DESIGNING A FORECAST MODEL FOR ECONOMIC GROWTH OF JAPAN USING COMPETITIVE (HYBRID ANN VS MULTIPLE REGRESSION) MODELS

Ahmet DEMIR
Ishik University, Erbil, Iraq,
ahmet.demir@ishik.edu.iq
AtabekSHADMANOV
Ishik University, Erbil, Iraq,
atabek.shadmanov@ishik.edu.iq
CumhurAYDINLI
Ipek University, Ankara, Turkey,
cumhur_a@hotmail.com
Okan ERAY
International Black Sea University, Tbilisi, Georgia,
okaneray@ibsu.edu.ge

Abstract
Artificial neural network models have been already used on many different fields successfully. However, many researches show that ANN models provide better optimum results than other competitive models in most of the researches. But does it provide optimum solutions in case ANN is proposed as hybrid model? The answer of this question is given in this research by using these models on modelling a forecast for GDP growth of Japan. Multiple regression models utilized as competitive models versus hybrid ANN (ANN + multiple regression models). Results have shown that hybrid model gives better responds than multiple regression models. However, variables, which were significantly affecting GDP growth, were determined and some of the variables, which were assumed to be affecting GDP growth of Japan,were eliminated statistically.

Keywords: *Artificial Neural Network, Hybrid Model, GDP Growth of Japan, Modelling Forecast, Variable Determination*

I. INTRODUCTION

For the past several decades ANN has been successfully used in a wide range of study of fields (Demir and Ozsoy, 2014; Demir, 2014). The recent upsurge in research activities into artificial neural networks (ANNs) has proven that neural networks have powerful pattern classification and prediction capabilities (G. Peter Zhang, 2004). Yet, neural networks are rather young in forecasting macroeconomic variables and becoming a more attractive tool alongside traditional econometric models.

Kuan and White (1994) have theoretically defined practicability of neural networks in forecasting econometric models. Swanson and White (1997) following Maasoumi, Khotanzad, and Abaye (1994) have applied artificial neural networks to model nine different macroeconomic variables. This paper compares the relative usefulness of different linear and nonlinear models using a wide array of out-of-sample forecasting performance indicators. Their major conclusion is that multivariate linear models are marginally better overall(Greg Tkacz and Sarah Hu,1999).

Kohzadi et al. (1995) compared ANN and ARIMA models to forecast the future of cereal and demonstrated that the forecast error of neural network model is between 18 to 40 per cent lower than that of ARIMA. DÜZGÜN R. (2008) forecasted GDP Growth of Turkey and compared the results of ANN and ARIMA models and found that ARIMA model performed better than neural network model. Moshiri and Cameron(2000) used Back Propagation Artificial Neural Network(BPN) models in comparison with econometric models to forecast inflation rate. The performance of hybrid BPN was as well as traditional econometric models and even better in some cases.

Krugman (2009) states that one of the several outbreaks during the 1980s was Japan’s bubbles which were mainly financed by bank loans. Kaihatsu(2014) supports Krugman(2009), arguing that “Great Stagnation” of Japan in 1990s was due to the dysfunction of the financial system stemming from the collapse of asset price bubbles in the early 1990s. Hayashi and Prescott(2002) believe that the reduction of workweek length in the 1988-1993 periods has an important impact on the economic stagnation which brought the revision of Labor Standard's Law. Cargill and Parker(2004) summarize the causes of Japan's downturn in a seven-step sequence which is attributed to tight monetary policy applied by Bank of Japan. The author holds Bank of Japan some responsibility for the stagnation.

