



Hybrid multi-model forecasting system: A case study on display market



Chen-Chun Lin^a, Chun-Ling Lin^{b,*}, Joseph Z. Shyu^a

^a Institute of Management of Technology, National Chiao Tung University, No. 1001, Ta Hsueh Rd., Hsinchu 300, Taiwan

^b Department of Electrical Engineering, Ming Chi University of Technology, No. 84, Gongzhuang Rd., Taishan Dist., New Taipei City 243, Taiwan

ARTICLE INFO

Article history:

Received 10 January 2012

Received in revised form 5 August 2014

Accepted 2014

Available online 11 August 2014

Keywords:

Hybrid multi-model forecasting system

Prediction

Display markets

Mean square error (MSE)

Mean absolute percentage error (MAPE)

Average square root error (ASRE)

ABSTRACT

This paper provides a novel hybrid multi-model forecasting system, with a special focus on the changing regional market demand in the display markets. Through an intensive case study of the ups and downs of the display industry, this paper examines the panel makers suffered from low panel price and unstable market demand, then they have changed to react to the rapid demand in the market or have lower panel stock for keeping supply and demand more balanced. In addition, this paper suggests a co-evolution forecasting process of sales and market factor. It can automatically apply various combinations of both linear and nonlinear models, and which alternatives deliver the lowest statistical error and produce a good estimate for the prediction of markets.

Moreover, this article shows how the system is modeled and its accuracy is proved by means of experimental results; and judged by 3 evaluation criteria, including the mean square error (MSE), the mean absolute percentage error (MAPE), and the average square root error (ASRE) were used as the performance criteria to automatically select the optimal forecasting model. Finally, the results showed that the proposed system had considerably better predictive performance than previous and individual models. To summarize, the proposed system can reduce the user's effort for easier obtaining the desired forecasting results and create high quality forecasts.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The flat-panel display (FPD) is a landmark sector all over the world in terms of technology innovation. This market is growing based on the competitiveness of three major technologies: thin-film transistor-liquid crystal displays (TFT-LCD), plasma display panels (PDP) and organic light-emitting diodes (OLED). TFT-LCD has the largest market share. This technology dominates the market, as it can be used in different types of applications, ranging from small devices including mobile phones to large applications including televisions.

However, TFT-LCD manufacture has high risk and low affixation. Because high-risk industry where failure for market estimation can lead to the elimination of an enterprise and where a timely, large-scale investment is essential; industry where large companies that should have the capacity to mobilize large capital are fully equipped with necessary parts and materials.

Research on flat panel displays (FPD), which started in the 1960s, has finally reached the commercialization stage in the form of large plasma display panels (PDPs) and liquid crystal displays

(LCDs). Japanese companies led initial technological development in the LCD upstream industry in the 1990s. But since 2000, Korea and Taiwan have made bold investments, and they are leading the global market. In 2010 China started to join this market, global manufacturers of TFT-LCD panels have established the majority of LCD module assembly plants in China to take advantage of lower labor costs. With investment in display production facilities likely to decline in other countries, production of TFT-LCD manufacturing equipment in China will account for a greater share of the world market. Now China has become a major hub in TFT-LCD manufacturing, and the TFT-LCD industry is one of the most dynamic industries.

Major manufacturers by country are Korea (Samsung, LG Display), Taiwan (AUO, Innolux, CPT, Hannstar), Japan (Sharp, TMD, NEC, Hitachi), China (BOE-OT, CEC-Panda, CSOT, Tianma), etc. Now the industry has diverse upstream markets as following: (1) small and mid-sized products including smart phones, IT products (i.e. monitors, tablet PC, desktop PC, laptops and automotive heads-up display) and (2) large-sized products including household appliances (i.e. LCD TV and monitor).

After the fall of 2008 and the European debt crisis, which began in late 2009, the shift still significantly influences the demand ratio of the TFT-LCD regional markets in the world until now. The

* Corresponding author. Tel.: +886 2 2908 9899x4819.

E-mail address: ginnylin@mail.mcut.edu.tw (C.-L. Lin).

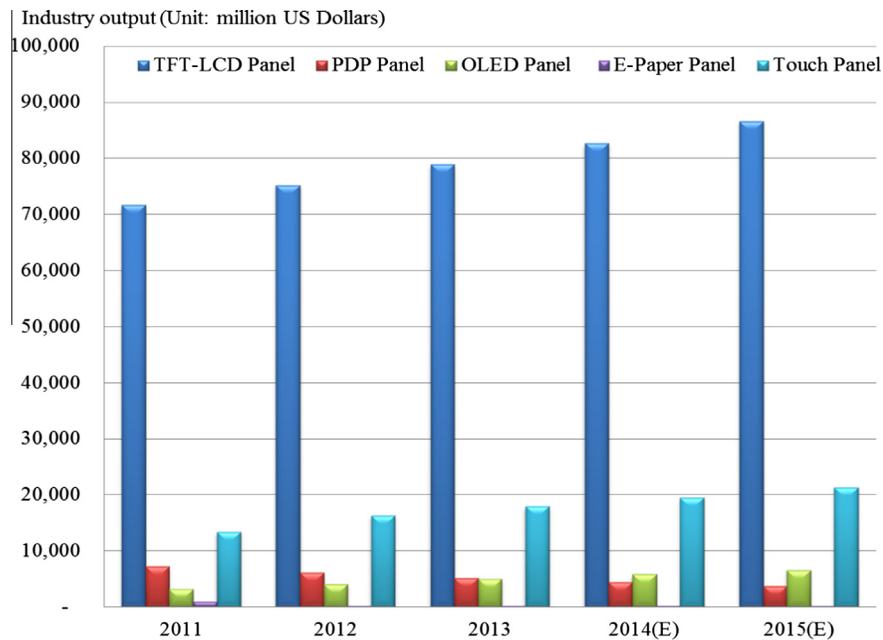


Fig. 1. The global flat panel display industry output (Data Resource: Photonics Industry and Technology Development Association (PIDA) in Taiwan, 2014. <http://www.pida.org.tw/usub/en/index.asp>).

TFT-LCD industry also has witnessed drastic changes in the intensity of competition in 2008–2010. This industry is undergoing a turbulent transformation as it becomes a mature industry as Fig. 1. However, TFT-LCD panel manufacturers are undoubtedly looking forward to sustainable growth, but they cannot simply wait for demand to increase and then react to that increase to generate revenue. These companies should examine the various revenue sources in the regional markets and seek new opportunities to increase revenue, or build a sound foundation in shaping effective output planning for the potential market. Hence, the future forecast should analyze historical data and forecast projections to deliver the most detailed information and insights available.

Due to those reasons mentioned above, this research hopes to develop an efficient process, and a functional tool to predict the rapid development of the TFT-LCD market. Of interest to note is forecasting is a problem that arises in many economic and managerial contexts, and hundreds of forecasting procedures have been developed over the years for many different purposes, both in business enterprises and elsewhere. Previous forecasting studies relied on qualitative methods or patent analysis, which can be quite useful for other forecasting problems but have been shown to be inappropriate for industrial development forecasting [5,29]. Recently, both theoretical and empirical results have suggested that combining forecasting methods can be an effective way to achieve better predictive performance over individual models [9,23]. Contributions from many researchers have improved the quality of the predictions and provided combined forecasting models for decision makers [1,16,26,30,35]. As a result, there have been profound changes in the forecasting field. Combining various linear and nonlinear models offers solutions in which models are combined in an optimal way and can be applied in real-world situations such as forecasting macroeconomic time series [32], tourist demand [6], and exchange rates [2].

