

# A Rule-based Forecasting System Integrating combining and Single Forecast for Decision Making

Chun-Ling Lin

Institute of Electrical  
and Control  
Engineering  
National Chiao Tung  
University, Hsinchu,  
Taiwan

Chen-Chun Lin

Institute of Management  
of Technology  
National Chiao Tung  
University, Hsinchu,  
Taiwan

Joseph Z. Shyu

Institute of  
Management of  
Technology  
National Chiao Tung  
University, Hsinchu,  
Taiwan

Chin-Teng Lin

Institute of Electrical  
and Control  
Engineering  
National Chiao Tung  
University, Hsinchu,  
Taiwan

## ABSTRACT

Many studies demonstrated that combining forecasts produces consistent but modest gains in accuracy. However, little researches define well the conditions under neither which combining is most effective nor how methods should be combined in each situation. In this paper, a rule-based forecasting system industry is proposed in order to compare forecast performance between combining forecasts and single forecasts, then define these conditions and to specify more effective combinations, finally suggest the best methods. Two comparative case studies for the telecommunications and TFT-LCD industry are proposed to examine the performance of the proposed system. Results from this study indicate that combining forecasts outperform single forecasts only when data set is data have various nonlinear characteristics. In this research, empirical evidence shows that rules based on causal forces improved the selection of forecasting methods, the structuring of time series, and the assessment of prediction intervals.

## Keywords

Combining Forecast, Single Forecast, Accuracy, Multiple Forecast, Rule-based Forecasting.

## 1. INTRODUCTION

Many studies mentioned that combined forecasts with important contribution by among others [1-5]. Compared with combined forecasting, the single approaches, such as Delphi, Forecasting by Analogy, Growth Curves, Trend Extrapolation, Measures of Technology, Correlation Methods, Causal Models, Probabilistic Methods etc. [6-8], are inconsistency between forecasts [9-10], and easily fail to achieve the level of capability forecast. Due to each of them can be instead of another method whose forecasters may have overlooked.

However, none of each combining forecast is a universal model that is suitable for all situations. Because it is difficult to completely know the linear or nonlinear characteristics of the time series data in a real problem. For this reason, an important motivation to combine different models is the fundamental assumption that one cannot identify the true process exactly, but different models may play a complementary role in the approximation of the data generating process. Hence, a rule-based forecasting system could be suggested to define these conditions and to specify more effective combinations in this research.

## 2. A RULE-BASED FORECASTING SYSTEM

Traditional extrapolation forecasting methods have two important limitations. First, they do not incorporate existing knowledge that shows which extrapolation methods are the best in various conditions. Second, they ignore the managers' knowledge about the situation.

In computer science, rule-based systems (RBS) can provide an efficient way to store and manipulate knowledge to interpret information in a useful way. Rule-based forecasting (RBF) is developed based on the concept of RBS. RBF is a type of expert system and can translate the forecasting expertise into a set of rules. From the combination which consists of simple extrapolation methods, the rules of RBF use the managers' domain knowledge and the characteristics of the data to produce a forecast. In addition, RBF is described in a rule-based forecasting process using judgment criteria in time-series extrapolation. Although the evidence for RBF's accuracy as a forecasting method is limited, the results are promising. Many original researches were done by previous research evidences.

The major advantage of RBF is to decide what methods to be used for various types of data. Thus, the findings can be applied to existing extrapolation programs. For example, the contrary series rule can be easily applied to other trend-based extrapolation models. RBF also provides a test bed for new or modified forecasting rules.

### 2.1 Judgment Criteria

In this paper, the mean square error (MSE) is employed for judgment the basic measure of accuracy. In statistics, MSE is used to measure the deviation between the actual and predicted values [11]. The smaller MSE corresponds to the better forecasting accuracy. Thus, MSE can be described by following equation:

$$MSE(f, x) = \frac{1}{n} \sum_{i=1}^n (f_i - x_i)^2 \dots \dots \dots (1)$$

where  $f_i$  and  $x_i$  denote the  $i_{th}$  forecast and  $i_{th}$  observation, respectively, and  $n$  presents the number of forecast.