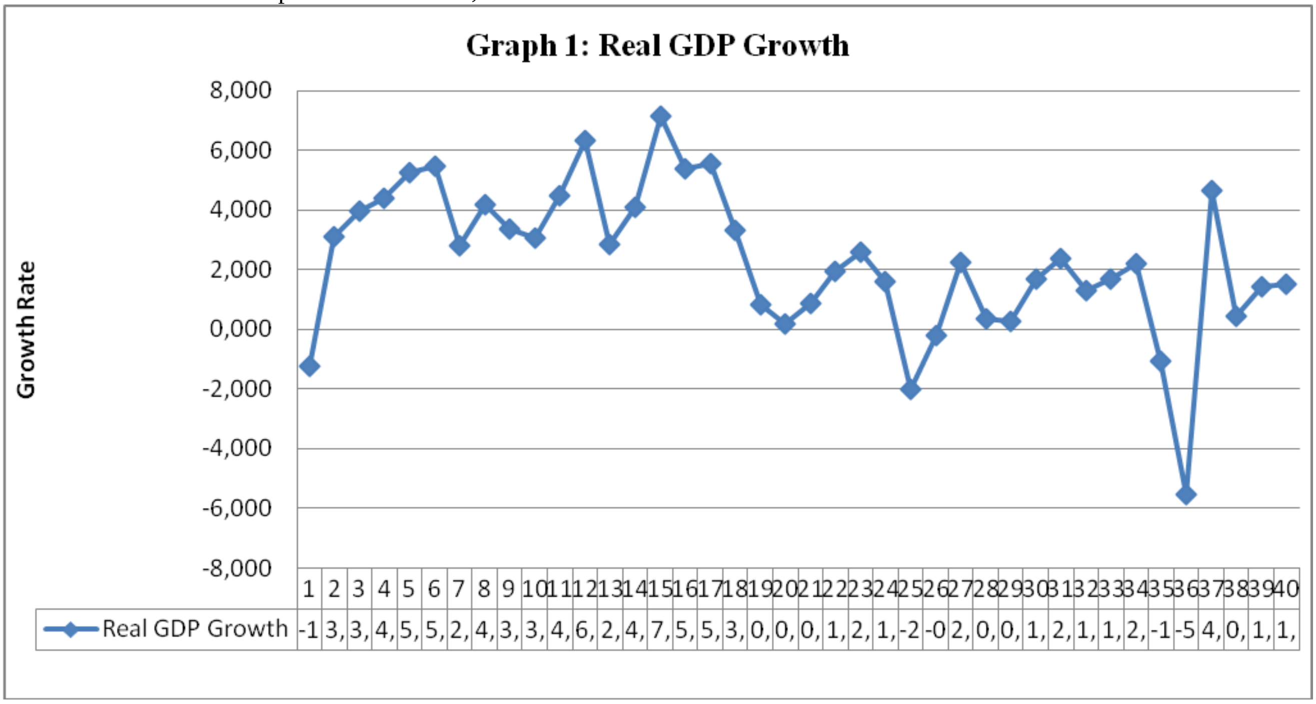
After the recovery of Japan's economy in 2003 it has experienced a severe economic contraction in the fourth quarter of 2008. Exports fell by 12.5 per cent in 2008 followed by 36.8 per cent fall in in the first quarter of 2009(Kawai and Takagi, 2009). In addition, one of the remarkable features of 2008 crisis is the collapse of international trade (Amiti, 2009).

In this research, it was observed that GDP growth of Japan was decreased dramatically (-5.527) in 2009. Based on the information above, it can be concluded that this decrease is because of global crisis that started in 2008. On the other hand, data indicates that GDP growth in 2010 was 4.652. These two impulses seem to be abnormal and makes forecast error increase. Furthermore information about data will be given the section 3.1.

II. Methodology

Data

Secondary data was taken from the original site of World Bank. Data includes the GDP growth of Japan from 1974 until 2013. There is plot of data below;



It can be observed from the graph that GDP growth data had a normal fluctuation until 2008 global economic crisis but after economic crisis, in 2009, GDP growth dramatically decreased. This decrease was an abnormal impulse and models had difficulty to catch this impulse. On the other hand, in 2010 Japanese GDP growth has jumped up and in these two consequent years a gap has occurred and this gap has increased the error level of forecast models. Out of these abnormal impulses, both multiple regression models and hybrid ANN models succeeded catching the points. Finally, models were compared despite of those abnormal fluctuations and results were discussed at the conclusion section.

III. MODELS USED FOR FORECASTING

This part introduces the models that were used in the research. Models are multiple regression models and artificial neural network models. Artificial neural network model is used hybrid with the support of multiple regression models.

