}$$

$$i = 1 \sim K, \tag{8}$$

$$l = 1 \sim m \text{ (the number of hidden-layer nodes)}. \tag{9}$$

After enumerating a number of possible values for them, the goal programming problem is solved many times. In these optimization results, the best one giving the minimum upper bound is chosen.

In a similar way, the following GP problem is solved to determine the lower bound of each parameter (e.g., $w_{il1}^h(t)$, $\theta_{l1}^h(t)$, $w_{l1}^o(t)$ and $\theta_1^o(t)$):

(GP II)

$$\text{Min} \quad \sum_{\text{all } t} \Psi_t(g) \tag{10}$$

s.t.

$$\ln\left(\frac{1}{C_{t1}} - 1\right) = \theta_1^o(t) - \sum_{\text{all } l} w_{l1}^o(t)h_{l1}'(t), \tag{11}$$

$$\sum_{\text{all } l} w_{l1}^o(t)h_{l1}'(t) - \theta_1^o(t) \leq -\ln(1/\Psi_t - 1), \tag{12}$$

$$\sum_{\text{all } l} w_{l1}^o(t)h_{l1}'(t) \leq \theta_1^o(t) - \ln\left(\frac{1 - s_L(g)}{a_t - s_L(g)C_{t2}} - 1\right), \tag{13}$$

$$\sum_{\text{all } l} w_{l1}^o(t)h_{l1}'(t) \leq \theta_1^o(t) - \ln\left(\frac{1}{a_t} - 1\right), \tag{14}$$

$$\sum_{\text{all } i} w_{il1}^h(t)x_i(t) - \theta_{l1}^h(t) \geq -\ln(1/h_{l1}'(t) - 1), \tag{15}$$

$$\sum_{\text{all } i} w_{il1}^h(t)x_i(t) - \theta_{l1}^h(t) \leq -\ln(1/h_{l1}'(t) - 1), \tag{16}$$

$$i = 1 \sim K, \tag{17}$$

$$l = 1 \sim m \text{ (the number of hidden-layer nodes)}. \tag{18}$$

As a result, all actual values fall within the scopes of the fuzzy forecasts. If we have the integral access to the data, then the forecasting results by the FBPN approach are as shown in Figure 5. However, such a "robust" property no longer holds under a distributed processing or decision making environment in which a decision maker may only have access to part of the data, and the forecasting results by the FBPN approach will be as shown in Figure 6. There are some outliers (indicated with red O) in Figure 6. To solve this problem, the forecasting results by all decision makers can be communicated to each other, so that they can modify their settings, and generate more robust forecasts as if all data are taken into account. To this end, the GP problems are modified so that a collaborative formulation can be made in the next section.

2.3. **Step 3: Collaboration of virtual experts.** The view of a virtual expert is indicated with $VS_g = \{\Psi_t(g), s_R(g), s_L(g)\}$, $g \in [1\ G]$, and will be packaged into information

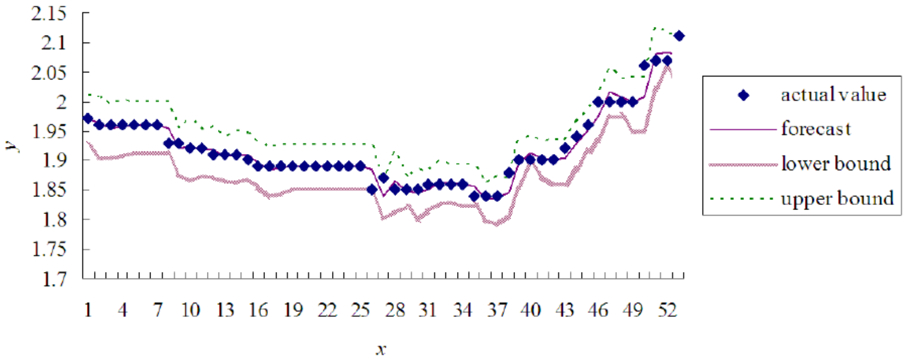

FIGURE 5. The upper and lower bounds established by the GP models

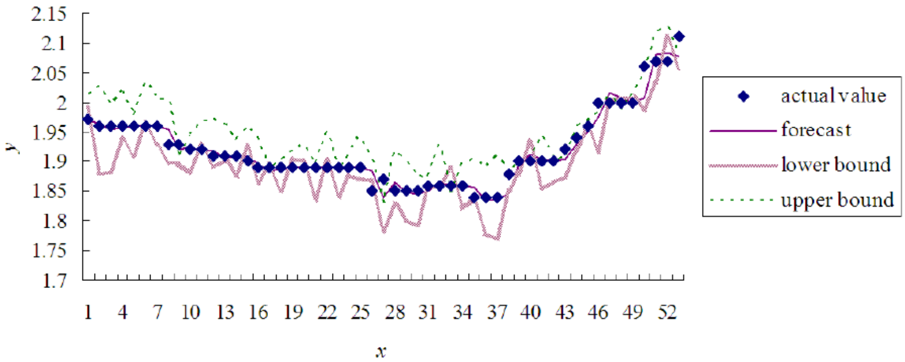

FIGURE 6. The forecasting results by the FBPN approach (with partial data access)

granules, which are then encoded using extensible markup language (XML). Subsequently, a software agent is used to convey information granules among virtual experts through a centralized point-to-point (P2P) architecture. The communication protocol is as follows:

Input: Expert E_g, $1 \leq g \leq G$, provides input data $\tilde{C}_t$ for T periods, where $n \leq t \leq T + n - 1$. In case of computing the FBPN output, the view vector VS_g is public.