Based on the parameters selected, the combined forecasting models can be roughly classified into three categories: (1) linear/equally weighted combined forecasts; (2) nonlinear/unequally weighted combined forecasts; and (3) combined forecasts from linear and nonlinear models. Linear methods such as the Bayesian method typically place equal weight on each of the sub-classifiers

in each time frame, regardless of their global or local accuracy [4,12,15,18,38]. Nonlinear methods apply unequal weights for the averaging of past observations (i.e., more recent observations are given more weight in forecasting than older observations); examples include neural networks, adaptive neuro-fuzzy inference systems, and fuzzy set methods [20,24,25]. Combining linear and nonlinear methods can retain the robustness while reducing the complexity considerably, such as in the combination of artificial neural networks (ANNs) and auto-regressive integrated moving average (ARIMA) methods. Among these, nonlinear combined forecasts and combined forecasts from linear and nonlinear models have proven to be very effective for demand forecasting in addition to other linear or nonlinear applications [6,27,34].

Several algorithms commonly found in the literature have the potential to surpass the performance of an individual predictor by combining the outputs of a collection of complementary predictors. In bagging, various methods are generated by applying a learning algorithm to independent bootstrap tests of the primary training data [3,17]. Boosting is another popular ensemble algorithm and was originally developed for classification problems. A sequence of models is obtained from a given dataset using an adaptive learning algorithm and different parameters for various training cases [17]. Adaptive neuro-fuzzy inference systems (ANFISs) outperform other individual methods, and the forecasting accuracy can be improved effectively using combined forecasts, as was done for a panel manufacturer [39].

Based on the forecasting performance of combined methods, Wang and Nie [41] proposed the combination of a back-propagation neural network (BPNN) and support vector machines (SVMs), which have the best forecasting performance, and showed that the combined forecasting model can greatly enhance the accuracy of predictions of a stock index. Other research showed that the adaptive neuro-fuzzy inference system method outperforms other methods in forecasting panel demand [39], automobile sales [40], and the demand for telecommunication technology [21].

However, none of these methods 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 an actual problem. An important motivation to combine the forecasts

from different models is the fundamental assumption that one cannot identify the true process exactly, and thus, different models may play a complementary role in the approximation of the process that produced the data.

Hence, this paper presents a hybrid multi-model forecasting system that combines the forecasts from various linear and nonlinear models and compares the performance of this system with that of nonlinear combined-forecast models. The main advantage of the proposed method is that it can be used to select better forecasting modules for greater performance.

The remainder of this paper is organized as follows. In Section 2, the exponential smoothing (ES) method, the ARIMA, the back-propagation neural network (BPNN), the adaptive neuro-fuzzy inference system (ANFIS), the support vector regression (SVR) method, and the combination methodology are described. Section 3 describes the data source and the evaluation criteria used for comparing the forecasting techniques. Section 4 compares the results obtained from the combinations of models against the nonlinear combined forecasts and discusses the forecasting system software. Finally, Section 5 provides concluding remarks.

2. Methodology

In this paper, we propose a systematic approach to combine different efficient methodologies for improving forecasting performance, especially focus on combined forecasting technique from linear and nonlinear models. Both the nonlinear combining forecast and the combining forecast from linear and nonlinear models have achieved successes in their own linear or nonlinear problems [6,11,34]. Each combining forecast method has its own advantages and disadvantages. In order to take advantage of the strengths of each combining method to develop the best forecast possible alternatives, we introduce a multiple forecasting system that can be used to select better combining forecasting modules for better forecasting improvement.

2.1. The combining forecasts from linear and nonlinear models

Two individual linear methods (Exponential smoothing/ES and Autoregressive Integrated Moving Average/ARIMA) and other two individual nonlinear methods (Back-propagation neural network/BPNN and Support vector regression/SVR), are selected as techniques for optimizing dynamic forecasting problem based on the combining forecast from linear and nonlinear models. Because a lot of work on using ES, ARIMA, BPNN and SVR as techniques for dynamic combined forecasting often have been previously reported with good predicting performance [6,14,24,33].

Then, this system is majorly proposed to provide four kinds of “combining forecasts from linear and nonlinear models” as follow: (1) ES_BPNN; (2) ES_SCR; (3) ARIMA_BPNN; and (4) ARIMA_SVR, which all have both linear and nonlinear modeling capabilities that can be a good strategy for practical use. Then, this study compares these results with the results from the nonlinear combining forecast by ANFIS and the individual forecast by ANFIS.

Previous researches in combining forecasts from linear and nonlinear models believe that it may be reasonable to consider a time series to be composed of a linear autocorrelation structure and a nonlinear component [6,42]. That is:

$$Y_t = M_t + N_t \tag{1}$$

where M_t is the linear component and N_t is the nonlinear component of the combination models. Both M_t and N_t have to be estimated from the data set. First, linear model (ES and ARIMA) is used to model the linear part of data set, and then the residuals from the linear model will contain only the nonlinear relationship.

Let E_t represent the residual at time t as obtained from the linear model, then:

$$E_t = Y_t - \widehat{M}_t \tag{2}$$

where \widehat{M}_t denotes the forecast value of the linear model at time t . In order to model the nonlinear residuals from the linear model, the nonlinear model (SVR and BPNN) can be used. In this study, the author built four various combination models with the following input layers:

$$E_i^{linear} = f^{nonlinear} (E_{i-1}^{linear}, E_{i-2}^{linear}, E_{i-3}^{linear}) \tag{3}$$

where E_i^{linear} represent the residual at time t from the linear models (ES and ARIMA), $f^{nonlinear}$ a nonlinear function determined by the nonlinear models (SVR and BPNN). Here, this study proposed four combined models, and called them as ES_BPNN, ARIMA_BPNN, ES_SVR, ARIMA_SVR. Therefore, the combined forecast will be:

$$\widehat{Y}_i = \widehat{M}_i + \widehat{N}_i \tag{4}$$

where \widehat{N}_i is the forecast value of Eq. (1).

The individual linear and nonlinear methods can be described by the following sections.

2.1.1. Exponential smoothing (ES)

ES is a widely used linear model that can be applied to time series data on the basis of the moving average technique [33]. It uses a weighted average of past data as the basis for a forecast. The procedure gives heaviest weight to recent information and smaller weights to observations in the more distant past. The reason for this is that the future is more dependent on the recent past than on the distant past. The formula for exponential smoothing is

$$F_{i+1} = \alpha X_i + (1 - \alpha)F_i \tag{5}$$

where F_{i+1} is the forecast value of period $i + 1$; α is the smoothing constant ($0 < \alpha < 1$), X_i is the actual value of period i , and F_i is the forecast value of period i . α can be thought of as the weight given to past history. The method is called “exponential” since the forecasted value is the discrete convolution of the observed sequence with an exponential curve with a time constant $1/(1 - \alpha)$. Alternately, if the value of X_i becomes fixed, the error $X_i - F_i$ decays exponentially.