Multiple regression models

Multiple regression models were widely utilized in economical researches (Ismail, Yahya, and Sabri, 2009; Kirshin, Maleev, and Pachkova, 2014; Pappu, Vijayakumar, and Ramamurthy, 2013). Multiple regression analysis helps a researcher to determine the effects of each independent variable on dependent variable (Xu, 2003). If there is an output like production, GDP growth…etc., there must be some variables which affect the level or quality of that output. In our case the GDP growth of Japan is taken as output. In this case the factors that affect the GDP growth will be tried to be determined. Effect of each factor or variable might be shown as;

$Y = \beta0+\beta1w1+\beta2w2+\beta3w3+\beta4w4+\beta5w5+ \beta5w5+ \beta6w6…. \beta nwn$ (1)

Here the meanings of the variables are;

Y: is the dependent variable that is searched for in the next period.

β0: is the constant of regression line.

β (1-n): is the coefficient of the concerning parameter on dependent variable. It doesn't have to be the same for each independent variable. It depends on how much this independent variable effect the dependent variable.

W (1-n): is the unit of each independent variable (Interest Rate, Inflation, M1…etc.) that effect the dependent variable (GDP Growth).

In this research the dependent variable is taken as GDP Growth and independent variables are initially determined as 12 variables such as year, long term interest rate, short term interest rate, interest rate yield spread, consumer price index, inflation expectation, real short term rate, M1, real M1, M2, real M2, and real GDP. Initially, coefficients of all variables those researchers think has an impact on GDP Growth were analyzed by IBM SPSS 20. The performance of the ENTER module of the software is shown as;

Table 1: Model Summary of Multiple Regression Anlysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.836	0.700	0.610	1.551442	2.560
Predictors: (Constant), Real_GDP, Interest_Rate_Yield_Spread, Real_Short_Term_Rate, Real_M1, Long_Term_Interest_Rate, Real_M2, Consumer_Price_Index, YEAR_, M2					
Dependent Variable: GDPGROWTH					

The table above shows that totally dependent variables explain 83.6% of the total factors. This is a normal level for the acceptance. Durbin-Watson test result is 2.56. It shows that there is no autocorrelation within the factors. Normally the acceptable level of this test should be between 1.5-2.5 but this result is a bit above than the upper control limit. F test of the regression model is shown below;

Table 2: ANOVA

Model		Sum of Squares	DF	Mean Square	F	Sig.
1	Regression	168.203	9	18.689	7.765	0.000
	Residual	72.209	30	2.407		
		240.412	39			
	Total					

Results of ANOVA test show that the model is significant because the P<=0.05 (0.000). but still we should consider t test results of the analysis because F test show overall model significance but each variables' significance also should be determined. If there is non-significant variable, it should be excluded from the equation. The t test results of the model are;

Table 3: Coefficients of the Model

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1188,105	553,628		2,146	,040
	YEAR_	-,619	,285	-2,916	-2,177	,038
	Long_Term_Interest_Rate	,143	,329	,166	,435	,667
	Interest_Rate_Yield_Spread	,712	,412	,425	1,728	,094
	Consumer_Price_Index	,188	,189	1,083	,992	,329
	Real_Short_Term_Rate	,059	,193	,105	,306	,762
	Real_M1	1,721E-12	,000	1,172	1,416	,167
	M2	-9,794E-14	,000	-13,479	-3,793	,001
	Real_M2	1,073E-11	,000	12,247	3,780	,001
	Real_GDP	3,660E-14	,000	1,523	1,847	,075

The table above shows the coefficients of the regression equation. T test results show the significance level of each variable. It is blatant that some of the factors that have somehow effect on GDP growth are not significant. This means that those variables might be taken out and only significant variables might be entered to the equation. To do this, the same regression model is performed by **“Backward”** module of IBM SPSS 20. The results are as;