### 2.2 System Design

In this paper, a rule-based forecasting system is proposed. It consists of four different forecasting modules which are

showed in Fig.1. The detail of these systems is described as follows:

(1) The data management subsystem: The data management sub system is an effective mechanism for storing, modifying, deleting, appending data. This subsystem can be managed by software called the database management system. In this paper, data management subsystem is designed to provide the two case studies with the capture and maintenance of current and historical data. Then, the data will be used as a base for forecasts;

(2) The model management subsystem: The model management subsystem can provide system's analytical capabilities and appropriate software management [12], by combining four various forecasting modules. When errors arise from faulty assumptions, bias, or mistakes in data, the errors can be reduced;

(3) The knowledge management subsystem: The knowledge management subsystem can support the operation of other subsystems or operate independently according to unequally-weighting-index database [12]. In this subsystem, the function and activities will aim at gathering, storing and giving access to information on current states of past data. The function and activities will then accumulate the estimation of unequally-weighting parameters for using in the forecasting model and currently established operational procedures. These unequally-weightings parameters will be used to predict. This procedure will be helpful to overcome the defects from static evaluation of emergency plans and regular weight assignment and can further reduce the impact of the model flaw on evaluation results;

(4) The dialogue subsystem: The dialogue subsystem can provide a user-friendly interface for decision makers to communicate with the forecasting system. It is responsible for the internal and external communications. A user sends requests to the system through this subsystem and the internal system sends feedback accordingly.

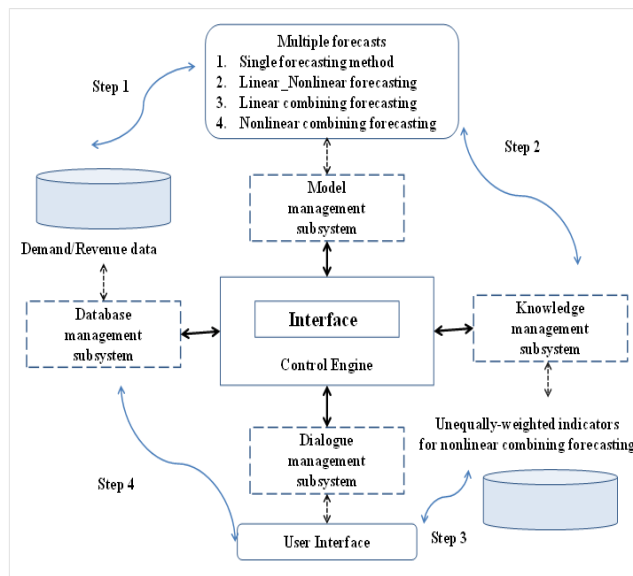


Fig.1. Rule-based forecasting system design

### 3. MULTIPLE FORECASTS

Single and combining forecasting models both have achieved successes in their own linear or nonlinear problems. In order to

develop a better forecasting model, a multiple forecasts, which collect the advantages from each combining method, is adopted to construct the rule-based forecasting system. It can be used to select better forecasting modules for better forecasting improvement.

Further, in order to tackle various forecasting problems (linear or nonlinear data characteristics) and improve forecasting accuracy, the process of the multiple forecasts is proposed in this section which is shown in Fig. 2. This system develops four different forecasting modules. Totally, there are 14 forecasting methods provided in those four different forecasting modules.

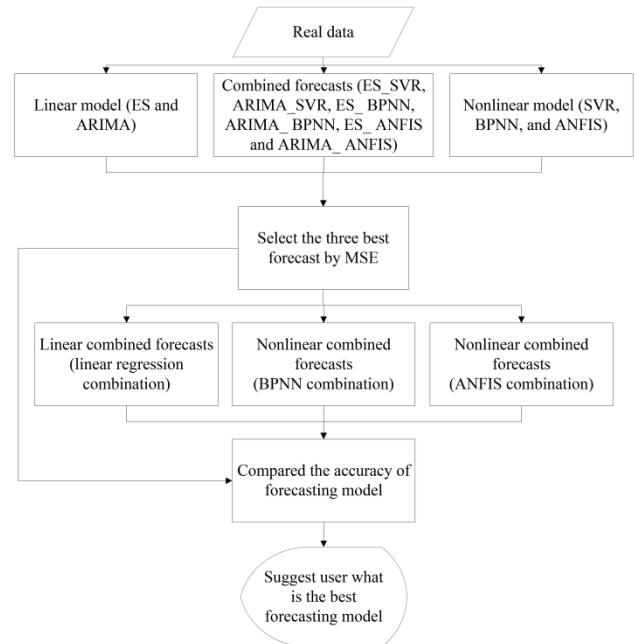


Fig.2. The process of multiple forecasts

The first module consists of two linear and three nonlinear forecasting methods: (1) The two linear methods are autoregressive integrated moving average model (ARIMA) and exponential smoothing model (ES); (2) The three Non-linear methods are propagation neural network (BPNN), adaptive neuro-fuzzy inference system (ANFIS) and support vector regression (SVR).

