Journal of
Economics and Political Economy
www.kspjournals.org
Volume 3 September 2016 Issue 3

Projections of Inflation Dynamics for Pakistan: GMDH Approach

By Shahid IQBAL [a†] & Maqbool H. SIAL [b]

Abstract. This study is focused on identifying, based on various forecast accuracy criteria, best inflation forecasting model for Pakistan using the in sample projections for Pakistan inflation from 2006II to 2009II. To resolve the important issue of degree of contribution in forecasting performance of the two monetary aggregates in forecasting inflation, three main predictors: real GDP, interest rate and one out of the two monetary aggregate have been used, thus constructing two models; one with Divisia Monetary Index (DMI) and other with Simple sum monetary aggregate (SSMA). It is revealed that, though both of the monetary aggregates are important predictors in forecasting inflation, but DMAs provide better fit and improved forecasts as compared to their simple sum counterpart. Hence, the evidence is established that monetary aggregates still play a dominant role in predicting inflation for Pakistan economy. The study recommends the construction, publication, and use of high frequency DMAs by the State Bank of Pakistan (SBP) for forecasting inflation in Pakistan instead of SSMAs. Finally, to identify the improvement in forecast accuracy w.r.t. different forecasts combination, these forecasts have been combined and compared. It is revealed that when the structure of an empirically observed underlying series has complex nonlinear structure then forecasts based on single nonlinear model may fail to capture these diverse complexities. The best strategy is then to use various nonlinear models and combine these forecasts. Further the study concluded that if the complex nonlinear structure of an observed series is, a priory, unknown then universal approximators like Group Method of Data Handling (GMDH)- Polynomial Neural Networks (PNNs) and GMDH-Combinatorially Optimized (CO) could provide outstandingly accurate forecasts yet avoiding 'overfitting' even for small sample size. Specifically, it recommends the use of nonlinear non-parametric universal approximators for forecasting inflation in Pakistan by the SBP.
Keywords. Monetary aggregate, Nonparametric nonlinear models, Universal approximators, Forecasting performance, Forecasts combination.
JEL. E31, E47, E51, E52.

1. Introduction

Accomplishment of price stability, in the sense of a low and steady inflation, is a key to economic growth of the economy and it is one of the objectives of almost every central bank throughout the world. Monetary authorities constantly need to monitor and forecast the prices evolution; hence, central banks

[a†] Department of Statistics, Higher Education Department Govt. Punjab, Pakistan.
☎. +923228600650
✉. shahidiqbal.uos@gmail.com
[b] Department of Economics, University of Sargodha, Pakistan.
☎. +923247729079
✉. maqsial@yahoo.com

necessitate a good model to forecast inflation. Therefore, the worldwide central banks have long conscious practice in forecasting inflation. Trichet (ECB, 2003) argued that inflation forecasts are "useful, even indispensable, ingredients of monetary policy strategy." For most central banks, inflation is at least one monetary policy objective. Some central banks have even resorted to inflation forecast targets presumably based on very reliable inflation forecasts. Yet, even in circumstances where structural relationships are not up to the mark with regard to stability and data quality is in the way of progress, inflation forecasts can provide valuable information on future economic activity scenarios of the economy, which may further need to be combined with supplementary analysis beyond econometrics.

Simple sum measure of money referred as Simple sum monetary aggregate (SSMA) is conventionally build upon simply summing all the component assets of the money stock with unit weighting and these continue to be the official framework used by any central banks. These SSMAs are generally being used to guide monetary policy decisions, although this method has demonstrably been identified with stringent faults. Divisia Monetary Aggregate (DMA) is envisaged as improved measure of the combined monetary service flow and these have demonstrably worked better than SSMAs shown in many studies across the world. Nevertheless, the performance of DMAs in providing improved out-of-sample inflation projections is yet to have overwhelmingly valid empirical evidences across the globe and it has developed a new area of research. Many recent studies; e.g. Schunk (2001) and Drake & Mills (2005), have raised the issue of both the aggregates for the US. But SW (1999) take up the wider issue: whether monetary aggregates could be used to ameliorate inflation forecasts models. They found that SSMAs give marginal ameliorations in some measures of price levels over some sample time spans, but the accuracy of CPI forecasts significantly dropped from 1970 to early 1980s for the US. Attempts to forecast inflation using monetary aggregates is done by Dorsey (2000) Elger et al. (2006), Binner et al. (2010) Azevedo & Pereira (2010a) Berger & Österholm (2011) Kovanen (2011) among others in the recent past and for Pakistan Bokil & Schimmelpfennig (2006), Haider & Hanif (2009), among a few in the recent past.

