

Enhancing Growth Curve Approach Using CGPANN for Predicting the Sustainability of New Food Products

Jawad Ali, Gul Khan, Sahibzada Mahmud

► **To cite this version:**

Jawad Ali, Gul Khan, Sahibzada Mahmud. Enhancing Growth Curve Approach Using CGPANN for Predicting the Sustainability of New Food Products. Lazaros Iliadis; Ilias Maglogiannis; Harris Papadopoulos. 10th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Sep 2014, Rhodes, Greece. Springer, IFIP Advances in Information and Communication Technology, AICT-436, pp.286-297, 2014, Artificial Intelligence Applications and Innovations. <10.1007/978-3-662-44654-6_28>. <hal-01391325>

HAL Id: hal-01391325

<https://hal.inria.fr/hal-01391325>

Submitted on 3 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Enhancing Growth Curve Approach using CGPANN for Predicting the Sustainability of New Food Products

Jawad Ali¹, Gul Muhammad Khan² and Sahibzada Ali Mahmud³

^{1,2,3}Department of Electrical Engineering,
University of Engineering and Technology Peshawar, Pakistan 25000.

¹ jawad.ali@nwfpuet.edu.pk

² gk502@nwfpuet.edu.pk

³ sahibzada.mahmud@nwfpuet.edu.pk

Abstract. *An enhancement to the growth curve approach based on neuro evolution is proposed to develop various forecasting models to investigate the state and worth of the producer, to market a new product. The forecasting model is obtained using a newly introduced neuro evolutionary approach called Cartesian Genetic Programming based ANN (CGPANN). CGPANN helps in obtaining an optimum model for all the necessary parameters of an ANN. An accurate and computationally efficient model is obtained, achieving an accuracy as high as 93.37% on the time devised terrains, providing a general mechanism for forecasting models in mathematical agreement to its application in econometrics. Comparison with other contemporary model evidences the perfection of the proposed model thus its vital power in developing the growth curve approach for predicting the sustainability of new products.*

Keywords: Growth curve approach, Cartesian Genetic Programming (CGP), Artificial Neural Network (ANN), Product sustainability

1 Introduction

Marketing a product needs knowledge about the anticipation of the end user and forecasts make it possible to respond to these changes. Different types of forecasts are made in order to mend the stability of a firm in the market including: demand forecast, supply forecast, sales forecast, production forecast, profit forecast, advertisement and promotional expense forecast.

Food products are always in the phase of innovation so that the user remains adhered to the products of a specific company. It is the job of a food company to know the exact limits of demand for a new product before its commissioning, machinery planting and management of other capital resources [10]. Trends in the previous products and their variation in different ingredients results a timely development or evolution of new products. Firms having past experience and marketing setup therefore always keep using previous knowledge for establishing

a new product, that is a resultant of some alteration in the present or previously developed product(s).

Data from the previous products that correlates the developing products is of great importance while dealing demand forecast. Without proper formulation of the past facts, the theories constituted for product forecast are meaningless¹. Economists use different traditions strategies and probabilistic models to evaluate the core consumption of a product before its initiation [2,3]. The general concept of economists circle around one point that is the net income from a product which is to be commercialized [9]. So the financial market can be studied completely by its two major aspects: fundamental and technical analysis. The fundamental analysis involves the study of economic factors of the product on a firm, while the technical analysis has to do with the future worth of a product and its sustainability. Joel Dean [4] has explained the achievement of these two goals in a six point formula by exploring evolutionary approach, substitution approach, opinion poll approach, sales experience approach, growth curve approach and vicarious approach.

This paper is using the growth curve approach [4] for the prediction of the sustainability of new products. The uniqueness in the work is that the growth curve of the future sales and stock is used for the prediction of the sustainability of the product by utilizing a neuro evolutionary approach of Cartesian Genetic programming evolved ANN (CGPANN). The paper is organized in the following manner. Section 2 provides an overview of the past work done in the field. It explains factors involved in product design and prior marketing steps. Section 3 gives an insight of the basic strategies followed in the development and commercialization of market products whereas section 4 covers the concept of CGP evolved ANN and their implementation mechanism. The Experiments, Results confirmed by the experiments and concluding remarks are added in later sections of the paper.