2.1.2. Autoregressive Integrated Moving Average (ARIMA)

ARIMA is the most popular linear model that fitted to time series data either to better understand the data or to predict future forecasting. It is applied in some cases where data show evidence of non-stationarity, where an initial differencing step can be applied to remove the non-stationarity [6,36]. The model is generally referred to as an ARIMA(p, d, q) model where p , d , and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. When one of the terms is zero, it is usual to drop AR, I or MA. For example, an I(1) model is ARIMA(0, 1, 0), and a MA(1) model is ARIMA(0, 0, 1).

$$\varnothing(B)\nabla^d x_i = \theta(B)\varepsilon_i \tag{6}$$

where x_i and ε_i represent the number of visitors and random error terms at period i respectively. B is a backward shift operator defined by $Bx_i = x_{i-1}$, and related to ∇ by $\nabla = 1 - B$, $\nabla^d = (1 - B)^d$, d is the order of differencing. $\varnothing(B)$ and $\theta(B)$ are autoregressive (AR) and moving averages (MA) operators of orders p and q , respectively, and are defined as:

$$\varnothing(B) = 1 - \varnothing_1 B - \varnothing_2 B^2 - \dots - \varnothing_p B^p \tag{7}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \tag{8}$$

where $\phi_1, \phi_2, \dots, \phi_p$ are the autoregressive coefficients and $\theta_1, \theta_2, \dots, \theta_q$ are the moving average coefficients. In order to fit an ARIMA model to the raw data, the ARIMA model involves the following four-step iterative cycles [6]: (1) Identification of the ARIMA (p, q, d) structure; (2) estimation of the unknown parameters; (3) goodness-of-fit tests on the estimated residuals; and (4) forecast future outcomes based on the known raw data. The ε_i should be independently and identically distributed as normal random variables with mean = 0 and constant variance σ^2 . The roots of $\phi_p(x_i) = 0$ and $\theta_q(x_i) = 0$ should all lie outside the unit circle.

2.1.3. Back-propagation neural network (BPNN)

BPNN is a nonlinear method. It is one of the most popular neural network models for use in business applications, especially with time series prediction problems such as sales forecasting [14]. We purpose to use BPNN as the selected nonlinear models to combine with the linear ones. The key motivation for doing so is due to the truth that BPNN do not make any assumption about the data. Instead, they try to learn the functional form of the true model from the data itself. For these reasons, we use BPNN to combine the linear model in forecasting the revenue trend.

Based on the algorithm of BPNN, it typically employs three or more layers of processing elements: an input layer, an output layer, and at least one hidden layer. The back propagation learning algorithm involves a forward-propagation step followed by a backward-propagation step. Both the forward and backward-propagation steps are done for each signal presentation during training.

2.1.3.1. Forward-propagation algorithm. This forward-propagation step is initiated when an input signal is presented to the network. Incoming connections to unit j are at the left and originate at units in the layer below. Output values from these units arrive at unit j and are summed by

$$S_j = \sum_{i=1}^n x_i w_{ji} \tag{9}$$

where x_i is the activation level of unit i , and w_{ji} is the weight from unit i to unit j . After the incoming sum S_j is computed, a sigmoid function F is used to compute $F(S_j)$. After the sigmoid function is computed, the resulting value becomes the activation level of unit j . This value, the output of unit j , is sent along the output interconnections.

2.1.3.2. Backward-propagation algorithm. Here, the error (δ) values are calculated for all processing elements and weight changes are calculated for all interconnections. The calculations begin at the output layer and progress backward through the network to the input layer. The error-correction step takes place after a signal is presented at the input layer and the forward-propagation step is complete. Then, the weights are adjusted for all interconnections that go into the hidden layer. The process is continued until the last layer of weights has been adjusted. If unit j is in the output layer, then its error value is

$$\delta_j = (t_j - a_j) \times F'(S_j) \tag{10}$$

where t_j is the target value for unit j , a_j is the output value for unit j , $F'(x)$ is the derivative of the sigmoid function F , and S_j is the weighted sum of inputs to j .

2.1.4. Support vector regression (SVR)

SVR is a nonlinear method. In most real-world problems, linear function approximation is of limited practical use. The solution is to map the input data in a higher dimensional feature space, in which the training data may exhibit linearity, and then to perform linear regression in this feature space [24]. Let x_i be mapped into a

feature space by a nonlinear function $\phi(x)$; the decision function can be written as:

$$f(w, b) = w \times \phi(x) + b \tag{11}$$

Similarly, the nonlinear regression problem can be expressed as the following optimization problem.

$$\begin{aligned} & \min_{w, b, \xi, \xi^*} \frac{1}{2} w^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\ & \text{Subject to } y_i - (w \times \phi(x_i) + b) \leq \varepsilon + \xi_i \\ & (w \times \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, l \end{aligned} \tag{12}$$

Then, the dual form of the nonlinear SVR can be expressed as:

$$\begin{aligned} & \min_{\alpha_i, \bar{\alpha}_i} \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \bar{\alpha}_i)(\alpha_j - \bar{\alpha}_j) \phi(x_i) \times \phi(x_j) + \varepsilon \sum_{i=1}^l (\alpha_i - \bar{\alpha}_i) \\ & - y_i \sum_{j=1}^l (\alpha_i - \bar{\alpha}_i) \\ & \text{Subject to } \sum_{i=1}^l (\alpha_i - \alpha_i) = 0 \\ & 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, l \\ & 0 \leq \bar{\alpha}_i \leq C, \quad i = 1, 2, \dots, l \end{aligned} \tag{13}$$

Little knowledge may be available as a basis for selecting an appropriate nonlinear function $\phi(x_i)$, and further, the computation of $\phi(x_i) \times \phi(x_j)$ in the feature space may be too complex to perform. An advantage of SVR is that the nonlinear function $\phi(x_i)$ need not be used. The computation in input space can be performed using a “kernel” function $K(x_i, x_j) = \phi(x_i) \times \phi(x_j)$ to yield the inner products in feature space, circumventing the problems intrinsic in evaluating the feature space. Functions that meet Mercer’s condition can be proven to correspond to dot products in a feature space. Therefore, any functions that satisfy Mercer’s theorem can be used as a kernel. In this study, we chose the radial basis function as kernel function:

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \tag{14}$$

Finally, the kernel function allows the decision function of nonlinear SVR to be expressed as follows.

$$f(x_i) = \sum_{k=1}^l (-\alpha_k - \bar{\alpha}_k) K(x_i, x_k) + b \tag{15}$$

The parameters that dominate the nonlinear SVR are the cost constant C , the radius of the insensitive tube ε , and the kernel parameters. These parameters are mutually dependent so changing the value of one parameter changes other parameters. The meanings of parameters C and ε can be interpreted. The parameter C controls the smoothness or flatness of the approximation function. A greater C value, corresponding to a greater penalty of errors, indicates that the objective is only to minimize the empirical risk, which makes the learning machine more complex. In this study, determining appropriate values of C and ε is often a heuristic trial-and-error process. The optimal values of SVR parameters may vary substantially among cases.