The second module is developed by combining linear forecasting methods and nonlinear forecasting methods. The data (linear) will be first forecasted by linear forecasting methods, and the rest data (non-linear) will then be forecasted by non-linear forecasting methods. This module can solve an approximation of a nonlinear objective function with linear constraints using a series of linear programs [18]. There are 6 models as follow: ES\_BPNN, ARIMA\_BPNN, ES\_SVR, ARIMA\_SVR, ES\_ANFIS, and ARIMA\_ANFIS.

The third module is constructed by linear combining forecasting methods. The linear regression combining forecast is also involved in this module. In the proposed method, we assume that the linear regression combining forecast have three inputs. These inputs of the linear regression combining forecast are selected from the forecasts of mentioned forecasting methods according to their MSE. For each forecast of the forecasting methods, the linear regression combining forecast are typically assigned with the same weighted at each

time frame, regardless of either global or local accuracy [13-15].

The final module is belonged to the nonlinear combining forecasting module. The final module includes two combining forecasts: ANFIS combining forecast and BPNN combining forecast [16-17]. We also assume that there are three inputs for the two combining forecasts. These inputs of the combining forecasts are selected from the forecasts of mentioned forecasting methods according to their MSE. For each forecast of forecasting methods, different weights are applied to both nonlinear combining forecasting methods at each time frame.

## 4. CASE STUDY

### 4.1 Case 1: Demand Forecasting of the Telecommunication Industry

In the first case, the subject of this research mainly focuses on forecasting the future demand on the telecommunication in Taiwan, to prove the efficiency of the proposed system, and this case majorly focus on the understanding of likely market trends. The original data (from January 2006-June 2008) is from Telecommunications Carriers Association subscriber database in Taiwan in 2008 [19].

Results show that data characteristic in case 1 are linear as Fig. 3B. The findings also show that SVR and ARIMA-ANFIS outperform other listing methods numerically as Fig. 3A. It can be explained by that employing useful SVR (single nonlinear methods) is highly efficient to forecast linear data. Besides, we argue that the single nonlinear method can efficiently reduce forecasting error than other combining forecasting methods when analyzing linear data.

### 4.2 Case2: Revenue Forecasting of the Global TFT-LCD Industry

The case 2 is revenue forecasting for the TFT-LCD industry. In the last five years, TFT-LCD industry has witnessed drastic changes in the intensity of competition. The industry is going through turbulent transformation, and stepping in to the mature stage of an industry life cycle. Most companies are actually relooking at the revenues and desperate for growth. At the moment, forecasting with high accuracy is an important aid in effective and efficient planning. This paper attempts to analyze various regions fluctuating revenue data to provide a basic strategy for future demand. The monthly revenue data from October, 2009 to February, 2011 are obtained from Asia Pacific, China and North American regions.

From the experimental results listed in Table 1, we can find the effectiveness of multiple forecasts. The smaller MSE corresponds to the better forecasting accuracy. Summary results can be seen that revenue time series fluctuated in Asia Pacific, China and North American regions, data are all with fluctuating nonlinear characteristics and forecasting results are shown in Fig. 3D, Fig. 3F and Fig. 3H. The forecast performances of multiple forecasts are shown in Fig. 3C, Fig. 3E and Fig. 3G. The prediction results including Case 1 and Case 2 are summarized as Table 1 to present that which methods are higher MSE performance on forecasting. Again, according to the results in Table 1, it is clear that no each model is appropriate for all situations. The assumption reason

for this is that various data characteristics from different regions that also can be seen as Table 1. It also implicates that the data characteristics and corresponding metrics are mapped to forecasting performance evaluation results to construct rules for forecasting method selection. To take advantage of each method, we introduce a rule-based forecasting system that can be used to select the best method.