2 Literature Review

Expected future market of products and financial time series forecasting has been explored in the past using various methods including econometric, statistical, probabilistic, expert systems and various other ANN based techniques. The most popular econometric model is the logit model [13], based on qualitative analysis of the product. Mixed logit with repeated choices that is an extension of the logit model dealing with the dependencies within each nest is proposed in [16]. The logit model and nested logit [1] model neglects the dependency of the prediction by using independence from irrelevant alternatives property. [3] deals the innovative perspective of new product by measuring the impact of society in choosing the right size of automobile and issues like downsizing of cars etc. [2] deals real world cases, including stock problem and inventory issues, while applying statistical approaches for making estimation about the consequences

¹ <http://www.statsoft.com/Textbook/Demand-Forecastin>

and likelihood of some future events. Fuzzy logic based production time-lining is developed in [15] for rescheduling of the production in the presence of varying perturbations.

Neural Networks are always comparable to other prediction models when it comes to time series analysis. Back-propagating Neural Network model is used in [12] for forecasting the demand of sea food material demand. SVM is used for financial time series forecasting in [8]. Financial time series analysis has been covered in [6] using Support Vector Machine (SVM) with modified features. The prediction model in [5] uses neural networks while it takes time series of bankruptcy data, predicting likelihood of bankruptcy. The stock market forecaster design is also obtained using Modular Neural Network in [5]. Neural Networks are employed in stock market for future trading volumes in [9]. This work is different from the past applications of the Neural Networks to the industrial development in the sense that it uses the trading volume and the change in the market value of the production of the firms for predicting the sustainability of the firm. This whole technique is collectively termed as growth curve approach.

3 New product development - An insight

Product design is the core issue for the deployment of new product in the market. New product should be consummate in terms of its function, environment, cost, ergonomics, material, customers need and industrial production feasibility². It should be unique and carry the identity of the producer in case of cash-cow³ companies. The more relevance of the product to the discussed attributes makes it more user friendly, affordable and attractive. In case, the product that is to be commercialized, is the renovation of some already existing product(s) or product that were developed in the past, the deployment should follow the production and consumption trends of these previous versions of the products. Forecasting the peak consumption era depends on the need of a product, for which the producer must be ready to place the product in the market at the exact time for the exact duration. The key discrepancies can be removed easily if the products are ready in their development.

Investigating the company's ability to introduce a successful product can be done by analyzing the past statistics of the firm [8]. The variation in the production of the company can be studied from the daily market values, variation in the sales with time and the total development to the date of announcing new products. The behavior of the time series is the resultant of the features that constitutes the growth of a product or company. This particular time series is thus used for predicting the near future and ensembles of the product profile.

² <http://www.deaconlloyd.com/services/market-feasibility-studies.html>

³ http://www.wikinvest.com/wiki/Cash_cow_business

4 Cartesian Genetic Programming evolved Artificial Neural Network (CGPANN)

The dynamicity of Artificial Neural Network is enhanced using different evolutionary techniques. These techniques including NEAT [7], TWEANN [17] and co-evolutionary method including adaptive and self-adaptive differential evolution systems [18]. The application space of ANN is increased when it is evolved using CGP.

Architectural overview of CGPANN A typical CGP evolved ANN consists of CGP Nodes connected with each other and to the inputs/output(s). The nodes or neurons thus are used as processing elements. A node can either be in input layer, it can form hidden layer or it can be a part of the output layer. In feed-forward CGPANN, this node takes its input either from the preceding sister nodes or from the system input, depending upon its location in the network. Each node takes a defined number of inputs that is specified in the initialization step for the network evolution. These input are then scaled (defined by Eq. 1) and summed before passing through an activation function.

$$\forall W_\varphi \in PRG([1, -1]) \wedge \varphi = [1, 2, 3, \dots, n] \quad (1)$$

φ is the total number of input connections to a node and the weights that are generated by PRG are assigned to these input connections respectively. All the connections are scaled before applying to a random node. Input nodes take their inputs from the system inputs directly without scaling. Intermediate nodes take its input(s) from preceding nodes in case of feed forward ANNs. The activation $f(x)$ is chosen to be log-sigmoid for all nodes in this network, defined by the Eq. 2. Junk nodes have no connectivity hence contribution to the network output.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