2.2. Nonlinear combining forecast by ANFIS

The ANFIS, first proposed by [19], combined the benefits of artificial neural network (ANN), and fuzzy inference systems. In this study, the processing of the nonlinear combining forecast by ANFIS was same as the one of the individual forecast by ANFIS. Those two

methods only had the difference in input data. The individual forecast by ANFIS adopted the training data as the inputs directly. The nonlinear combining forecast by ANFIS selected the estimated values from three forecasting model as the inputs to build the nonlinear combining forecasting system. The three forecasting models can be selected from three nonlinear forecasts models and four combining forecasts models which mentioned above in Section 2.1. The MSE, MAPE, ASRE values of three forecasting model were the lowest of three values than other models. The first-order Sugeno fuzzy model has become a common practice on ANFIS implements in the past. Thus, we used the same model. The five steps process shows as follows: (1) In layer 1, each node is called an input linguistic node and corresponds to one input linguistic variable. The nodes transmit input forecasts to the next layer directly. Each node function can be modeled by fuzzy membership function. Here, the generalized bell membership function and Gaussian membership function are used; (2) in layer 2, each node in this layer calculates the firing strength of a rule via multiplication; (3) in layer 3: The i th node in this layer calculates the ratio of the n th rule's firing strength to the sum of all the rules' firing strengths. The result would be the normalized firing strengths. For convenience, the output of this layer will be called the normalized firing strengths; (4) in layer 4: Each node in this layer is a square node with a node function. Parameters in this layer will be referred to as consequent parameters by node function; (5) in layer 5: The single node in this layer computes the final combining forecast as the summation of all incoming forecasts. Here, we assumed ANFIS as having two inputs, x and y , and one output f to delineate the regular framework of ANFIS in Fig. 2.

2.3. The combining forecasts methodology

Both “nonlinear combining forecasts” and “combining forecasts from linear and nonlinear models” have achieved successes in their own linear or nonlinear problems. Each combining forecast has its own advantages and disadvantages. In order to take advantage of the strengths of each combining method to develop the best forecast possible alternatives, we introduce a hybrid multi-Model forecasting system that can be used to select better forecasting modules for better forecasting improvement. This system is majorly proposed based on four kinds of “combining forecasts from linear and nonlinear models” (e.g. ES_BPNN, ES_SCR, ARIMA_BPNN, ARIMA_SVR), and “nonlinear combining model forecasting by ANFIS, which all have both linear and nonlinear modeling capabilities that can be a good strategy for practical use.

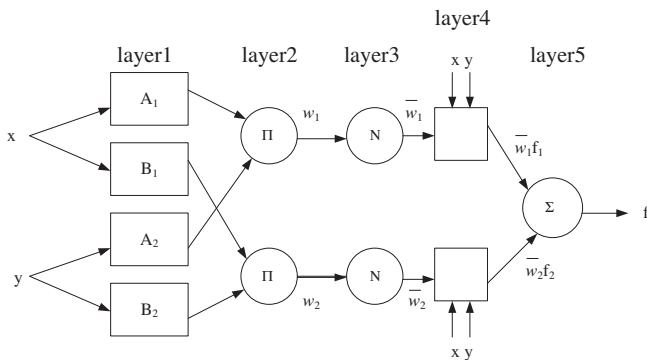


Fig. 2. ANFIS architecture.

3. Data set and performance criteria

To test the effectiveness of this proposed system, monthly TFT-LCD revenue data for four regions (the Asia-Pacific region, China, North America and Eastern Europe) were gathered as a representative sampling, during this experiment and then analyzed to form a valid testing. The data were obtained between January 2008 and June 2011, from a leading global market research and consulting firm “display search”.

The data were divided into two sets, a training set (in-sample data) and a test set (out-of-sample data), for each TFT-LCD revenue time series to assess the performance of the forecasting methods selected for this research. To achieve a more reliable and accurate result, a longer period was used for training.

Once the training stage was complete, the various methodologies were applied to the test data. The ANFIS was trained using the lowest MSE and the lowest MAPE, which are defined in the following subsection, as the performance criteria.

3.1. Quantitative evaluations

There were two main categories in hybrid multi-model forecasting system. The first category included four models that combined forecasts from linear and nonlinear models (ES- BPNN, ES-SCR, ARIMA- BPNN, ARIMA- SVR), and the second category consisted of one nonlinear combined forecasting model (ANFIS). These five models were tested with the revenue time series data for TFT-LCD panel manufacturers in each region. We calculated the mean square error (MSE), the mean absolute percentage error (MAPE), and the average square root error (ASRE) to compare the accuracy of these five methods. Previous studies showed that the MSE, the MAPE and the ASRE are frequently used measures of the difference between actual and predicted values for a process that is being modeled [7,28]. In statistics, the MSE, the MAPE and the ASRE are used to measure the difference between actual and predicted values and can provide a measure of the level of agreement between the observed and predicted values. The lower the values of the MSE, the MAPE and the ASRE, the closer the predicted values are to the actual values. We can calculate the MSE by squaring each of the individual errors, e_t , and taking the average of those squared values, i.e.,

$$MSE = \frac{\sum_{i=1}^n e_i^2}{n} \tag{16}$$

The MAPE is computed by dividing the absolute errors by the corresponding true values and then averaging the deviations and multiplying by 100, i.e.,

$$MAPE = \frac{\sum_{i=1}^n |X_t - F_t|/X_t}{n} \times 100 \tag{17}$$

The ASRE of a model is defined as the square root of the mean squared error, where X_{obs} are observed values and X_{model} are the predicted values at a time or place i .

$$ASRE = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \tag{18}$$

4. Results

4.1. Comparative results

Because revenues have been strongly influenced by several factors such as the financial crisis in the United States, the European debt crisis and the rapid development of the TFT-LCD industry in China, we chose the time series of revenues in the TFT-LCD markets

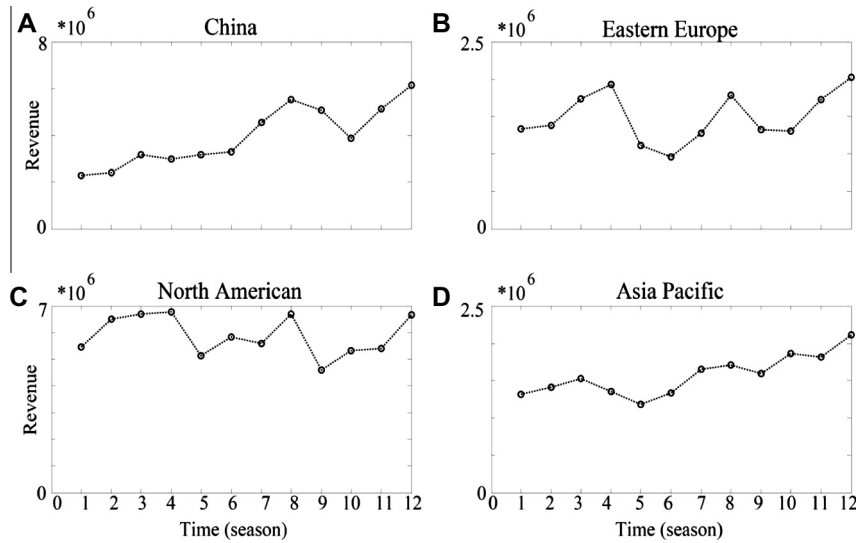


Fig. 3. The time series of revenue in the four regions.

Table 1
The statistics of the four regions.

	Mean	STD	Confidence interval (Confidence level = 95%)	Correlation coefficient
Asia	1.57E + 06	2.71E + 05	2.3852E + 006–7.6029E + 005	0.83 (Linear)
China	3.97E + 06	1.28E + 06	7.8070E + 006–1.3042E + 005	0.89 (Linear)
Eastern Europe	1.49E + 06	3.37E + 05	2.5042E + 006–4.8288E + 005	0.26 (Nonlinear)
North America	5.89E + 06	7.49E + 05	8.1374E + 006–3.6435E + 006	–0.21 (Nonlinear)

Table 2.1
Prediction errors of the ten models for the China region.