Output: Expert E_g, $1 \leq g \leq G$, learns $\dfrac{D(\tilde{C}_t) - a_t}{a_t}$ without anything else, where $D(\widetilde{C}_t)$ is computed using the center-of-gravity method. After receiving this information, if it reveals that the forecasting performance of a virtual expert is very prominent, the others may change their settings, so that their views will move closer:

(1) Favoring mechanism: During the collaborative global CO2 concentration forecasting, some virtual experts favor the views of other virtual experts, and modify their earlier views to be as close as possible. To facilitate such actions, a favoring mechanism is designed. A favoring mechanism helps to achieve the consensus in a number of virtual experts in order to avoid losing the real global CO2 concentration (see Figure 7).

(2) Disfavoring mechanism: On the contrary, some virtual experts disfavor the views of other virtual experts, and modify their views to be as far as possible. To facilitate such actions, a disfavoring mechanism is designed. If the consensus of virtual experts is already high, then a disfavoring mechanism helps reduce the area in search of the actual global CO2 concentration, to improve the precision and accuracy of forecasts (see Figure 8).

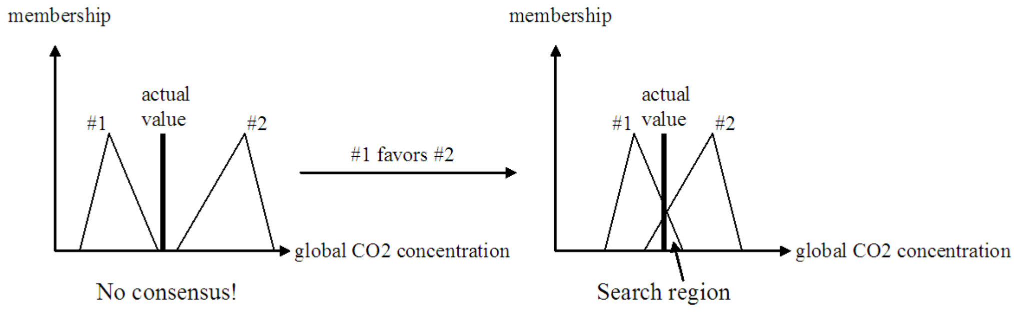

FIGURE 7. Effects of a favoring mechanism

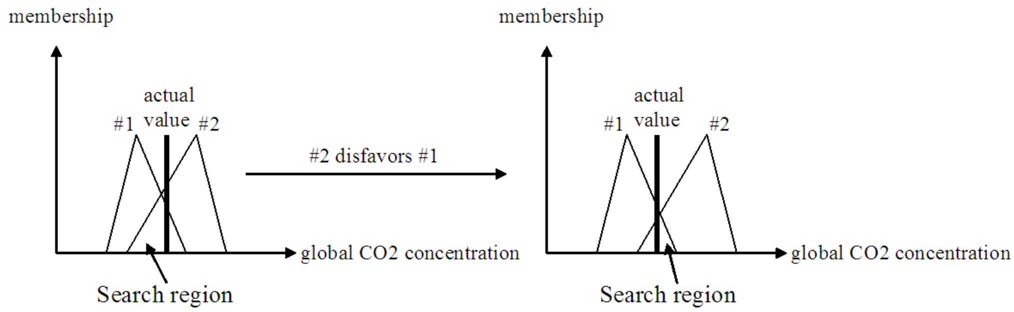

FIGURE 8. Effects of a disfavoring mechanism

After communication virtual expert g refits the corresponding FBPN with the following GP models:
(GP III)

$$\text{Min} \sum_{\text{all } t} \Psi_t(gg), \tag{19}$$

s.t.

$$\ln\left(\frac{1}{C_{t3}} - 1\right) = \theta_3^o(t) - \sum_{\text{all } l} w_{l3}^o(t) h_{l3}'(t), \tag{20}$$

$$\sum_{\text{all } l} w_{l3}^o(t) h_{l3}'(t) - \theta_3^o(t) \le \max_{qq \in t^c(gg)} (-\ln(1/\Psi_t(qq) - 1)), \tag{21}$$

$$\sum_{\text{all } l} w_{l3}^o(t) h_{l3}'(t) - \theta_3^o(t) \le \max_{qq \in t^c(gg)} \left(-\ln\left(\frac{1 - s_U(qq)}{a_t - s_U(qq) C_{t2}} - 1\right)\right), \tag{22}$$