Here, x is the sum of scaled inputs of the randomly chosen inputs. The x is given by the Eq. 3

$$x = \sum_{i=1}^n [I_i] \times [W_i] \quad (3)$$

Where the number of inputs ranges from 1 to n , given by matrix $[I]$ in Eq. 4. The matrix $[W]$ is the collection of respected random weights to the node, given by Eq. 5. Eq. 3 adds the product of all inputs and their respected weights.

$$I = [i_1, i_2, i_3, \dots, i_n] \quad (4)$$

$$W = [w_1, w_2, w_3, \dots, w_n] \quad (5)$$

Any summing junction in ANN can be represented by

$$y' = \sum_{i=1}^N x_i \quad (6)$$

where x_i is the input to the junction. If the same input is scaled with a randomly assigned weight w_i then Eq. 6 takes the form

$$y' = \sum_{i=1}^N x_i w_i \quad (7)$$

Let for N inputs to a node, we have y_j output such that

$$y_j = f^j(y'_j) = f^j\left(\sum_{i=1}^N x_i w_i\right) \quad (8)$$

Here f is an activation function, particular to node j . If total number of nodes in the network are N_T then j is defined by

$$\{j | j \in N \wedge 1 \leq j \leq N_T\} \quad (9)$$

Let I be the input set to a unique genotype network G_k , consisting unique entries i such that, for all inputs

$$i \in R \wedge 0 \leq i \leq 1 \quad (10)$$

in

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

so, network G_k is the a set of random selection from inputs $[I]$, outputs of nodes y_j , for a single output O_p such that

$$O_p = \frac{1}{n} \sum_{i=1}^N \left(f \left(\sum (y_j W_j + y_{j-1} W_{j-1} + \dots + y_1 W_1 + I W_k) \right) \right) \quad (11)$$

here, W_j, W_{j-1}, \dots, W_1 are Random subsets of W such that W_k is a subset W_j and

$$W_j = \{w_k | w_k \in R \wedge -1 \leq w_k \leq 1\} \quad (12)$$

Now

$$G_k = \{I, y_j, y_{j-1}, \dots, y_1, O_p\} \quad (13)$$

$$\begin{aligned} y_j &= \left(f \left(\sum (y_{j-1} W_{j-1} + y_{j-2} W_{j-2} + \dots + y_1 W_1 + I W) \right) \right) \\ y_{j-1} &= \left(f \left(\sum (y_{j-2} W_{j-2} + y_{j-3} W_{j-3} + \dots + y_1 W_1 + I W) \right) \right) \\ &\vdots \\ y_2 &= \left(f \left(\sum (y_1 W_1 + I W) \right) \right) \\ y_1 &= \left(f \left(\sum I W \right) \right) \end{aligned} \quad (14)$$

Let G_k and G_l be the two successive genotypes. G_l is produced from G_k by mutating $\mu\%$ weights of connections in G_k , nodal connectivity, functions defining each node or the combination of these.

Let the total entries in G_k , including weights, connections and functions be N_{cwf} then

$$N'_{cwf} = \mu \times N_{cwf} \quad (15)$$

N'_{cwf} represents genotype entries that are to be randomized to get G_l .

Let r be a unique entry to the set ζ that contains value to be changed to get mutated G_k or G_l defined by r .

$$\{r | r \in \zeta \forall (W, I, y_j)\} \quad (16)$$

each value in r is defined by a Pseudo Random Generator (PRG) that takes N'_{cwf} values from available N_{cwf} entries using the relation

$$\left\{ r_i | r_i \in \{1, 2, 3, \dots, N_{cwf}\} \wedge r_i \subset \{1, 2, 3, \dots, N_{cwf}\} \right\}, i = 1, 2, 3, \dots, N'_{cwf} \quad (17)$$

where r_i belongs to $\{1, 2, 3, \dots, N_{cwf}\}$ so each entry in ζ is

$$\Gamma(i) = \left\{ r_i | r_i \in \{1, 2, 3, \dots, N_{cwf}\} \wedge \{1 \leq r_i \leq N_{cwf}\} \right\} \quad (18)$$

The value that is replaced in G_k for N'_{cwf} values uses Pseudo Random Number Generator (PRG). Here, each input I_i is the output of some preceding node or the network input given by the Eq. 19.

$$I(G, L, N) = PRG([O(G, L_p, N_p)] : [I(G, L_0)]) \quad (19)$$

G represent a specific genotype, L is the Layer in which the node N is situated. L_p and N_p are representing the preceding Layers and Nodes respectively whereas L_0 represent the input matrix to the Genotype G .