China	MSE	MAPE	ASRE
ES	1.2E + 12	22.95713	1,095,019
ARIMA	1.95E + 12	34.6286	1,396,370
SVR	9.94E + 11	21.48865	997189.3
BPNN	2.72E + 12	44.84088	1,649,544
ES_BPNN	5.43E + 12	63.23275	2,329,950
ES_SVR	6.15E + 12	68.16248	2,479,872
ARIMA_BPNN	4.68E + 10	3.795475	216426.2
ARIMA_SVR	1.05E + 12	20.88024	1,024,746
ANFIS	4.93E + 12	75.81456	2,220,437
ANFIS combination	2.76E + 12	36.27864	1,660,335

The significance of bold values indicates the mean that the best predicating model.

Table 2.2
Prediction errors of the ten models for the Eastern Europe region.

Eastern Europe	MSE	MAPE	ASRE
ES	1.72E + 11	28.18267	414752.2
ARIMA	2.53E + 11	35.52441	503,129
SVR	7.21E + 10	14.34171	268592.8
BPNN	4.58E + 10	12.52652	213904.2
ES_BPNN	2.55E + 11	30.93597	505157.3
ES_SVR	1.23E + 10	5.890067	110770.8
ARIMA_BPNN	5.84E + 11	55.00434	764046.3
ARIMA_SVR	1.99E + 11	22.96733	446616.2
ANFIS	1.98E + 11	22.77794	445174.1
ANFIS combination	1.09E + 12	132.4728	1,045,702

The significance of bold values indicates the mean that the best predicating model.

in the Asia-Pacific region, China, Eastern Europe and North America for this research. Fig. 3 shows the time series of revenue in the four regions.

Table 1 shows the mean, the standard deviation (STD), the confidence interval (CI) and the correlation coefficient for each of the four

Table 2.3
Prediction errors of the ten models for the North America region.

North America	MSE	MAPE	ASRE
ES	8.52E + 11	13.16447	922961.3
ARIMA	6.57E + 11	11.46021	810648.5
SVR	2.54E + 12	29.42889	1,592,481
BPNN	8.8E + 11	14.34333	938096.5
ES_BPNN	6.43E + 11	12.56918	801913.4
ES_SVR	1.48E + 12	20.78015	1,215,228
ARIMA_BPNN	1.74E + 12	17.04229	1,320,151
ARIMA_SVR	8.91E + 11	13.59256	944057.6
ANFIS	5.44E + 10	2.995633	233204.5
ANFIS combination	2.26E + 12	27.26805	1,504,988

Table 2.4
Prediction errors of the ten models for the Asia region.

Asia	MSE	MAPE	ASRE
ES	5.3E + 10	10.52902	230295.4
ARIMA	1.72E + 11	22.83422	415313.9
SVR	6.65E + 09	4.357999	81565.13
BPNN	1.3E + 11	22.2264	359918.1
ES_BPNN	2.41E + 10	7.533161	155244.9
ES_SVR	7.17E + 09	3.898338	84668.39
ARIMA_BPNN	7.52E + 11	54.6065	867,080
ARIMA_SVR	2.28E + 11	26.78322	477,049
ANFIS	4.44E + 11	43.9903	665994.2
ANFIS combination	4E + 10	9.811785	199905.5

regional markets [10]. The correlation coefficient was used in this study to estimate whether the data were linear or nonlinear. If the value of the correlation coefficient was between –0.3 and 0.3, the revenue was considered a nonlinear function. A common formula for calculating the correlation coefficient [8] is shown in Eq. (19).

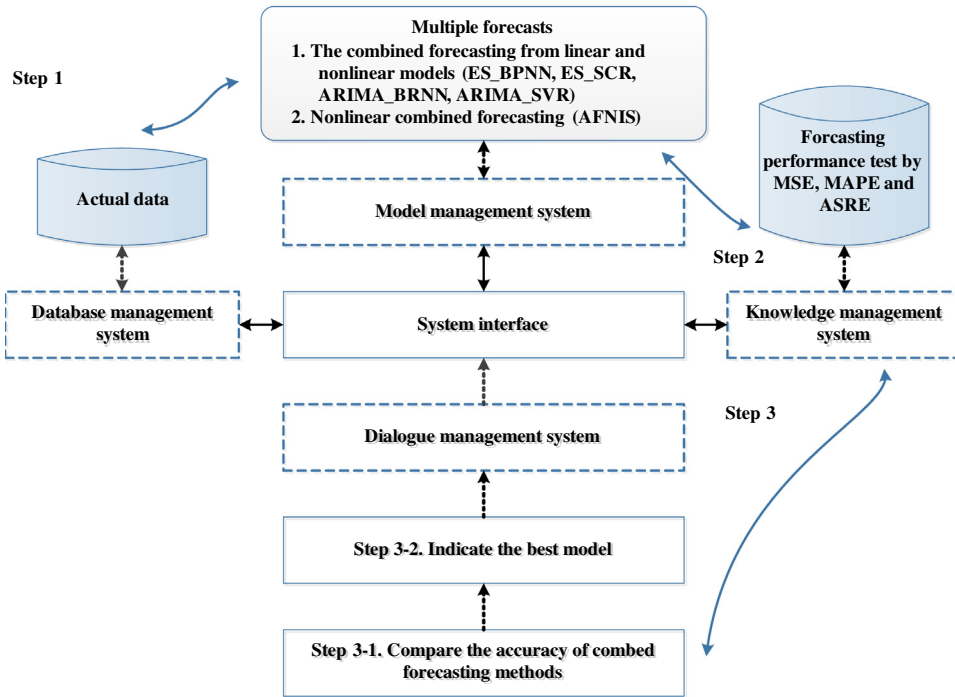


Fig. 4. The multiple-model forecasting system structure and process.

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right) \left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}} \quad (19)$$

4.2. Analysis

First, we applied five single-model forecasting methods to obtain the forecasts. Two methods, ES and ARIMA, were linear, and the other three models, SCR, BPNN and ANFIS, were nonlinear. Second, we created four combined models by pairing the linear and

nonlinear models as follows: the ES and BPNN models, the ES and SCR models, the ARIMA and BPNN models and the ARIMA and SVR models. Finally, we compared and analyzed the results using the MSE, the MAPE and the ASRE. The training period was from January 2008 to June 2011 (season 1–season 10), and the test period was from July 2011 to December 2011 (season 11–season 12). In summary, 10 different forecasting models were used to predict the revenues in the four regions. Tables 2.1–2.4 compare the results of the 10 forecasting methods in the four regions.