Table 4: Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,836[a]	,700	,610	1,551442	
2	,836[b]	,699	,621	1,528588	
3	,835[c]	,697	,630	1,509471	
4	,822[d]	,676	,617	1,536640	2,407

a. Predictors: (Constant), Real_GDP, Interest_Rate_Yield_Spread, Real_Short_Term_Rate, Real_M1, Long_Term_Interest_Rate, Real_M2, Consumer_Price_Index, YEAR_, M2

b. Predictors: (Constant), Real_GDP, Interest_Rate_Yield_Spread, Real_M1, Long_Term_Interest_Rate, Real_M2, Consumer_Price_Index, YEAR_, M2

c. Predictors: (Constant), Real_GDP, Interest_Rate_Yield_Spread, Real_M1, Real_M2, Consumer_Price_Index, YEAR_, M2

d. Predictors: (Constant), Real_GDP, Interest_Rate_Yield_Spread, Real_M2, Consumer_Price_Index, YEAR_, M2

e. Dependent Variable: GDPGROWTH

The results above show that the 4th model is the significant model which explains the 82.2% of the overall variance and the Durbin-Watson test result show that there is no autocorrelation among independent variables. F test results show that the model is overall significant (P=0.000).

Table 5: Coefficients of the Equation

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
4	(Constant)	463,017	175,705		2,635	,013
	YEAR_	-,250	,090	-1,175	-2,782	,009
	Interest_Rate_Yield_Spread	,652	,211	,390	3,091	,004
	Consumer_Price_Index	,146	,068	,842	2,134	,040
	M2	-1,023E-13	,000	-14,073	-5,779	,000
	Real_M2	1,067E-11	,000	12,169	5,208	,000
	Real_GDP	3,841E-14	,000	1,598	2,004	,053

T test results above show that the parameters that is included in the 4th model are significantly has impact on the GDP Growth. On the other hand, other factors (Short term interest rate, inflation expectation, M1, real short term rate, long term interest rate, real M1) are excluded from the equation. Only significant factors are included into the equation. finally, the model can be stated as ;

Y= 463.017+ (-0.250) Year+0.652I.R.Y.S. +0.146C.P.I. + (-1.023E-13) M2+1.067E-11RealM2+3.841E-14RealGDP (2)

Multilayer Perceptron

Artificial neural network models are flexible and nonparametric tools of calculation (Tang and Chi, 2005). These models can be used for forecasting, classifying, setting…etc. aims. Artificial neural network models are capable to find adequate solutions for many different kinds of problems.

Number of Input Layers: This layer consists of independent variables of the problem. 12 different independent variables were filtered by using multiple regression analysis. Results of the analysis have shown that only 6 variables out of 12 had significant effect on the result. That is why 6 variables were used as independent variable for the input layer. The results of multiple regression analysis also were considered as another independent variable. By this way, our model became a hybrid model. Totally 7 variables were used as input layer entries.

Number of Hidden Layers: This layer is called as black box of neural network models. In the literature it has been found that one or two hidden layer number is the most common ones. Otherwise it has been seen that the models were memorizing instead of learning. However, it takes more time to calculate in case more hidden layers were chosen. For these reasons we have chosen only one hidden layer.

Number of Neurons: There is no specific calculation about finding the best number of neurons. Using trial and error method, we have specified four neurons finally.

Output Layer: This layer is the objective layer of the problem. In our case, we have one output layer and called as GDP Growth.

Totally two scenarios were used to find the best neural network model. As first scenario, 7 independent variables (Year, Interest Rate Yield Spread, Consumer Price Index, M2, Real M2, Real GDP, and result of Backward Multiple Regression Model) were used. Table 6 shows the information about the model and;

Table 6: Network Information

Input Layer	Covariates	1	YEAR_
		2	Interest_Rate_Yield_Spread
		3	Consumer_Price_Index
		4	M2
		5	Real_M2
		6	Real_GDP
		7	BACKWARD_REGRESSION
	Number of Units[a]		7
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		4
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	GDPGROWTH
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity

a. Excluding the bias unit

Here there are the calculations of each hidden layer and finally the neural network model is stated below;

H (1.1) = 1.354+ (-0.041) Year+ (0.632) I.R.Y.S. + (0.805) C.P.I. + 0.630 M2+ 0.035 RealM2+ 0.418 Real GDP+ (-0.507) BACKWARD R.A. *(3)*