The mutation rate is kept 10% for its better results in [11]. In each generation, connections are altered and weight are varied according to the relevance of the network output with the desired values that are also given to the network for its training. Junk nodes that contribute none to the network and having no influence on the network performance are excinded using CGP in an efficient way. Thus the final phenotype might have node number less than the number of nodes that are initialized at the start of evolutionary phase. The data is taken for three different dairy product producing companies i.e. Nestle (NSRGY)⁴, Dean Food Company (DF)⁵ and Saputo (SAPIF).

The growth curve estimation model works in two major steps. The first step involves the training of the model. The daily stock market values of NSRGY for the years 1998-99 are used to train the particular CGPANN model. The model

⁴ <http://money.rediff.com/companies/Nestle-India-Ltd/11120007/nse/day/chart>

⁵ <https://finance.yahoo.com/q?s=DF>

develops an intrinsic intuition for predicting the next half day stock value while taking 7 days data as its input sequence. **10** different networks, on the basis of the number of their nodes, were initialized for this purpose. The number of nodes were varied between 50 and 500 with a step size of 50. Inputs are taken from the provided stock value time series to the network where the CGPANN calculates the output $V_{forecasted}$. This output is compared with the given desired value V_{actual} and the generation proceeds to minimize this difference between the output from the network and the desired output. The evolution stops after reaching to the maximum number of generation, initialized at the start of this training session or if the desired fitness is achieved. The fitness of 10 mutants is measured at the end of each generation using Eq. 20. This training phase took 20 hours when performed with 3.4GHz corei7-2600 processing unit with 16GB RAM and 8MB L3 Cache.

$$Fitness = 100\% - MAPE \quad (20)$$

Where MAPE is the mean absolute percentage error that is calculated as

$$MAPE = \left| \frac{V_{Actual} - V_{forecasted}}{V_{Actual}} \right| \times 100\% \quad (21)$$

After selecting the best mutants among the evolved genotypes, these networks are evaluated with the time series data of normalized stock values, change in stock and total sales to data. Sliding window mechanism is used for the prediction of whole sequence. The model takes **14** inputs and gives forecast value for the nested inputs.

5 Experiment Results

Selecting network: MAPE values of the networks for the years 1998-1999 given for each network in Table 1. Table 1 is the numerical evidence of the fact that although all networks are giving a minimal MAPE due to their better adaptability features while each network is evolved for its 1M generations, the network that is initialized with 450 nodes outperforms over other networks. Although the

Table 1. MAPE values for training session - 14 instants input

	Single	7	14		Single	7	14
Nodes	output	outputs	outputs	Nodes	output	outputs	outputs
50	1.8016	2.5912	2.681	300	1.8045	2.6114	2.694
100	1.8046	2.6059	2.692	350	1.8109	2.6222	2.705
150	1.8107	2.6000	2.687	400	1.8040	2.6228	2.703
200	1.8028	2.5981	2.703	450	1.7992	2.6041	2.694
250	1.8072	2.5929	2.687	500	1.8198	2.6018	2.713

error increases when the forecasting horizon is elevated yet the model is useful

as this MAPE value is competent to other models quoted in the literature. Table 2 contains MAPE results for the year 2002 for different companies with various data sets. The model takes 14 instant inputs and gives single output. Results are tabulated with respect to the initialized nodes for network.

Table 2. MAPE results for Stock value, Change in stock and Cumulative stock with network inputs ranging from 50 to 500 with step size of 50

Nodes	50	100	150	200	250	300	350	400	450	500
SAPIF(S)	6.087	5.372	5.588	5.478	5.523	5.685	6.105	5.788	5.365	5.730
SAPIF(CH)	12.806	11.101	11.810	11.441	11.535	11.994	13.094	11.970	11.042	12.08
SAPIF(CS)	9.334	8.104	8.519	8.261	8.333	8.666	9.386	8.843	8.089	8.750
DF(S)	7.624	6.633	7.000	6.806	6.864	7.117	7.709	7.186	6.624	7.178
DF(CH)	11.603	10.068	10.701	10.370	10.456	10.869	11.772	10.877	10.032	10.96
DF(CS)	9.310	8.069	8.514	8.254	8.327	8.661	9.38	8.802	8.064	8.746
NSRGY(S)	7.861	6.875	7.215	7.023	7.084	7.342	7.898	7.459	6.867	7.417
NSRGY(CH)	6.663	6.4366	6.528	6.462	6.449	6.542	6.554	6.564	6.382	6.521
NSRGY(CS)	8.949	7.7498	8.182	7.934	8.003	8.323	9.010	8.451	7.747	8.402

here: S=Stock value, CH=Change in stock and CS=Cumulative stock
The Standard deviation and Peak to Average ratio (P/A) values in Table 3