From the results in Tables 2.1–2.4, we can observe the effectiveness of the combined models, which had lower values of the MSE, the MAPE and the ASRE. The results showed the following: (1) for

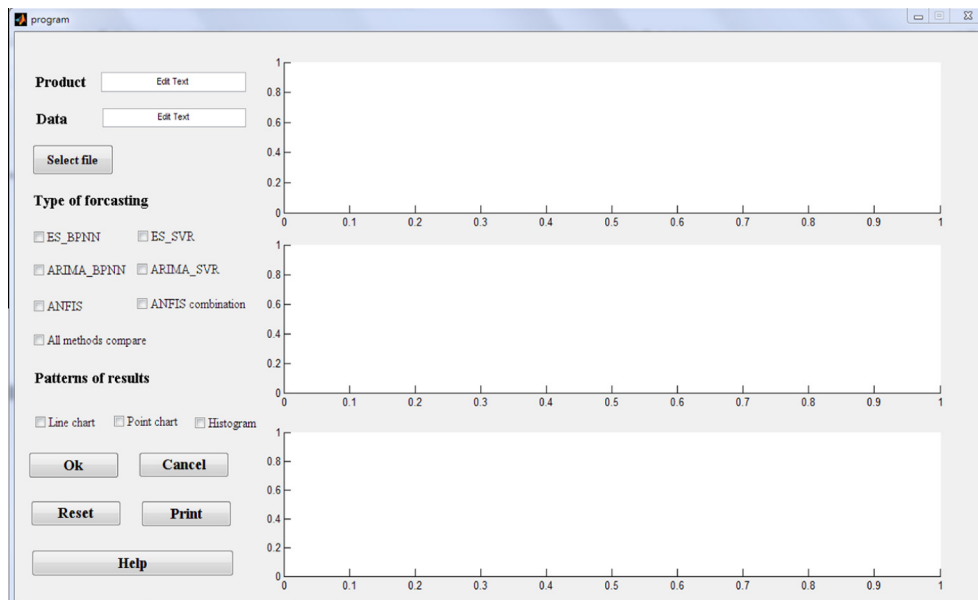


Fig. 5. System interface at start-up.

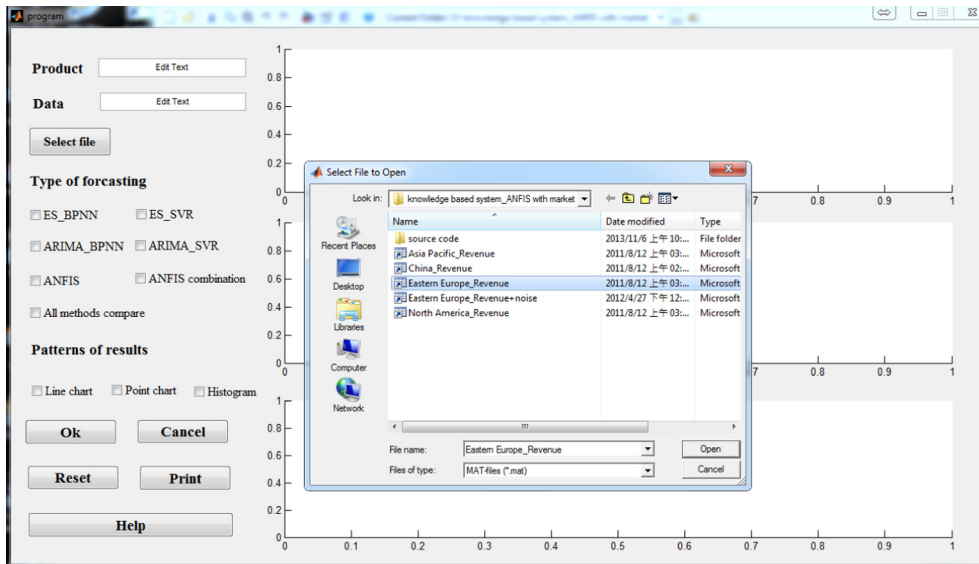


Fig. 6. System interface – importing data.

the Chinese region, the ARIMA-BPNN model was superior to the other methods; (2) for the Eastern European region, the ES-SVR model was the best; (3) for the North American region, the ANFIS was the best; and (4) for the Asia-Pacific region, the ES-SVR model was the best.

The results in Tables 2.1–2.4 showed that the best forecasting model was different for each region. However, there is substantial evidence to demonstrate that combined forecasts from linear and nonlinear models improve the forecasting accuracy. Furthermore, combinations of linear and nonlinear models were more accurate in forecasting the revenues in the four regions than the individual linear models. Research has not yet revealed the conditions or the methods for the optimal combinations of forecasts. We infer that there were differences in the information among the four regions. It follows that the extracted data characteristics and the corresponding metrics should be mapped to the forecasting perfor-

mance evaluation to construct rules for selecting the forecasting method. To take advantage of the strengths of each method to develop the best forecast, we introduce a hybrid multi-model forecasting system that can be used to select forecasting models to improve the forecasts.

4.3. Hybrid multi-model forecasting system (HMFS)

In this section, an operational multi-model forecasting system is designed to integrate all five combined forecasting methods into one forecasting process. Fig. 4 shows the structure of hybrid multi-model forecasting system. The main idea of the hybrid multi-model forecasting system is based on previous studies [41]. Several studies have shown multiple-model systems to be effective for software engineering, in which forecasting has become increasingly important [5,37]. In previous studies [22], this type of system

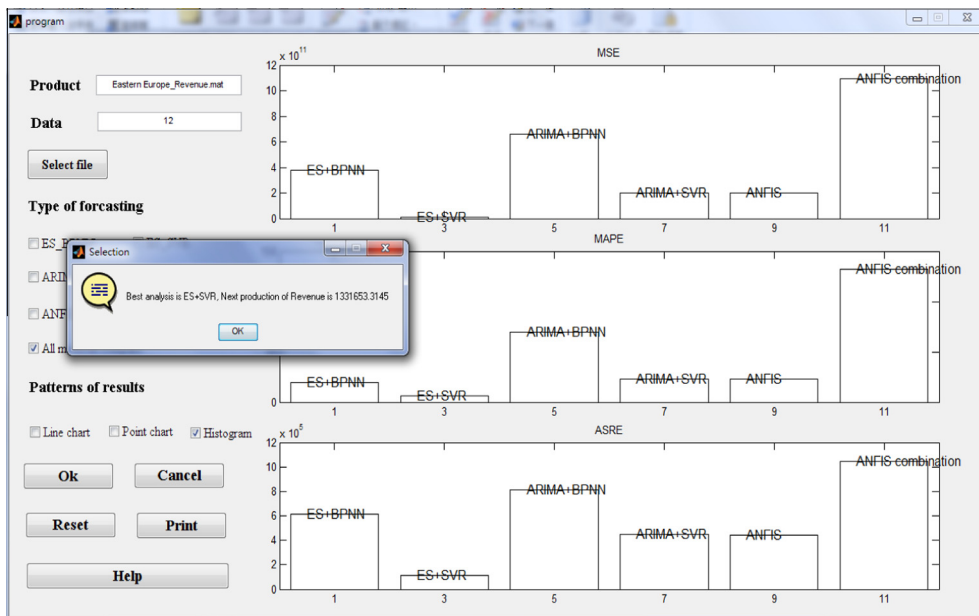


Fig. 7. System interface – model comparison.