$$\sum_{\text{all } l} w_{l3}^o(t) h_{l3}'(t) \ge \theta_3^o(t) - \ln\left(\frac{1}{a_t} - 1\right), \tag{23}$$

$$\sum_{\text{all } i} w_{il3}^h(t) x_i(t) - \theta_{l3}^h(t) \ge -\ln(1/h_{l3}'(t) - 1), \tag{24}$$

$$\sum_{\text{all } i} w_{il3}^h(t) x_i(t) - \theta_{l3}^h(t) \le -\ln(1/h_{l3}'(t) - 1), \tag{25}$$

$$i = 1 \sim K, \tag{26}$$

$$l = 1 \sim m \text{ (the number of hidden-layer nodes)}, \tag{27}$$

(GP IV)

$$\text{Min} \sum_{\text{all } t} \Psi_t(gg) \tag{28}$$

s.t.

$$\ln\left(\frac{1}{C_{t1}} - 1\right) = \theta_1^o(t) - \sum_{\text{all } l} w_{l1}^o(t) h'_{l1}(t), \tag{29}$$

$$\sum_{\text{all } l} w_{l1}^o(t) h'_{l1}(t) - \theta_1^o(t) \le \max_{qq \in t^c(gg)} (-\ln(1/\Psi_t(qq) - 1)), \tag{30}$$

$$\sum_{\text{all } l} w_{l1}^o(t) h'_{l1}(t) - \theta_1^o(t) \le \max_{qq \in t^c(gg)} \left(-\ln\left(\frac{1 - s_L(qq)}{a_t - s_L(qq) C_{t2}} - 1\right)\right), \tag{31}$$

$$\sum_{\text{all } l} w_{l1}^o(t) h'_{l1}(t) \le \theta_1^o(t) - \ln\left(\frac{1}{a_t} - 1\right), \tag{32}$$

$$\sum_{\text{all } i} w_{il1}^h(t) x_i(t) - \theta_{l1}^h(t) \ge -\ln(1/h'_{l1}(t) - 1), \tag{33}$$

$$\sum_{\text{all } i} w_{il1}^h(t) x_i(t) - \theta_{l1}^h(t) \le -\ln(1/h'_{l1}(t) - 1), \tag{34}$$

$$i = 1 \sim K, \tag{35}$$

$$l = 1 \sim m \text{ (the number of hidden-layer nodes)}, \tag{36}$$

where $VS_{qq} = \{\Psi_t(qq), s_R(qq), s_L(qq)\}$ is the view of virtual expert q and so on, following the notations suggested by Pedrycz [14]; $t(qq)$ includes the time indexes of all data in the part assessed by virtual expert q. $t^c(qq)$ is the complement of $t(qq)$, i.e., $t^c(qq) = [1\ T] - t(qq)$; $s(ii)$ is the satisfaction level requested by virtual expert i. Constraint (21) and (22) force the upper bound on the fuzzy forecast to be greater than those made by other virtual experts for a period the datum is lacking. On the contrary, in (30) and (31), the lower bound should be less than those made by other virtual experts for the same period.

If there are L virtual experts, then after incorporating the virtual experts' opinions into the FBPN approach there will be at most $2L$ GP problems to be solved. After solving every two GP problems, the optimal solution is used to construct a corresponding FBPN. Eventually, there will be at most L FBPNs, each of which generates a fuzzy global CO2 concentration forecast. A mechanism is therefore required to combine these fuzzy global CO2 concentration forecasts. In the proposed methodology, the concept of FGLT is proposed for the global CO2 concentration forecasting.

2.4. **Step 4: Deriving a representative value with FGLT.** In the existing methods, deriving a representative value from a group of fuzzy forecasts consists of two tasks: aggregation and defuzzification. This process may be time-consuming in the absence of consensus, which requires a lot of time for them to adjust their settings to gradually close to each other. To solve this problem, FGLT is employed in the collaborative fuzzy-neural system.

The concept of FGLT is natural. Because these virtual experts will eventually reach a consensus, we can treat it as a learning process, which will converge to the final value. If the final value can be estimated somehow, then you do not need to go another round of communication. Starting from the initial set up around the root, the FGLT places the final values on the top, indicating that these experts have reached a consensus.

In FGLT, the changes in the setting of a virtual expert follow a learning process:

$$VS_g(u) = FVS_g \cdot e^{-\frac{b_g}{u}+r(u)}, \tag{37}$$

where $VS_g(u)$ is the view of virtual expert g after the u-th round of communication; FVS_g is the final value; b_g is the learning speed; $r(u)$ is a homoscedastical, serially non-correlated error term. Fitting historical data to obtain the parameters in (37) is the conventional approach. In logarithms,

$$\ln(VS_g(u)) = \ln FVS_g - \frac{b_g}{u} + r(u), \tag{38}$$

which can be solved by using the simple, linear regression, under the assumption of