Table 1. Summary of Results

Region	Data characteristics	The best MSE forecast performances and method
<b>Case 1: Telecommunication Industry</b>		
Taiwan	Linear	MSE=1.98E+11, by SVR.
<b>Case 2: TFT-LCD Industry</b>		
Asia Pacific	Non-linear	MSE=7.17E+09, by ES_SVR
China	Non-linear	MSE=7.89E+11, by ANFIS non-linear combination.
North American	Non-linear	MSE=5.44E+10, by SVR.

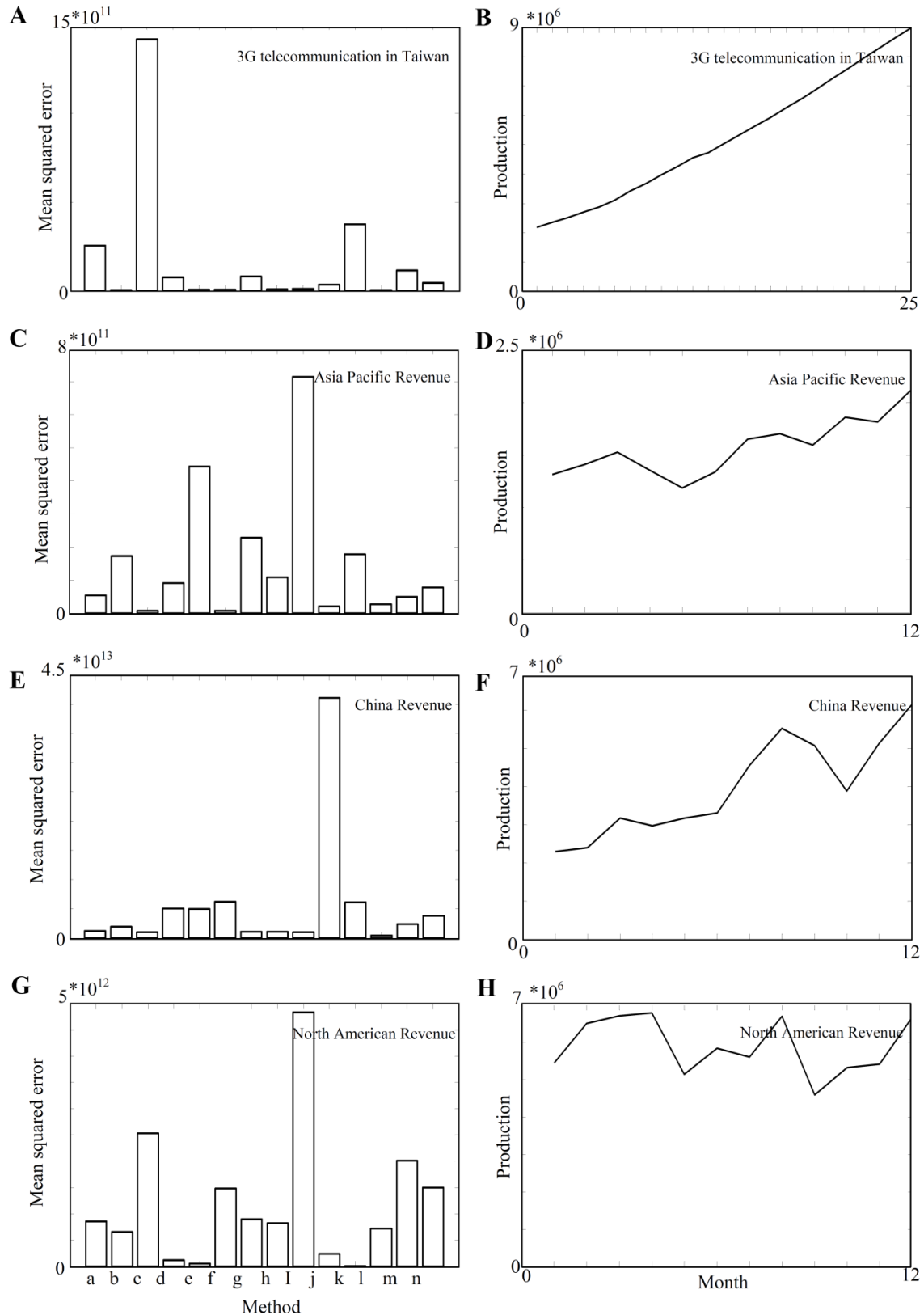
### 4.3 System Interface

This system has developed a procedure to optimize the choice of rule-based forecasting system. The initial screen for the system is given in Fig. 4A. Once the dataset is chosen as Fig. 4B, the system automatically attempts to enter the historical data time, and a new screen allows the user to choose the forecasting mode to develop a custom forecasting model. Then the system will decide which method is the best method to forecast data by the MSE performance as Fig. 4C. The best method to forecast future revenue is the ES\_SVR method shown in Fig. 4C, and its MSE performance is 7.17E+09.

## 5. CONCLUSIONS

For previous studies, the dominant paradigm for forecasting research has been statistical modeling. These efforts have done little to incorporate domain knowledge into extrapolations. In order to bridge the research gap, a useful technical rule for forecasting is supported for variants data analysis and is proposed in this paper. As expected, the proposed rule-based forecasting system does better than single and combining forecasts in linear or non-linear data