H (1.2) = *1.989+ (-1.322) Year+ (0.175) I.R.Y.S. + (0.347) C.P.I. + (-0.244) M2+ 0.368 RealM2+ 1.094 Real GDP+ (1.331) BACKWARD R.A.* *(4)*

H (1.3) = 0.853+ (-0.998) Year+ (0.935) I.R.Y.S. + (0.046) C.P.I. + 0.156 M2+ 0.095 RealM2+ 0.364 Real GDP+ (-0.749) BACKWARD R.A. *(5)*

H (1.4) = *-0.088+ (0.430) Year+ (0.343) I.R.Y.S. + (-0.122) C.P.I. + 0.124 M2+ 0.240 RealM2+ (-0.086) Real GDP+ (0.097) BACKWARD R.A.*

Neural network model;

Y= I (-0.738 + (-0.752) h1+1,542h2+ (-0.416) h3+0.327h4 *(6)*

Here I (.) can be thought as an activation function in order to convert the rescaled data into the original series.
As a result, mean absolute error (MAE) was calculated as 0.732. On the other hand, mean squared error level was 0.959. The comparison will be done at the conclusion part.

IV. FINDINGS AND CONCLUSION

As a result, it was observed that the filtered multiple regression analysis model performs better than the simple multiple regression model. On the other hand, hybrid artificial neural network model performs better than both filtered multiple regression model or simple multiple regression model. Here below there is a graph that shows the performances and table show about the error levels of each model;

Graph 2: Performance graph of the models

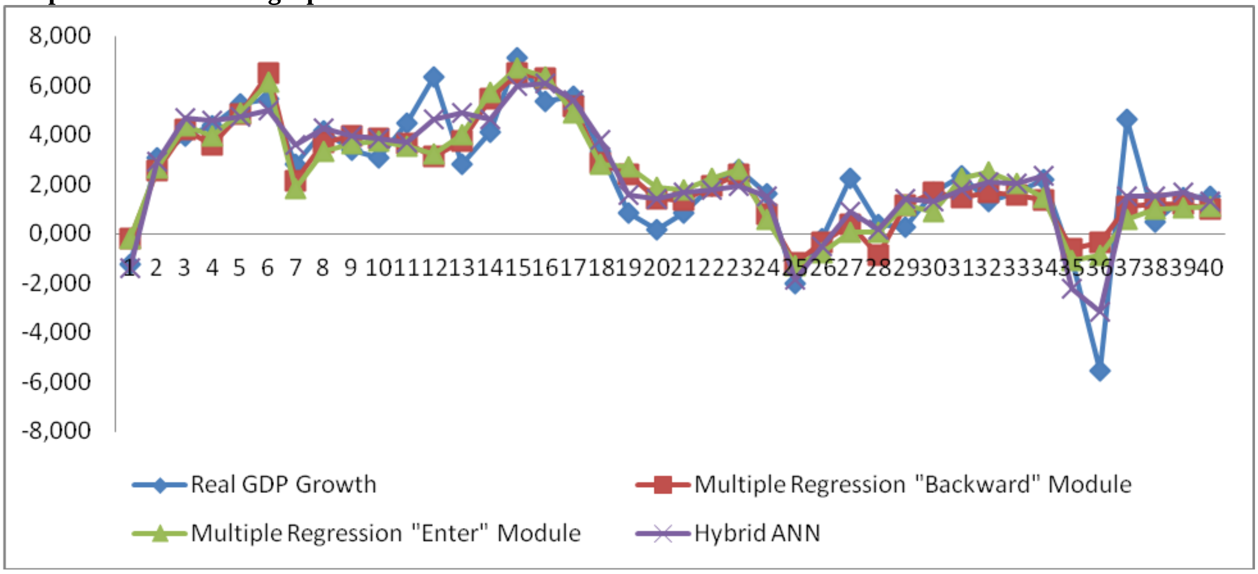


Table 7: Error levels of the models

Model Name	MAE	MSE
Simple Multiple Regression Model	0.9789	1.919972
Filtered Multiple Regression Model	0.9344	1.833406
Hybrid ANN Model	0.7324	0.959231

It is blatant from the error levels that the hybrid ANN model gives better results than other models. From this point, GDP Growth of Japan can be calculated using “Hybrid ANN” rather than multiple regression models. This doesn’t mean that Hybrid ANN model performs the best every time. In order to say this, all forecasting models must be compared with this model. But in this case it was proven that the Hybrid model performed better. Furthermore, despite of 12 independent variables used for the modelling, it was observed that not all of them significantly affected the model. That is why; non-significant independent variables have been subtracted from the model.