$$r(u) \sim \text{Normal } (0, \sigma^2),\ \hat{r}(u) = 0, \tag{39}$$

where "ˆ" denotes that it is an estimate. We can obtain

$$\hat{b}_g = -\frac{S_{xy}}{S_{xx}}, \quad \widehat{FVS_g} = e^{\frac{1}{U}\sum\limits_{u=1}^{U}\ln(VS_g(u))+\frac{\hat{b}_g}{U}\sum\limits_{u=1}^{U}\frac{1}{u}}, \tag{40}$$

where

$$\begin{aligned} S_{xx} &= \sum_{u=1}^{U}\frac{1}{u^2} - \frac{\left(\sum\limits_{u=1}^{U}\frac{1}{u}\right)^2}{U}, \\ S_{xy} &= \sum_{u=1}^{U}\frac{\ln(VS_g(u))}{u} - \frac{1}{U}\left(\sum_{u=1}^{U}\frac{1}{u}\right)\left(\sum_{u=1}^{U}\ln(VS_g(u))\right). \end{aligned} \tag{41}$$

If the final values can be reliably estimated, then the communication is terminated:

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} \text{ (from 0 (not reliable) to 1 (reliable))}, \tag{42}$$

where

$$S_{yy} = \sum_{u=1}^{U}\ln(VS_g(u))^2 - \frac{\left(\sum\limits_{u=1}^{U}\ln(VS_g(u))\right)^2}{U}. \tag{43}$$

The convergence criteria for the FGLT is

$$\sup_{all\ g}\min_{w=1\sim3} R^2(g,w) \geq \varphi, \tag{44}$$

where $R^2(g,w)$ is the coefficient of determination of virtual expert g along the w-th dimension; $\varphi \in [0\ 1]$ is a threshold. Since then, the corners of the forecasts by all virtual experts are fed into a multiple linear regression equation to derive the representative value:

$$C_t = w_1C_{t1}(1) + w_2\mu_{C_{t1}} + w_3C_{t2}(1) + w_4\mu_{C_{t2}} \dots w_{6G-1}C_{t3}(G) + w_{6G}\mu_{C_{t3}}, \tag{45}$$

where $\tilde{C}_t(g) = (C_{t1}(g), C_{t2}(g), C_{t3}(g))$ is the fuzzy global CO2 concentration forecast by virtual expert g; $g = 1 \sim G$. Equation (45) can be fitted by the traditional way of minimizing the sum of squared error.

An example is given in Table 1. There are three virtual experts. Because these virtual experts will eventually reach a consensus, we can estimate the final views based on the first few values. As a result, they do not need to go another round of communication, and the collaboration process can be shortened. Subsequently, these final views are fed into the corresponding FBPN approaches to generate fuzzy global CO2 concentration forecasts, as shown in Table 2. For each period, the three corners of the fuzzy forecast by each virtual expert are recorded, and used to construct a multiple linear regression equation to derive the representative value that can be directly compared with the actual value (see Table 3).

Table 1. The views of some virtual experts

Virtual Expert #	View #1	View #2	...	Final View	Estimated Final View
1	(10, 0.3, 0.4)	(9, 0.25, 0.4)	...	?	(9.3, 0.26, 0.38)
2	(8, 0.2, 0.5)	(8.5, 0.25, 0.45)	...	?	(8.6, 0.27, 0.43)
3	(13, 0.45, 0.45)	(13, 0.5, 0.4)	...	?	(12.8, 0.45, 0.43)

Table 2. The fuzzy forecasts by using the FBPN approaches

Virtual Expert#	Period #1	Period #2	...	Period #12
1	(350, 353, 358)	(351, 355, 359)	...	(362, 367, 370)
2	(346, 350, 353)	(347, 351, 354)	...	(358, 364, 365)
3	(347, 349, 354)	(348, 350, 355)	...	(359, 364, 366)

Table 3. The corners of the fuzzy forecasts and the multiple linear regression equation

Period #	Corners	Equation
1	(350, 0), (353, 1), (358, 0), (346, 0), (350, 1), (353, 0), (347, 0), (349, 1), (354, 0)	$C_1 = w_1 \cdot 350 + w_2 \cdot 0 + w_3 \cdot 353 + \ldots + w_{17} \cdot 349 + w_{18} \cdot 0$
2	(351, 0), (355, 1), (359, 0), (347, 0), (351, 1), (354, 0), (348, 0), (350, 1), (354, 0)	$C_2 = w_1 \cdot 351 + w_2 \cdot 0 + w_3 \cdot 355 + \ldots + w_{17} \cdot 354 + w_{18} \cdot 0$
...		...
12	(362, 0), (367, 1), (370, 0), (358, 0), (364, 1), (365, 0), (359, 0), (364, 1), (366, 0)	$C_{12} = w_1 \cdot 362 + w_2 \cdot 0 + w_3 \cdot 367 + \ldots + w_{17} \cdot 366 + w_{18} \cdot 0$

3. **Experiment and Analyses.** To demonstrate the application of the proposed methodology, the real data of the global CO2 concentration from 2005∼2009 were used [1] (see Figure 9). The raw data contained 14 columns and 60 rows. There was obvious some seasonality in the original data, which has been removed.