Table 3. Standard deviation and Peak to average ration of the data sets

Attribute	DF(S)	DF(CH)	NSRGY(S)	NSRGY(CH)	SAPIF(S)	SAPIF(CH)
P/A	1.766	1.282	1.799	1.825	2.008	1.226
SD	0.234	0.035	0.225	0.0775	0.206	0.0440

describe the distribution of the data on the given time line. It is the beauty of the model to predict up to the excellent level of accuracy as given in Table 2. It is evident from the Table 2 that different data sets (of exclusive data trends) gives best results on network with 450 nodes. The best results obtained in all the models for the year 2002 for different data trends are given by in Table 4. The

Table 4. MAPE results for different data sets - 450 initial nodes - 14 instances input - Single output

Time series	Company		
	SAPIF(%)	DF(%)	NSRGY(%)
Stock Value	5.365	6.624	6.867
Change in Stock	11.042	10.032	6.382
Total Sales to date	8.089	8.064	7.747

comparison of the proposed CGP evolved ANN with different other forecaster models is performed in Table 5.

Table 5. Comparison of proposed CGPANN model with different models[6,14]

Testing Accuracy		Testing Accuracy	
Model(s)	(%age)	Model(s)	(%age)
MDA	79.13	RBF SVM	83.06
Logit	78.30	Polynomial SVM(1)	82.01
BPNN(1)	81.48	Polynomial SVM(2)	77.24
BPNN(2)	71.16	Sigmoid SVM	71.16
Linear SVM	77.24	CGPANN	93.37

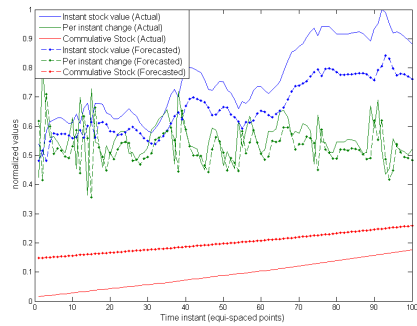


Fig. 1. Nestle - Instantaneous stock value, per instant change and cumulative stock prediction (14 inputs(7 days) 1 output)

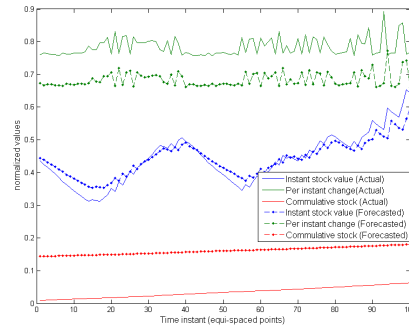


Fig. 2. Dean - Instantaneous stock value, per instant change and cumulative stock prediction (14 inputs(7 days) 1 output)

It can be inferred from Table 5 that CGPANN is a better techniques that can be used to evolve Artificial Neural Network for human biased certainty reluctant time series such as stock value and productions due to its self-adaptability and fast learning features. CGPANN, in comparison with other prediction models, performs better for its non-linear characteristics and non-conventional behavior. The non-conventional behavior of CGPANN increases its MAPE to 93% that is 22.21% in addition to the MAPE value of Sigmoid SVM, a non-linear model, conventional model. On the other hand, the Linear SVM that follows linear approach and possesses conventional behavior, gives 6.08% more effective MAPE values for a given timely devised data.

Fig. 1, 2 and 3 graphically explain the accuracy of the model by plotting the estimated stock values, instantaneous changes and the cumulative stocks of each company. The Fig. 1, 2 and 3 explain progress of each producer as well as the

change in their market stock time by time. It can be inferred from the estimated and the actual values of the entities that the proposed forecaster has learned the pattern that is experienced by each producer. The time series thus obtain is authenticated one and can be used for econometric analysis.