As a result, it was seen that Year, Interest Rate Yield Spread, Consumer Price Index, M2, Real M2, and Real GDP variables are significantly affected the GDP Growth of Japan. On the other hand, Short Term Interest Rate, Inflation Expectation, M1, Real Short Term Rate, Long Term Interest Rate and Real M1 variables didn’t affect the GDP Growth significantly and subtracted from the model. By this way we have determined some of the factors that affect GDP Growth of Japan and proven that the Hybrid ANN model performed better than multiple regression models.

V. REFERENCES

Ahmet Demir, S. O. (2014). Forecasting the Monthly Electricity Demand of Georgia using Competitive Models and Advices for the Strategic Planning.*International Journal of Academic Research in Economics and Management Sciences, 3*(5).

David F. Hendrya, M. P. C. (2003). "Economic forecasting: some lessons from recent research." Economic Modeling_20(1): 301–329.

Demir, A. (2014). The method of selection of parameters' values in the problem of determining the quality level in manufacturing, business and education.*International Journal of Social Sciences and Educational Studies, 1*(2).

DÜZGÜN, Y. D. D. R. (2008). "A comparison of artificial neural networks' and arima models' success in gdp forecast." Marmara Üniversitesi İ.İ.B.F. Dergisi 17(2).

Fumio Hayashi, E. C. P. (2002). "The 1990s in Japan: A Lost Decade." Review of Economic Dynamics 5: 206_235.

Greg Tkacz, S. H. (1999). Forecasting GDP Growth Using Artificial Neural Networks, Bank of Canada**:** 1-24.

Ismail, Z., A. Yahya and A. Shabri, 2009.Forecasting Gold Prices Using Multiple Linear Regression Method. Am. J. Applied Sci., 6: 1509-1514.

I.A. Kirshin, M.V. Maleev, O.V. Pachkova, Assessment of Impact of Domestic and External Demand Factors on Economic Growth in Russia on the Basis of Model of Multiple Regression Analysis, Procedia Economics and Finance, Volume 14, 2014, Pages 320-325

J. Sharon Mano Pappu, G. K. V., V. Ramamurthy (2013).Artificial neural network model for predicting production of Spirulinaplatensis in outdoor culture.*Bioresource Technology*.

Krugman, P. (2009). The return of depression economics and the crisis of 2008. USA, W. W. Norton & Company, Inc.

M. Amiti, D. E. W. (2009). EXPORTS AND FINANCIAL SHOCKS. Cambridge, England, NATIONAL BUREAU OF ECONOMIC RESEARCH.

S. Kaihatsu, T. K. (2014). "What caused Japan's Great Stagnation in the 1990s? Evidence from an estimated DSGE model."Journal of The Japanese and International Economies (34): 217-235.

Swanson, N. R. a. H. W. (1997)."A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks."REVIEW OF ECONOMICS AND STATISTICS 79(4): 540-550.

T.F Cargill, E. P. (2004). "Price deflation and consumption: central bank policy and Japan's economic and financial stagnation." Journal of Asian Economics, 15: 493–506.

Takagi, M. K. a. S. (2009). Why was Japan Hit So Hard by the Global Financial Crisis? Tokyo, Asian Development Bank Institute. 153.

Tang, T. C., & Chi, L. C. (2005). Neural Networks Analysis in Business Failure Prediction of Chinese Importers: A Between-Countries Approach, Expert Systems with Applications, 29, 244-255.

Xu Y., (2003). Using Data Mining in Education Research: A Comparison of Bayesian Network with Multiple Regression in Prediction, Department of Educational Psychology, The University of Arizona, Arizona.