Three virtual experts were asked to forecast the global CO2 concentration based on their local observations. However, these virtual experts may not share the data they own with each other. To implement the FBPN approach, there were six GP problems to be solved. From the optimization result of every two GP problems, a corresponding FBPN could be constructed. All three FBPNs were applied to forecast the global CO2

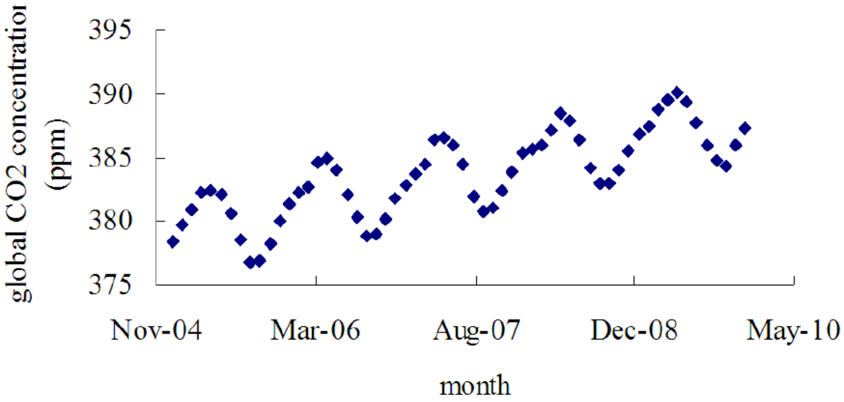

FIGURE 9. The original global CO2 concentration data

concentration. Finally, FGLT was applied to derive the representative value, i.e., the crisp global CO2 concentration forecast, from these fuzzy forecasts.

The proposed methodology was implemented on a PC with an Intel Dual CPU E2200 2.2 GHz and 2.0G RAM. The FBPN was implemented with the Neural Network Toolbox of MATLAB 2006a with the following conditions:

(1) Number of epochs per replication: 60000.
(2) Number of initial conditions/replications: 1000.
(3) Stop training if MSE $< 10^{-5}$ or 60000 epochs have been run.

Lingo 8.0 was applied to solve the GP problems with the following conditions:

(1) The initial opinions of the three virtual experts are randomly generated using the linear congruential generator (LCG).
(2) Successive linear programming (SLP) was used to compute new search directions. Namely, a linear approximation was used in search computation to speed iteration times.
(3) The computation method used for computing derivatives was the numerical method. Numerical derivatives were computed using finite differences.

The forecasting accuracy was measured with RMSE, MAE, and MAPE.

At the same time, the forecasting precision of a non-biased crisp approach can be measured with 6σ (or 6 RMSE), because theoretically the probability that such an interval contains the actual value is about 99.7%, based on the assumption that residuals follow a normal distribution. The precision of a fuzzy approach can be measured with the average spread (or range) of fuzzy global CO2 concentration forecasts if all of them contain the actual value.

To make a comparison with some existing approaches, MA, ES, BPN, ARIMA, Tanaka's FLR approach, and Peters' FLR approach were applied to the collected data. Each virtual expert applied each of the seven ways to predict the global CO2 concentration based on its local data access, and only the best of them is presented here. In MA, various numbers of moving periods (from 10 to 2 step 1) were tried. Among them, the best one was chosen for the subsequent analysis. The number of inputs in BPN was determined in a similar way. In ES, the value of the smoothing constant changed from 0 to 1 with an interval of 0.1, and then the value that contributed to the best accuracy was adopted. In ARIMA, there are three stages: model identification, model estimation, and model checking. The minimum information criterion (MINIC) method was used to identify the order in the ARIMA process. The augmented dickey fuller (ADF) unit root tests were used to test the stationarity and seasonal stationarity in the global CO2 concentration data. In the

BPN approach, there was one hidden layer with nodes twice as many as that of the inputs. The number of epochs was set to 60 000. In addition, the initial values of all parameters in the BPN were randomly generated 100 times. Among them, the best setting was kept for the subsequent analysis. A 4-fold evaluation process was applied to cross-validate the data. In the FLR approaches of Tanaka and Peters, a fuzzy global CO2 concentration forecast is defuzzified with the center-of-gravity (COG) formula:

$$d(\widetilde{W}_t) = \frac{W_{t,1} + W_{t,2} + W_{t,3}}{3}. \tag{46}$$

The accuracy and precision achieved by applying these approaches were recorded and compared in Figures 10-13.