Cumulative stock or the growth curve defines the exact position of the firm after each and every time instant which has been made possible by using the newly developed CGPANN. Basic fact about SNRGY food producing company can be explained from its cumulative stock value. Based on cumulative stock, SNRGY is leading among its competitors, shown in Fig. 4. The variations in the market stock value (defined by peak to average ratio and standard deviation) is also compared in Fig. 4. More fluctuation in the time series of the per-instant change of the company stock, as given by Table 3, reflects the *dynamicity* in the products of the company.

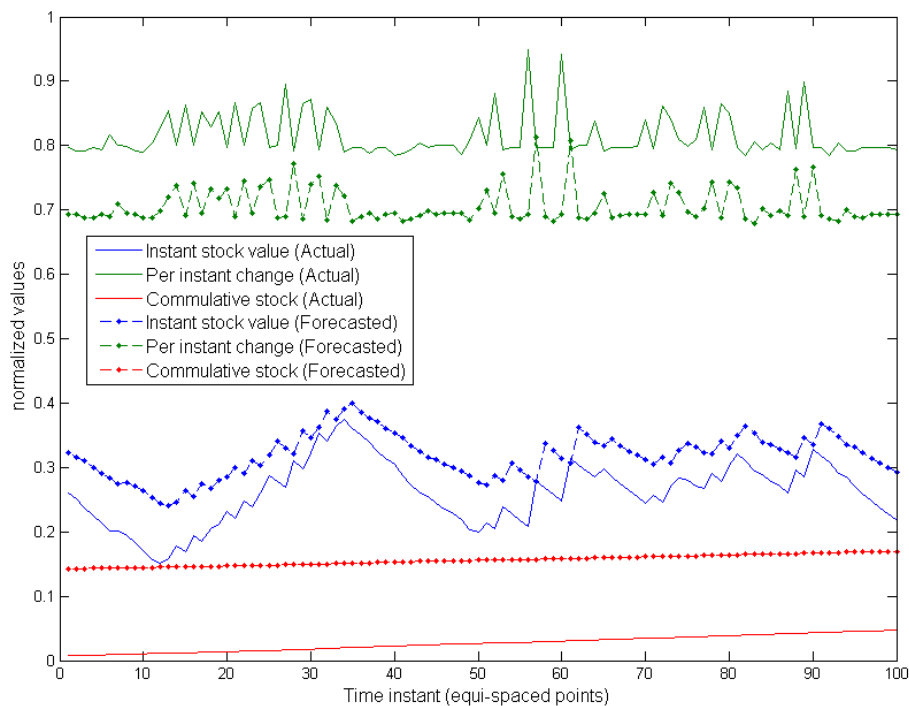


Fig. 3. Saputo - Instantaneous stock value, per instant change and cumulative stock prediction (14 inputs(7 days) 1 output)