To evaluate the in sample forecasting performance, the price inflation ismodelled and projected with a two non-parametric models namely: Group Method of Data Handling GMDH-PNNs and GMDH-CO involving Kolmogorov-Gabor (K-P) polynomial. For the remainder of the study, the literature is reviewed on forecasting in the next section, methodology for computation forecasting is detailed in section 3.Section 4 is focused on resultswith computations and comparisons of forecasting exercise and forecast combinations. The main findings and recommendations of the study are outlined in section5 and finally, references are added in the last section.

2. Literature

More explicitly, the inflation forecasting strength of standard Divisia and simple sum indices, with that of two newDivisia indices—adjusted for high financial innovation period to capture the true user costs of the component assets—was computed and compared by Binner et al. (2004) for Taiwan economy for time span of 1970Q_1-95Q_3. Three dummies relating to three spans of high inflation were used. The simple Neural Network (NN) model was extended to include further explanatory variables regarded as having forecasting potential such as GDP and interest rate. The dual DPI and 3-month deposits were interchanged with change of Divisia variant or simple sum money, respectively. The Divisia index was adjusted to take into account the financial liberalization of Taiwan since the 1970s. The preferred inflation forecasting NNs model-employed a DMA M_2 adjusted to fit in a learning mechanism to permit agents to change slowly their views of the raised productivity of Money-outperformed the conventional econometric system approach. The explanatory strength of the two innovation-adjusted DMIs dominated the SSMAs in the majority of cases. Drake & Mills (2005) built on SW (1999) in order to forecast growth of nominal GDP and price level for a time span of 1991–2001 utilizing SSAMs M_2 and M_2+ (contained stock and bond mutual funds plus M), DMA M_2, and a new empirically weighted (derived weights from long-run relation between monetary components and nominal GDP in level) monetary aggregatebased on observations from 1960:Q_2 to 2001:Q_2. They found that SSMA M_2 furnished better forecasts of nominal GDP growth, but the newly weighted monetary aggregate outperformed in forecasting inflation mainly at longer horizon. Unlike Schunk (2001), they found that SSMA M_2 outperformed DMA M_2 every time and it was suspected the potential reason for DMA M_2 low performance was the benchmark rate chosen by the Fed. of St. Louis. Hale & Jorda (2007) used the core CPI inflation, M_2, M_3 monetary aggregates, M3C (a rectified M_3), real GDP in the euro area, 4-month euro area Euribor, industrial production and US federal funds rate for the sample period of 1985M_1–2007M_1. They concluded that monetary aggregates had no predictive power with regard to forecasting inflation rather inflation forecasting potential of monetary aggregates seemed to be embedded in measures of past price level, interest rates, and economic activities for the US. For the euro area, over short (but not the long ones) horizons, inclusion of monetary aggregates in the inflation-forecasting model appeared useful; although it seemed probably small time span for the monetary aggregates to encompass a sizable impact.

Binner et al. (2010) compared the contribution of both the aggregates for forecasting inflation from narrow to broad levels of aggregation and explored many types of interest rate including different DMAs and SSMAswith different collections of included monetary assets building upon many recent studies such as Schunk (2001), Drake & Mills (2005) and Elger et al. (2006). Of the 541 (1960M_1 - 2005M_2) observations available, the first 433 observations trained the networks, the next 50 observations validated and the last 46 (2001M_5–2005M_2) were left for forecasting US inflation. They, utilizing Recurrent NNs (RNNs) and Kernel Recursive Least Squares Regression (KRLSR) models, found that RNNs operate with latent unlimited input memory, while the KRLSR was a limited memory predictor. The inflation forecasts of two competing models were then compared to random walk model forecasts and it was revealed that KRLSR was the best among the three compared, but evidence for the worth of monetary aggregates in forecasting inflation could not be established.