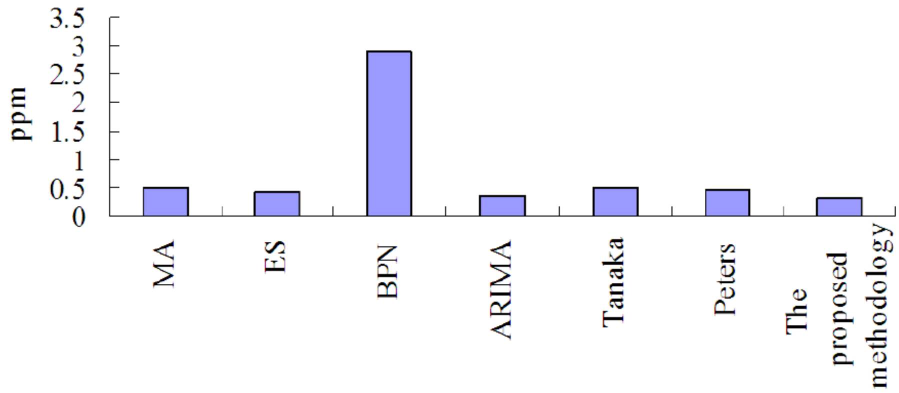

FIGURE 10. The accuracy of various approaches (RMSE)

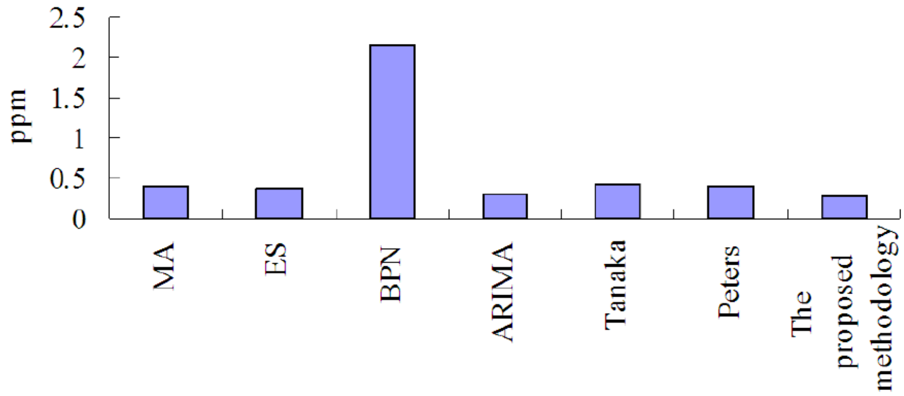

FIGURE 11. The accuracy of various approaches (MAE)

MA was adopted as the comparison basis. The experiment provided the following results,

(1) The accuracy of global CO2 concentration forecasting, measured in terms of RMSE, of the proposed approach, was significantly better than those of the traditional approaches by achieving a 28% reduction in RMSE over the comparison basis – MA. The advantages over ES, BPN and ARIMA were 18%, 87% and 3%, respectively. The accuracy of the proposed methodology with respect to MAE or MAPE was also significantly better than those of the other approaches.

(2) ES outperformed MA because there was an obvious (upward) trend in the collected data. The accuracy of the nonlinear approach BPN was poor because it could overreact when there was a linear trend in the data.

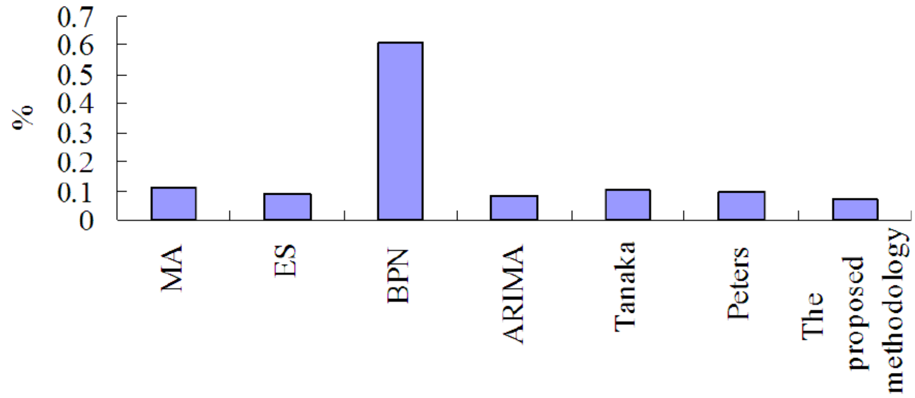

FIGURE 12. The accuracy of various approaches (MAPE)

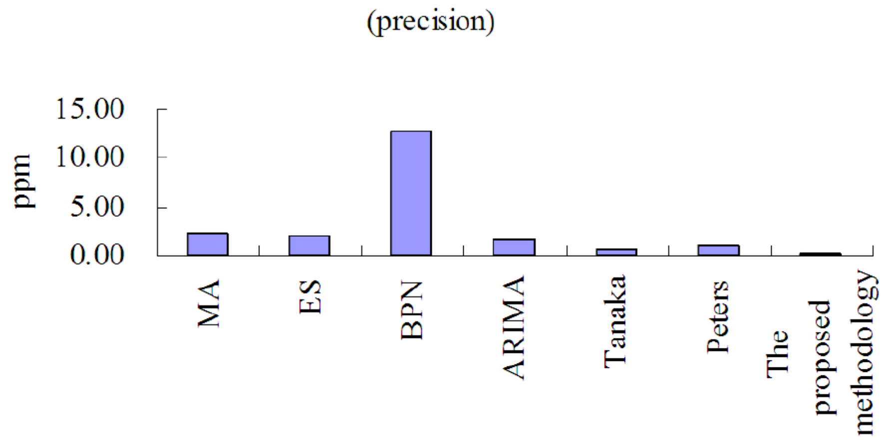

FIGURE 13. The precision of various approaches (the average range)