For forecasting researches, the proposed system has the contributions to identify the time series problems faster, to make better use of different forecasting methods for analyzing different data, and to choose the best method with high forecast performance that can lead to better decisions and better management of corporate assets. Moreover, the automating feature detection of the rule-based forecasting system has also introduced consistency and reliability into the forecasting process. The added reliability could contribute to further effort in the validation and refinement of rule-based forecasting. These results are encouraging considering the absence of domain knowledge about the series in the competition. Because these findings resulted from the application of theory and empirical testing, we are optimistic that continued refinement of this research program will produce further improvements.



**Fig 3: The performances of multiple forecasts and forecasting results**

Forecasting method codes: “a” method =ES, “b” method=ARIMA; “c” method=SVR; “d” method=BPNN; “e” method=ANFIS; “f” method=ES\_SVR; “g” method=ARIMA\_SVR; “h” method=ES\_BPNN; “i” method=ARIMA\_BPNN; “j” method=ES\_ANFIS; “k” method=ARIMA\_ANFIS; “l” method=Linear combined forecasting; “m” method= Nonlinear combined forecasting by BPNN; “n” method= Nonlinear combined forecasting by ANFIS

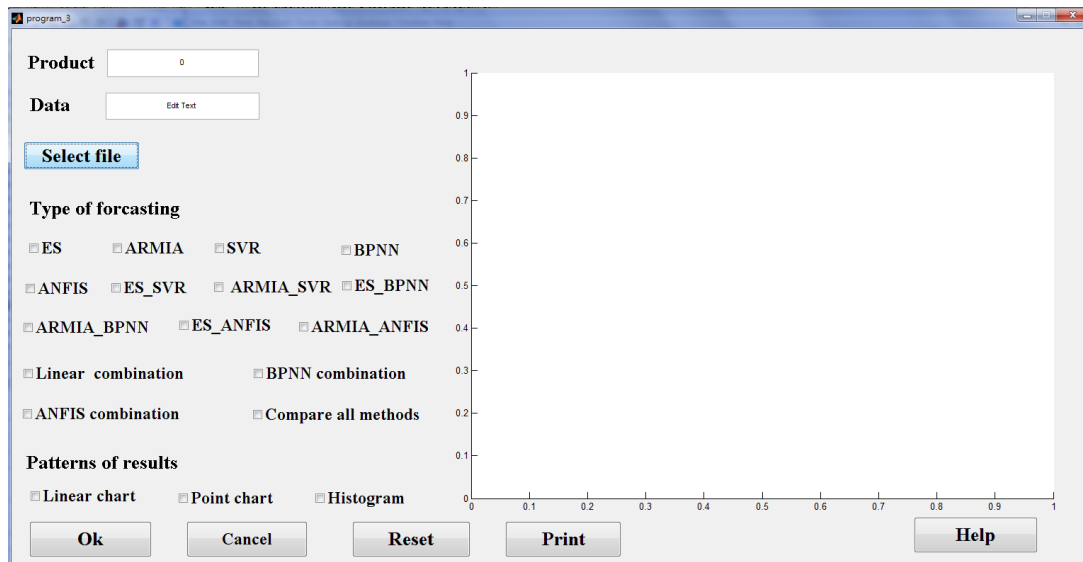


Fig.4A. System Interface (1)