The PNNs of GMDH typeemerged as a variant of NN, which is used in Generalized Regression NN (GRNN). Its key benefit lies in its ability to swiftly learn and quickly converge to the best regression surface essentially with large number of data sets making GRNN method to be the best model for the prediction in comparison with its close competitors. Generalization in GRNN is typically achieved by dividing available training data into three sets; one to be used for network training, the other to be used to verify training performance of algorithms as they are run, and the last one for running final independent test. Owing to its vigorous capacity for nonlinear mapping and better robustness, GRNN could attain the maximum sensitivity as it employs RBF set up, which consequently make GRNN as useful exploratory and predictive tool for the appraisal of rice biophysical parameters (Yang et al., 2009). Ahangar et al. (2010) gauged the active firms' stock price in Tehran stock exchange, Iran. Using both linear regression and

GRNN techniques for an architecture incorporating ten macroeconomic and thirty financial variables at the beginning, they were left with only three macro-economic variables and four financial significant variables at final stages. In order to determine the stock price, using independent components analysis by describing the equations of the two methods for the comparison, they demonstrated that artificial GRNNs method was more efficient than other method. Urwatul-Wutsqa et al. (2006) extended NN application to multivariate data, particularly in time series analysis and proposed VAR-NN by mixing the NNs and VARs that belonged to the genre of nonparametric and nonlinear model. Leung et al. (2000) applied GRNN in exchange forecasting to demonstrate that the GRNN outperformed the other forecasting methods.

Yao & Ni (2009) used Autoregressive (AR)-GMDH and analog complexing algorithms to forecast oil prices. The evidence on feasibility and validity of AR-GMDH was demonstrated by the comparing the performance with traditional techniques. It was and confirmed that AR-GMDH was more accurate in forecasting such complex systems. Zheng et al. (2010) found double trends; long-term upward trend and seasonal fluctuations trend, in monthly cigarette sales. It seemed impossible to model such complexity with a few linear and nonlinear models. It was envisaged that a more flexible forecasting model could deeply capture the characteristics of a complex system forecasting. So, they proposed a combination of ARIMA and GMDH models based on info-entropy method to get merits of both ARIMA and GMDH models in linear and nonlinear modeling. They empirically demonstrated that the proposed combined model was effective in improving forecast accuracy when compared with either of the individual models. Samsudin et al. (2011) developed a hybrid, GMDH embedded Least Squares Support Vector Machine (LSSVM) called GLSSVM forecasting model while GMDH to work out useful predictors in the time series forecasting for the LSSVM. Based on $1962M_1$ to $2008M_{12}$ for Selangor river and for Bernam river $1966M_1$ to $2008M_{12}$ river flow data, with monthly time-points below $2004M_{12}$ for training and from $2005M_1$ to $2008M_{12}$ for testing, the forecasting performance of this model was compared with the conventional NN models, ARIMA, GMDH and LSSVM models using RMSE and coefficient of correlation (R). The GLSSVM outperformed decisively the other models. Chaudhuri (2012) has recently modelled the annualized earning per share (a broad measure of a firm's entire marketable yield), capital, and turnover of an Indian Auto Major from 1953-54 to 2008-09 data. He used by computer-aided self-organization techniques, multilayer GMDH NNs and combinatorial algorithms. Beyond the theory of parametric econometrics, he used a "black box" method that requires no a priori knowledge or assumption of the inner mechanism. Thus, it was free from economic theory to determine the structure. The study demonstrated that the GMDH approach was straightforward, simple, and tremendously useful for trade off studies that lead to alternate economic (monetary and fiscal) policies. The GMDH models combinations of quadratic polynomials and regression analysis ensemble the underlying structure, thus it is an improvement for the analysis of economic phenomena such as earning per share (as economic indicators of economic growth) using non-stationary data. Once the models have been identified by GMDH methods the time-varying parameters can be estimated with help of fresh observations in GMDH algorithms.