(3) Among the fuzzy approaches, the proposed methodology surpassed the two traditional FLR approaches by improving the forecasting accuracy. Nevertheless, Tanaka's and Peters' FLR approaches still achieved fair performances in this respect.
(4) As expected, the forecasting accuracy of ARIMA was also very good and quite close to that of the proposed methodology.
(5) On the other hand, the precision of the proposed methodology was significantly better than those of the other approaches. The improvement over the baseline approach (MA) was up to 91%. The advantages over the other approaches ranged from 64% to 98%. In other words, with the proposed methodology, it is possible to come up with a very small range for the future global CO2 concentrations. This will make it possible to determine the lowest upper bound for the global CO2 concentration.
(6) To further illustrate the effectiveness of the proposed methodology, the variance accounted for (VAF) was also calculated:

$$\mathrm{VAF} = \left[1 - \frac{\mathrm{var}(Y_t - A_t)}{\mathrm{var}(A_t)}\right] \cdot 100\%. \tag{47}$$

where var() is the variance function. If VAF is 100% and RMSE is 0, the model is considered to be excellent. The result was 99%, which supported the proposed methodology as a good prediction model of the global CO2 concentration.

We noticed that the hit rate, accuracy, and precision of the collaborative fuzzy-neural system were better than those without collaboration. In addition, after collaboration, it was possible to maximize the hit rate with a very narrow average range of the fuzzy forecasts. Furthermore, compared with the results without collaboration, the aggregation result was more precise. It was not easy since the three virtual experts did not share the raw data they owned with each other.

4. **Conclusion and Directions for Future Research.** Due to technical constraints, security or privacy considerations, the integral access to the global CO2 concentration data is often limited. As a result, it becomes extremely difficult to accurately predict the global CO2 concentration. Forecasting the future global CO2 concentration is a crucial task. To this end, a collaborative fuzzy-neural approach was proposed in this study. In the proposed methodology, multiple virtual experts construct their own FBPNs from various viewpoints to forecast future global CO2 concentrations. These virtual experts are asked to predict the global CO2 concentration based on their local observations, and may not share the data they own with each other. A collaboration mechanism is therefore established. Each FBPN can be fitted with two equivalent GP problems to be solved. Subsequently, rather than the raw data, the forecasting results by a virtual expert are conveyed to others to modify their settings, so that the actual values will be contained in the fuzzy forecasts after collaboration. A communication protocol is also designed to achieve this purpose. The received information is then incorporated into the GP model to fit the FBPN. To combine the fuzzy global CO2 concentration forecasts by different experts, FGLT is applied.

A practical case containing the historical data of the global CO2 concentration was used to demonstrate the applicability of the proposed methodology. The results of the experiment were as follows:

(1) According to the experimental results, the hit rate, precision, and accuracy of forecasting the global CO2 concentration were considerably improved using the fuzzy collaborative forecasting approach for virtual experts without sharing the raw data they owned.
(2) It is possible to determine the lowest upper bound on the global CO2 concentration by using the proposed methodology.

The advantages of the proposed methodology over the existing approaches include:

(1) Unlike the existing approaches, the proposed methodology forecasts the global CO2 concentration in a collaborative way, which conforms to the practical operations in this field.
(2) A sufficient number of domain experts are no longer needed for group forecasting. In this way, there can be significant cost savings and convenience.
(3) Compared with approaches that also use domain experts (e.g., [21-26]), the proposed methodology improves the precision further, and reduces the risk of environmental planning.

Conversely, the possible deficiencies of the proposed methodology include:

(1) Compared with some existing approaches, more data are required with the proposed methodology due to the need to incorporate expert opinions.
(2) It takes more time to implement the collaboration process for the proposed methodology than for the other approaches.

More sophisticated global CO2 concentration forecasting approaches [27] can be developed in similar ways in future studies.

Acknowledgment. This work is partially supported by the National Science Council of Taiwan.

REFERENCES

[1] Union of Concerned Scientists, Global warming: Early warning signs, *http://www.climatehotmap.org/index.html*.

[2] National Assessment Synthesis Team, *Climate Change Impacts on the United States: The Potential Consequences of Climate Variability and Change*, US Global Change Research Program, Washington DC, 2000.

[3] I. C. Prentice, G. D. Farquhar, M. J. R. Fasham, M. L. Goulden, M. Heimann, V. J. Jaramillo, H. S. Kheshgi, C. LeQuéré, R. J. Scholes and D. W. R. Wallace, The carbon cycle and atmospheric carbon dioxide, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the 3rd Assessment Report of the Intergovernmental Panel on Climate Change*, J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell and C. A. Johnson (eds.), New York, Cambridge University Press, 2001.