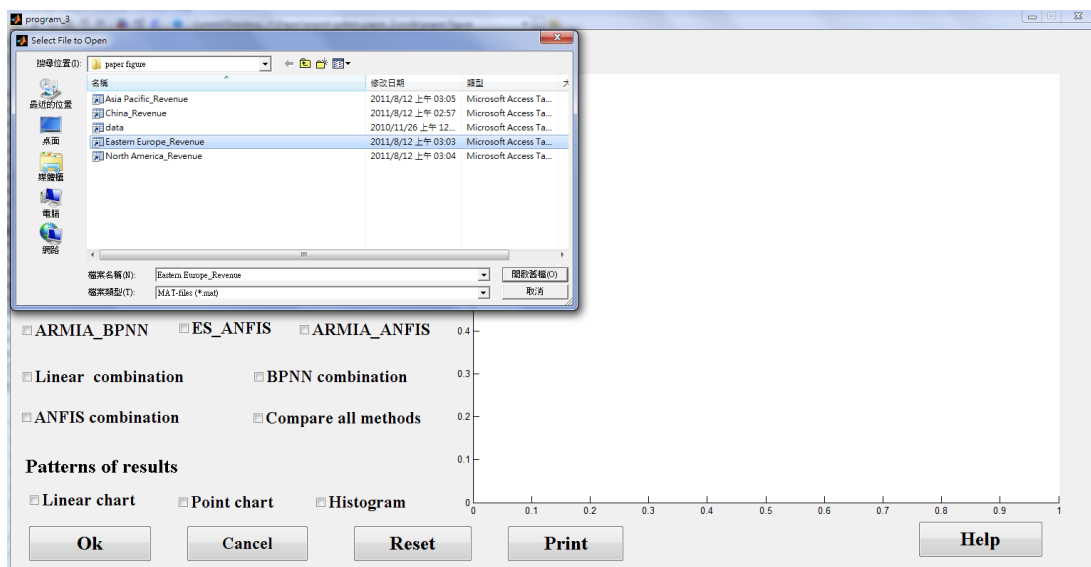


Fig.4B. System Interface (2)

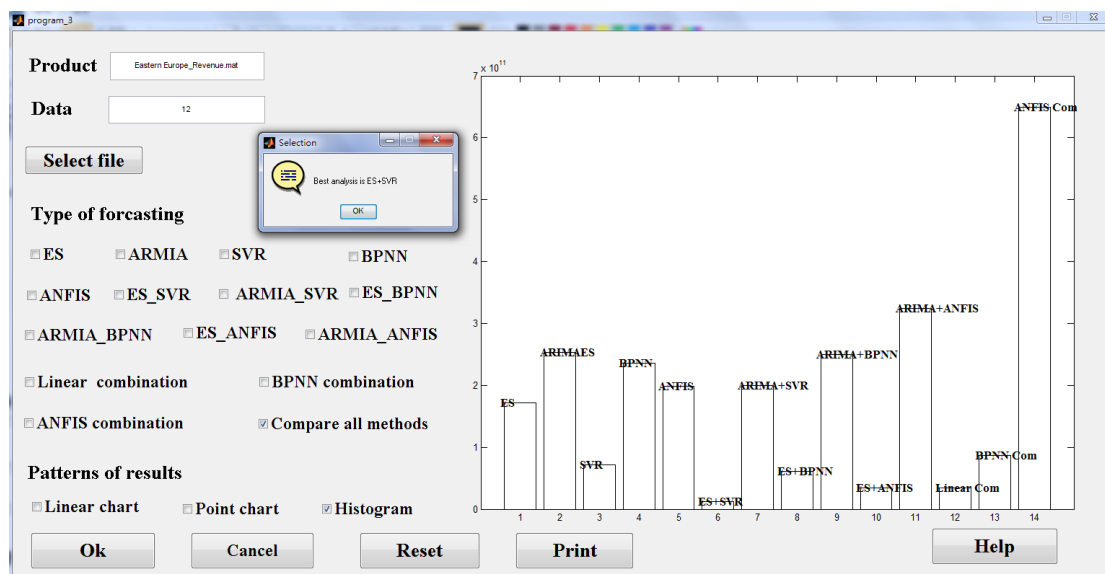


Fig.4C. System Interface (3)