Varahrami (2012) used Multi-Layer Feed Forward (MLFF) NNs with back-propagation learning algorithm and GMDH NNs with genetic algorithms of learning to predict the gas price of Henry Hob database for the period from 1st January 2004 to 13th July 2009 by employing moving average crossover inputs. The results confirmed a short-term dependence in gas price fluctuations. The GMDH NNs outperformed MLFF NNs in prediction accuracy.

The innovation of forecast combinations was initiated by Bates & Granger (1969) and Ried (1968) but many of its improvement and several review articles have appeared after two decades namely: Clemen (1989), Diebold & Lopez (1996) and Timmermann (2006) among others. It gained more popularity in the first decade of this century and was further introduced in the NNs by Hashem et al. (1993) and Hashem (1997). Acknowledgement with regard superiority of forecast combinations to their constituent forecasts is documented in the literature persistently (Timmermann, 2006).

A noteworthy result of Stock, & Watson (1999) was to combine the nonlinearly generated forecasts among the most accurate in rankings. For forecasting one month ahead, these were among the top five in 53% (specifications in levels) and 51% (specifications in differences) of all produced. For forecasting six and twelve months ahead, these percentages dropped between 30% and 34%. Comparable performance was found in the combinations comprising of all linear models at these horizons. No single model performed dominantly. Rather than trusting only on one nonlinear specification, it was found that the use of a larger number of nonlinear models with forecasts combination from these models increases accuracy. It appeared that combining forecasts might lead to better forecasting accuracy than what is attained by linear counterparts. Due to presence of some exploitable nonlinearity in macroeconomic data, it was too diffuse to be captured by only one nonlinear specification. A similar result was obtained from the study of Jaditz et al. (1998), which yielded superior forecasts by combining the nonparametrically generated individual forecasts. Marcellino (2002) compared 58 models for forecasting 480 series for the twelve countries of the European Monetary Union, but unlike SW in Marcellino (2004), the forecast combinations were not considered. Besides purely linear models in the study, he further utilized linear models with stochastic coefficients, each following a random walk, NNs, and LSTAR models. Terasvirta et al. (2005) combined forecasts and established evidence that in several cases, but without any system, these combinations consisting of pairs of forecasts, improved forecast accuracy in comparison with individual models. As only pairs were then combined only, so his assessment regarding usefulness of forecasts combinations cannot be regarded as overwhelmingly informative.

Interestingly, support for combining forecasts from nonlinear models recently emerged from the study of Kock (2009), who used 47 chronological data series from the G7 and Scandinavian economies. He used the K-G polynomial model -considered universal approximators -and found its forecasting performance nearly equal to that of NNs with logistic hidden units. The forecasts thus obtained had an accuracy edge over the ones from linear AR models. Aforementioned studies augment the better forecasting performance through forecast combination. Further, these studies augment the case of nonlinear models against linear counterparts in forecasting macroeconomic series. Recent applications on forecasts combination include Stock, & Watson (1999; 2003), Canova (2007), Ang et al. (2007), Inoue & Kilian (2008), and Clark & McCracken (2010) among others.

Abdullah & Khalim (2009) have investigated the key causal factors directly related to food inflation in Pakistan using JML technique to estimate long run results for the period from 1972 to 2008 namely: agriculture support price, food exports and imports, GDP per capita, and quantity of money. Bashir et al. (2011) examined demand and supply side determinants of inflation using VECM under JML in Pakistan and also to investigate causal relationships by GC test for period from 1972 to 2010. The long run CPI was found to be positively related to money supply, GDP, imports and government expenditures with long run elasticities of inflation w.r.t. the regressors were 0.61, 0.73, 0.41, and 0.32, respectively, whereas the government revenue inversely related with long run elasticities of inflation w. r. t. government revenue was found as -1.37. Lagged CPI and double lagged government revenue were directly inflating current year CPI. For a long run stable price level, recommendations included maintenance of optimal level for all of main determinants with increase in government expenditure and GDP.