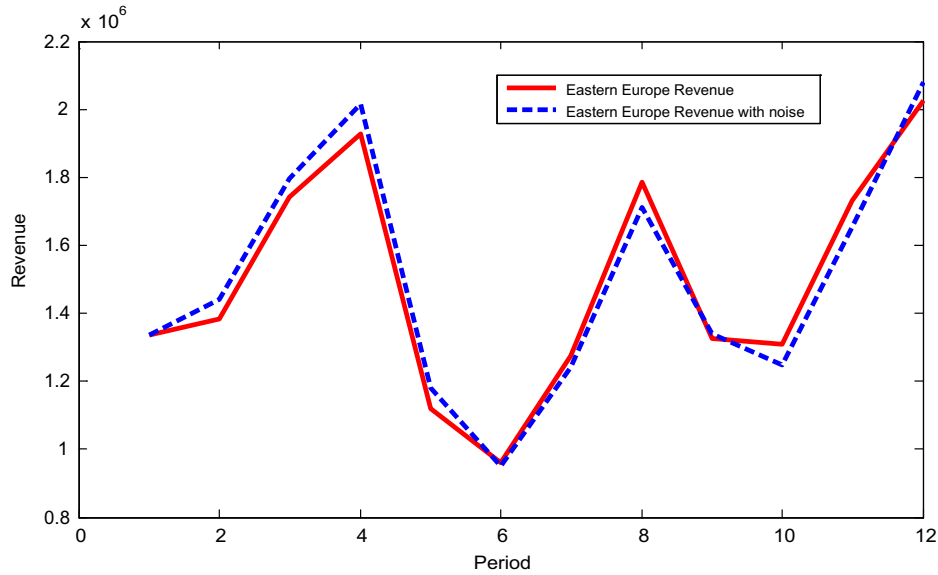


Fig. 8. The revenue forecast of the Eastern European region with and without noise.

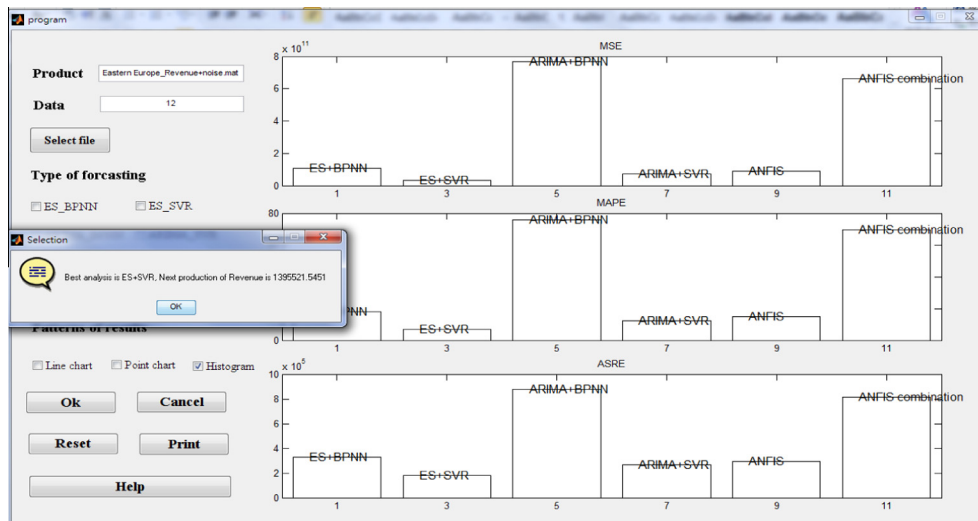


Fig. 9. The output of the hybrid multi-model forecasting system for the revenue of the Eastern European region with noise.

included information, model, knowledge and dialog management systems and combined forecasting software to obtain and analyze the data characteristics, develop the optimal forecast, and communicate the forecast to the user. All of the combined forecasting functions from the two categories were implemented in the MATLAB programming environment (MathWorks, Natick, Massachusetts, USA).

Fig. 5 shows the interface for the forecasting system software, which has a framework that integrates data management, monitoring facilities and five possible combined forecasting models with a user-friendly geographic information system (GIS) platform.

The system display for the initial step is shown in Fig. 6. Data on TFT-LCD production by region may be imported into the system and used as the input to the various combined forecasting models. The data are stored in the database of this hybrid multi-model forecasting system. The system provides automated tools to select the most accurate and reliable forecasting model for the input time series data. Fig. 5 shows that once the dataset is chosen, the system automatically attempts to load the historical data, and a separate

window allows the user to choose the forecasting mode to develop a custom combined forecasting model.

Fig. 7 shows the system interface providing guidelines to the user, typically a sales and marketing manager, in selecting suitable combined forecasting methods. The choice of the combined forecasting method depends on the forecasting performance, as determined by the MSE, the MAPE and the ASRE. The most important benefit is that the system can automatically organize and analyze large amounts of data to indicate the forecasting performance of each combined forecasting method, allowing the manager to make the final judgment. Additionally, this system provides forecasts based on past data, but the manager can adjust this forecast based on future events that the system did not consider.

An additional test of the system was conducted by adding noise (a random walk, [13,31]) to the revenue of one region (Eastern Europe) and verifying that the hybrid multi-model forecasting system could effectively forecast the revenue of that region (Fig. 8). Fig. 9 shows the output of the hybrid multi-model forecasting system for the revenue of the Eastern European region with noise. The hybrid

multi-model forecasting system selected the ES-SVR model to forecast the revenue of the Eastern European regional market with noise in the data.

5. Conclusion and further research

In fact, this research investigated the use of combinations of linear and nonlinear forecasting models to predict the revenues of the TFT-LCD industry in four regions from time series data. The results showed that the proposed models usually had better performance than the individual linear models (ES and ARIMA), the individual nonlinear models (SVR and BPNN) and the nonlinear combined forecasting model (ANFIS) in forecasting revenues in the Chinese, Eastern European and Asia-Pacific regional markets; the individual nonlinear forecasting model ANFIS had the best forecasting performance in predicting the North American market. In general, this research showed that combined forecasts from linear and nonlinear models had better performance compared with other methods.

Chen [6] reported that there is no clear evidence in favor of combined forecasts from linear and nonlinear models over other forecasting models in terms of forecasting performance [6]. Our results indicate that combined forecasts from linear and nonlinear models are better for predicting linear or nonlinear revenue time series. However, no one method was best for all series. To do so, a powerful hybrid multi-model forecasting system to generate more accurate forecasts from empirical comparisons of alternative forecasting methods was developed in this research. This hybrid multi-model forecasting system draws on several sources for forecasting inputs, including databases, documents, and a variety of forecasting methods. After processing the data from various sources, sophisticated forecasting systems integrate all the necessary data into a single spreadsheet, which the user can then manipulate by entering various projections such as different estimates of future revenue that the system will incorporate into a new output.

It is important to note the role of this flexible and sound architecture is crucial, particularly with fast-paced, rapidly developing forecasting techniques. If the base of the system is rigid or inadequate, it can be impossible to reconfigure the system to adjust to changing market conditions. Along the same lines, in other business forecasting methods and systems, it is also important to invest in systems that will remain useful over the long term, accommodating changes in the business world.

We conclude that this system has three major benefits over other forecasting systems: (1) this system has better accuracy than nonlinear combined forecasting methods or single methods; (2) this system assists in deciding which combined forecasting models are better, which alternatives are inactive, and which alternatives deliver the lowest statistical error and produce a good estimate of the variable of interest and (3) this system provides users with a graphical user interface, where the user can answer queries and can view the desired results in an integrated form. Hence, this system reduces the user's effort in obtaining the desired forecasting results from the regional revenue database.

There are many types of nonlinear models and combined forecasting methods for the prediction of markets. In this study, we considered 10 types of models and combined forecasting methods. Future work will include other types of models. Furthermore, the proposed hybrid multi-model forecasting system can be adapted to many other fields of market prediction.

Acknowledgement

This work was supported in part by National Science Council of Taiwan, Taiwan, R.O.C. under Grant NSC 103-2218-E-131-001.