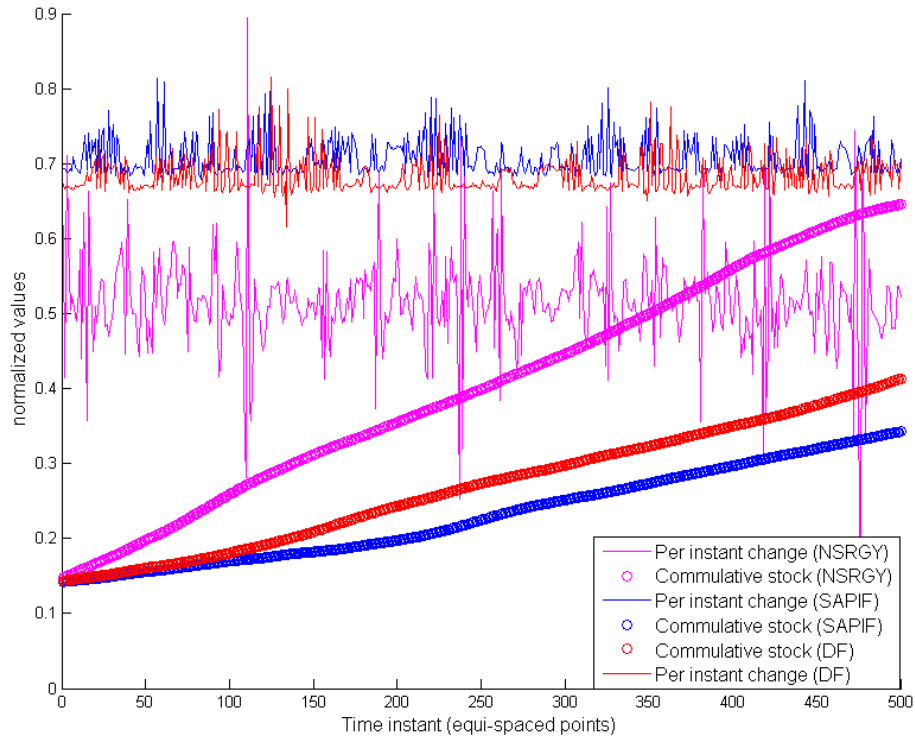


Fig. 4. Predicted growth curve using CGPANN (initialized with 450 nodes)

6 Conclusion

The paper uses CGPANN prediction model for forecasting the growth curve of a firm. This growth curve thus enables a particular producer, at any instant, to compare its production statistics with other similar producer and find out the margin to which they have to take risk in introducing new product in the market. The buying power of the consumer can be predicted by looking at the daily change in the stock usage that can be used in the estimation of the future production by a company. Thus the companies can estimate the raw material and manage their stock according to the future buying capacity of the present users. The model further enables newly developed firms to calculate their increment in market reputation while looking at the daily change in stock value. The model can be much effective if the consumption time series of a single product is employed for its future use. The producers can then manage their inventory while knowing the exact number of future consumer of each and every single product, supplied by them.

References

1. Ben-Akiva, M., Bolduc, D.: Multinomial probit with a logit kernel and a general parametric specification of the covariance structure. Working paper, Department of Economics, MIT (1996)
2. Brownstone, D., Train, K.: Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics* pp. 109–129
3. Cardell, N., Dunbar, F.: Measuring the societal impacts of automobile downsizing. *Transportation Research* 14(5,6), 423–434
4. Dean, J.: Demand forecasting for a new product, http://www.entranceguruji.in/read_matirial.php
5. Fletcher, D., Goss, E.: Forecasting with neural networks and application using bankruptcy data. *Information and Management* 24, 159–167 (1993)
6. Francis, E.H., L., C.: Modified support vector machines in financial time series forecasting. *Neurocomputing* 48, 847–861 (2002)
7. Hasanat, A.: Object class recognition using neat-evolved artificial neural network. Fifth International Conference on Computer Graphics, Imaging and Visualization, CGIV pp. 271–275 (2008)
8. J.Cao, L., Tay, H.: Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks* 14(6), 1506–1518 (2003)
9. Kaastra, I., Milton, S.: Forecasting futures trading volume using neural networks. *J. Futures Markets* 15, 853–970 (1995)
10. Kempf, K.G., Keskinocak, P., Uzsoy: Planning production and inventories in the extended enterprise. *International Series in Operations Research and Management Science* 152(2), 588–589 (2011)
11. Khan, G.M., Khan, S., Ullah., F.: Short-term daily peak load forecasting using fast learning neural network. *Intelligent Systems Design and Applications (ISDA)* pp. 843–848 (2011)
12. Lo, C.Y.: Back propagation neural network on the forecasting system of sea food material demand. In: Zhou, M., Tan, H. (eds.) *Advances in Computer Science and Education Applications, Communications in Computer and Information Science*, vol. 202, pp. 147–154. Springer Berlin Heidelberg (2011)
13. McFadden, D.: Conditional logit analysis of qualitative choice behavior. *Frontiers in econometrics*. Academic Press, New York (1973)
14. Min, J.H., Lee, Y.C.: Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters 48, 847–861 (2002)
15. Petrovic, D., Duenas, A.: A fuzzy logic based production scheduling/rescheduling in the presence of uncertain disruptions. *Fuzzy Sets and Systems* 157(16), 2273–2285 (2006)
16. Revelt, D., Train, K.: Mixed logit with repeated choices: Households choices of appliance efficiency level. *Review of Economics and Statistics* 80(4) (1998)
17. Yao, X.: Evolving artificial neural networks. In *Proceedings of the IEEE* 87(9), 1423–1447 (1999)
18. Yao, X., Islam, M.M.: Evolving artificial neural network ensembles. *Computational Intelligence Magazine, IEEE* 3(1), 31–42 (2008)