## 6. ACKNOWLEDGMENTS

We thanks for the experts who have contributed towards development of the template. This work was supported in part by the UST-UCSD International Center of Excellence in Advanced Bioengineering, in part by the Taiwan National Science Council I-RiCE Program under Grant Number NSC-100-2911-I-009-101, in part by the Aiming for the Top University Plan of National Chiao Tung University, the Ministry of Education, Taiwan, under Grant Number 100W9633, also supported NSC 100-2220-E-009 - 060(100N113).

## 7. REFERENCES

- [1] P. Poncela, J. Rodríguez, R. Sánchez-Mangas, E. Senra, Forecast combination through dimension reduction techniques, *International Journal of Forecasting* 26 (3) (2010) 445-626.
- [2] A.S. Jore, J. Mitchell, S.P. Vahey, Combining forecast densities from VARs with uncertain instabilities, *Journal of Applied economics* 25(4) (2010) 621-634.
- [3] J. Smith, K.F. Wallis, A Simple explanation of the forecast combination puzzle, *Oxford Bulletin of Economics and Statistics* 71 (2009) 331-355.
- [4] M.Y. Wang, W.T. Lan, Combined forecast process: Combining scenario analysis with the technological substitution model, *Technological Forecasting and Social Change* 74 (3) (2007) 357-378.
- [5] M. Singh, R. Singh, V. Shinde, Application of Software Packages for Monthly Stream Flow Forecasting of Kangsabati River in India, *International Journal of Computer Applications* 20(3) (2011) 7-14.
- [6] J.P. Martino, *Technological Forecasting for Decision Making* (3rd ed.), McGraw-Hill, New York, 1993.
- [7] V.A. Banuls, J.L. Salmeron, A scenario-based assessment model— SBAM, *Technological Forecasting and Social Change* 74(6) (2007), 750-762.
- [8] C.C. Holt, Forecasting seasonals and trends by exponentially weighted moving averages, *International Journal of Forecasting* 20(1) (2004) 5-10.
- [9] M.Y. Wang, W.T. Lan, Combined forecast process: Combining scenario analysis with the technological substitution model, *Technological Forecasting and Social Change* 74 (3) (2007) 357-378.
- [10] Saima H., J. Jaafar, S. Belhaouari, T.A. Jillani, ARIMA based Interval Type-2 Fuzzy Model for Forecasting, *International Journal of Computer Applications* 28(3) (2011) 17-21.
- [11] G.J. Lobo, R. D. Nair, Combining judgmental and statistical forecasts: An application of earnings forecasts, *Decision Sciences* 21(2) (1990) 446-460.
- [12] P. Gottschalk, *Knowledge Management in Law Enforcement: Technologies and Techniques*, Idea Group Publishing, Hershey, Pa, USA, 2007.
- [13] S.I. Gunter, Nonnegativity restricted least squares combinations. *International Journal of Forecasting* 8(1) (1992) 45-59.
- [14] R.J. Vokurka, B.E. Flores, S.P. Pearce, Automatic feature identification and graphical support in rule-based forecasting: a comparison, *International Journal of Forecasting*, 12(4) (1996) 495-512.
- [15] P.H. Franses, R. Legerstee, Combining SKU-level sales forecasts from models and experts. *Expert Systems with Applications*, 38(3) (2011) 2365-2370.
- [16] A. Tarsauliya, S. Kant, R. Kala, R. Tiwari, A. Shukla, Analysis of Artificial Neural Network for Financial Time Series Forecasting, *International Journal of Computer Applications* 9(5) (2010) 16-22.
- [17] G.J. Lobo, R. D. Nair, Combining judgmental and statistical forecasts: An application of earnings forecasts, *Decision Sciences* 21(2) (1990) 446-460.
- [18] K.Y. Chen, Combining linear and nonlinear model in forecasting tourism demand, *Expert Systems with Applications* 38(8) (2011) 10368-10376.
- [19] The Ministry of Transportation and Communications (MOTC) in Taiwan, <http://www.motc.gov.tw>, 2011.