[4] T. Chen and Y. C. Wang, A fuzzy-neural approach for global CO2 concentration forecasting, *Intelligent Data Analysis*, vol.15, no.5, pp.763-777, 2011.

[5] CO2now.org, Current data for atmospheric CO2, *http://co2now.org/index.php?option=com_content&task=view&id=22&Itemid=1*.

[6] I. S. Cheng, Y. Tsujimura, M. Gen and T. Tozawa, An efficient approach for large scale project planning based on fuzzy Delphi method, *Fuzzy Sets and Systems*, vol.76, pp.277-288, 1995.

[7] T. Chen, An intelligent hybrid system for wafer lot output time prediction, *Advanced Engineering Informatics*, vol.21, pp.55-65, 2007.

[8] T. Chen, A. Jeang and Y. C. Wang, A hybrid neural network and selective allowance approach for internal due date assignment in a wafer fabrication plant, *International Journal of Advanced Manufacturing Technology*, vol.36, pp.570-581, 2008.

[9] T. Chen, A SOM-FBPN-ensemble approach with error feedback to adjust classification for wafer-lot completion time prediction, *International Journal of Advanced Manufacturing Technology*, vol.37, pp.782-792, 2008.

[10] T. Chen, Incorporating fuzzy c-means and a back-propagation network ensemble to job completion time prediction in a semiconductor fabrication factory, *Fuzzy Sets and Systems*, vol.158, pp.2153-2168, 2007.

[11] T. Chen, An intelligent mechanism for lot output time prediction and achievability evaluation in a wafer fab, *Computers and Industrial Engineering*, vol.54, pp.77-94, 2008.

[12] T. Chen, A hybrid fuzzy and neural approach for DRAM price forecasting, *Computers in Industry*, vol.62, pp.196-204, 2011.

[13] T. Chen and Y. C. Lin, A fuzzy back-propagation-network ensemble with example classification for lot output time prediction in a wafer fab, *Applied Soft Computing*, vol.9, no.2, pp.658-666, 2009.

[14] W. Pedrycz, Collaborative architectures of fuzzy modeling, *Lecture Notes in Computer Science*, vol.5050, pp.117-139, 2008.

[15] R. Falcón, B. Depaire, K. Vanhoof and A. Abraham, Towards a suitable reconciliation of the findings in collaborative fuzzy clustering, *Proc. of the 8th International Conference on Intelligent Systems Design and Applications*, 2008.

[16] V. Loia, W. Pedrycz and S. Senatore, P-FCM: A proximity-based fuzzy clustering for user-centered web applications, *International Journal of Approximate Reasoning*, vol.34, pp.121-144, 2003.

[17] H. Selim, C. Araz and I. Ozkarahan, Collaborative production-distribution planning in supply chain: A fuzzy goal programming approach, *Transportation Research Part E: Logistics and Transportation Review*, vol.44, no.3, pp.396-419, 2008.

[18] T. Chen, Y. C. Wang and H. R. Tsai, Lot cycle time prediction in a ramping-up semiconductor manufacturing factory with a SOM-FBPN-ensemble approach with multiple buckets and partial normalization, *International Journal of Advanced Manufacturing Technology*, vol.42, no.11-12, pp.1206-1216, 2009.

[19] E. Eraslan, The estimation of product standard time by artificial neural networks in the molding industry, *Mathematics Problems in Engineering*, 2009.

[20] T. Chen and Y.-C. Wang, Incorporating the FCM-BPN approach with nonlinear programming for internal due date assignment in a wafer fabrication plant, *Robotics and Computer Integrated Manufacturing*, vol.26, pp.83-91, 2010.

[21] T. Chen, Applying a fuzzy and neural approach for forecasting the foreign exchange rate, *International Journal of Fuzzy System Applications*, vol.1, no.1, pp.36-48, 2011.

[22] T. Chen and Y. C. Lin, A fuzzy-neural system incorporating unequally important expert opinions for semiconductor yield forecasting, *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, vol.16, no.1, pp.35-58, 2008.
[23] T. Chen, A fuzzy-neural knowledge-based system for job completion time prediction and internal due date assignment in a wafer fabrication plant, *International Journal of Systems Science*, vol.40, no.8, pp.889-902, 2009.
[24] T. Chen, An online collaborative semiconductor yield forecasting system, *Expert Systems with Applications*, vol.36, no.3, pp.5830-5843, 2009.
[25] T. Chen, A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan, *Computers & Industrial Engineering*, in press.
[26] T. Chen, Applying the hybrid fuzzy c means-back propagation network approach to forecast the effective cost per die of a semiconductor product, *Computers & Industrial Engineering*, vol.61, no.3, pp.752-759, 2011.
[27] T. Chen, Evaluating the mid-term competitiveness of a product in a semiconductor fabrication factory with a systematic procedure, *Computers & Industrial Engineering*, vol.53, pp.499-513, 2007.
[28] R. R. Yager, On possibilistic and probabilistic information fusion, *International Journal of Fuzzy System Applications*, vol.1, no.3, pp.1-14, 2011.