References

- [1] J.M. Bates, C.W. Granger, The combination of forecasts, *OR* (1969) 451–468.
- [2] R.K. Bissoondeal, J.M. Binner, M. Bhuruth, A. Gazely, Forecasting exchange rates with linear and nonlinear models, *Glob. Bus. Econ. Rev.* 10 (2008) 414–429.
- [3] L. Breiman, Bagging predictors, *Mach. Learn.* 24 (1996) 123–140.
- [4] D.W. Bunn, A Bayesian approach to the linear combination of forecasts, *Operat. Res. Quart.* (1975) 325–329.
- [5] P.-C. Chang, C.-P. Wang, B.J. Yuan, K.-T. Chuang, Forecast of development trends in Taiwan's machinery industry, *Technol. Forecast. Soc. Chang.* 69 (2002) 781–802.
- [6] K.-Y. Chen, Combining linear and nonlinear model in forecasting tourism demand, *Expert Syst. Appl.* 38 (2011) 10368–10376.
- [7] F. Collopy, J.S. Armstrong, Rule-based forecasting: development and validation of an expert systems approach to combining time series extrapolations, *Manage. Sci.* 38 (1992) 1394–1414.
- [8] T.R. Derrick, B.T. Bates, J.S. Dufek, Evaluation of time-series data sets using the Pearson product-moment correlation coefficient, *Med. Sci. Sports Exerc.* 26 (1994) 919–928.
- [9] V. Fernandez, Wavelet-and SVM-based forecasts: an analysis of the US metal and materials manufacturing industry, *Resour. Policy* 32 (2007) 80–89.
- [10] T.A. Ferreira, G.C. Vasconcelos, P.J. Adeodato, A new intelligent system methodology for time series forecasting with artificial neural networks, *Neural Process. Lett.* 28 (2008) 113–129.
- [11] A. Fiordaliso, A nonlinear forecasts combination method based on Takagi-Sugeno fuzzy systems, *Int. J. Forecast.* 14 (1998) 367–379.
- [12] P.H. Franses, R. Legerstee, Combining SKU-level sales forecasts from models and experts, *Expert Syst. Appl.* 38 (2011) 2365–2370.
- [13] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *Comput. Learn. Theory* 904 (1995) 23–37.
- [14] A. Goh, Back-propagation neural networks for modeling complex systems, *Artif. Intell. Eng.* 9 (1995) 143–151.
- [15] S.I. Gunter, Nonnegativity restricted least squares combinations, *Int. J. Forecast.* 8 (1992) 45–59.
- [16] G. Guo, M.E. Roettger, T. Cai, The integration of genetic propensities into social-control models of delinquency and violence among male youths, *Am. Sociol. Rev.* 73 (2008) 543–568.
- [17] D. Hernández-Lobato, G. Martínez-Muñoz, A. Suárez, Empirical analysis and evaluation of approximate techniques for pruning regression bagging ensembles, *Neurocomputing* 74 (2011) 2250–2264.
- [18] K.-T. Hsu, Using a back propagation network combined with grey clustering to forecast policyholder decision to purchase investment-linked insurance, *Expert Syst. Appl.* 38 (2011) 6736–6747.
- [19] J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, *IEEE Trans. Syst., Man Cybernet.* 23 (1993) 665–685.
- [20] H. Kelley, J. Busemeyer, A comparison of models for learning how to dynamically integrate multiple cues in order to forecast continuous criteria, *J. Math. Psychol.* 52 (2008) 218–240.
- [21] C.-C. Lin, C.-L. Lin, J.Z. Shyu, C.-T. Lin, The ANFIS system for nonlinear combined forecasts in the telecommunications industry, *Int. J. Comp. Appl.* (2012) 37.
- [22] C.-L. Lin, C.-C. Lin, J.Z. Shyu, C.-T. Lin, A rule-based forecasting system integrating combining and single forecast for decision making, *Int. J. Comp. Appl.* 39 (2012) 1–6.
- [23] C.-T. Lin, S.-Y. Yang, Forecast of the output value of Taiwan's opto-electronics industry using the grey forecasting model, *Technol. Forecast. Soc. Chang.* 70 (2003) 177–186.
- [24] K.-P. Lin, P.-F. Pai, A fuzzy support vector regression model for business cycle predictions, *Expert Syst. Appl.* 37 (2010) 5430–5435.
- [25] G.J. Lobo, R. Nair, Combining judgmental and statistical forecasts: an application to earnings forecasts, *Dec. Sci.* 21 (1990) 446–460.
- [26] J.P. Martino, *Technological Forecasting for Decision Making*, McGraw-Hill, Inc, 1993.
- [27] H.-T. Pao, Comparing linear and nonlinear forecasts for Taiwan's electricity consumption, *Energy* 31 (2006) 2129–2141.
- [28] P.C. Pendharkar, A threshold varying bisection method for cost sensitive learning in neural networks, *Expert Syst. Appl.* 34 (2008) 1456–1464.
- [29] F.M. Scherer, The office of technology assessment and forecast industry concordance as a means of identifying industry technology origins, *World Patent Inf.* 4 (1982) 12–17.
- [30] G.H. Shakouri, R. Nadimi, F. Ghaderi, A hybrid TSK-FR model to study short-term variations of the electricity demand versus the temperature changes, *Expert Syst. Appl.* 36 (2009) 1765–1772.
- [31] R. Sitte, J. Sitte, Neural networks approach to the random walk dilemma of financial time series, *Appl. Intell.* 16 (2002) 163–171.
- [32] J.H. Stock, M.W. Watson, *A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series*, Oxford University Press, 1998, pp. 1–44.
- [33] J.W. Taylor, Exponential smoothing with a damped multiplicative trend, *Int. J. Forecast.* 19 (2003) 715–725.
- [34] N. Terui, H.K. Van Dijk, Combined forecasts from linear and nonlinear time series models, *Int. J. Forecast.* 18 (2002) 421–438.
- [35] F.-M. Tseng, Y.-C. Hu, Quadratic-interval Bass model for new product sales diffusion, *Expert Syst. Appl.* 36 (2009) 8496–8502.

- [36] F.-M. Tseng, H.-C. Yu, G.-H. Tzeng, Combining neural network model with seasonal time series ARIMA model, *Technol. Forecast. Soc. Chang.* 69 (2002) 71–87.
- [37] K. Vinay Kumar, V. Ravi, M. Carr, N. Raj Kiran, Software development cost estimation using wavelet neural networks, *J. Syst. Softw.* 81 (2008) 1853–1867.
- [38] R.J. Vokurka, B.E. Flores, S.L. Pearce, Automatic feature identification and graphical support in rule-based forecasting: a comparison, *Int. J. Forecast.* 12 (1996) 495–512.
- [39] F.-K. Wang, K.-K. Chang, Adaptive neuro-fuzzy inference system for combined forecasts in a panel manufacturer, *Expert Syst. Appl.* 37 (2010) 8119–8126.
- [40] F.-K. Wang, K.-K. Chang, C.-W. Tzeng, Using adaptive network-based fuzzy inference system to forecast automobile sales, *Expert Syst. Appl.* 38 (2011) 10587–10593.
- [41] W. Wang, S. Nie, The performance of several combining forecasts for stock index, in: *2008 International Seminar on Future Information Technology and Management Engineering*, 2008, pp. 450–455.
- [42] G.P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing* 50 (2003) 159–175.