For consistent out-of-sample YoY monthly inflation forecast for Pakistan for fiscal year (FY) 2008 based on period from 1993M7 to 2007M6, Haider & Hanif (2009) usedsimple univariate NN model. Model simulation used feed-forward with back-propagation architecture criterion. They found that forecast of inflation for the end of next FY2008 was higher as compared with FY2007. Further, NN based forecasts outperformed the AR(1) and autoregressive integrated moving average (ARIMA) models based forecasts utilizing RMSE criterion to check forecast accuracy.

At least, considering the global fitting and flexibility properties of the statistical model, the nonlinear NN would be likely to outperform the linear models provided the overfitting is suitably avoided. Since no such work has been carried out for forecasting inflation in Pakistan, except that of Haider & Hanif (2009), the universal approximators definitely are required play a role here. Hence, recent evidence recommends that simple, though nonlinear, models may be at least as competitive as linear ones for forecasting macro variables like inflation. Two such approximators, the K-G polynomial and the GMDH type NN are used to investigate the forecasting performance for Pakistan inflation.

We perform ex post forecasts on six months forecast horizon. These forecasts can be compared with actual figures to know and compare the forecasting performance of the competing models so that we are able to find and suggest, based on the results of this exercise, the most efficient and valid forecasting models and their combination for ex ante forecasts in Pakistan.

3. Methodology

Inflation is defined as a rising general overall level of prices in an area with passage of time. Four different price indices -CPI, GDP-IPD, SPI, and WPI- are used in Pakistan in fiscal year. Also, measures of core inflation, and Headline CPI inflation etc. are available. To measure general price level increase or decrease, mainly the CPI, GDP-IPD or both is used. CPI is based on Laspeyre's index and contains four types of biases: new-product; quality-change; substitution; and outlet; but the intensity of bias level assigned to each can be different. How big are these biases? In advanced countries, they are probably on the order of 1 to 2 percentage points, at most (BLS, 1997). However, in Emerging Market Economies (EMEs), where data collection is more difficult, these can be large. CPI biasis a rare investigated area of research in developing economies like Pakistan.

The GDP-IPD includes items barred from the CPI and contains extra items. Barred items include used consumer goods and imported goods prices—import prices growth rate is the chief inflation determinant in Pakistan, both in the short and long run (Chaudhary & Chaudhary, 2006). Therefore, GDP-IPD is a comprehensive measure of prices inflation, hence is used in this study.

3.1. Nonparametric Models

The specification and estimation of nonlinear models poses more difficulties for the researcher. First, the model should be fully specified using the appropriate order of the basis function. Second, estimation requires nonlinear optimization, which is even more difficult to handle even with the existence modern day computer.

The safest strategy to forecast with nonlinear underlying structure is to admit the doubt over the recognition of the specification and try to build a flexible approximation by allowing for very large set of specification to contain the underlying structure in the modelling space. This is situation where the universal approximators of the functions are to play a role. Evidence on NN forecasting in economics is growing rapidly with list of main applications including growth of GDP, stock returns, currency in circulation, demand for electricity, demand for construction and exchange rates. Many central banks are currently engaged in forecasting various macroeconomic indicators utilizing NNs (Haider & Hanif, 2009).

To cope with unknown nonlinearities in data is to use NNs, as these models are data driven nonparametric models capable of modelling/generating underlying nonlinear structures without prior information regarding inherent functional forms. Further, NNs, are highly flexible to approximate with any unknown nonlinear continuous function with high degree of accuracy such that a researcher fears overfitting rather than underfitting (underfitting is likely to be feared in linear or parametric nonlinear models). Finally, like parametric nonlinear specifications, the danger of making comparatively farthest forecasts rests with nonparametric specifications also, when forecasting from data points where the observations are relatively sparse (Kock & Teräsvirta, 2011). However, a few models that have been used in similar situations in the literature have been mentioned. Nevertheless, it cannot be guaranteed that these comprise the final layout. Multivariate linear and nonlinear models for inflation dynamics can be utilized to predict inflation namely: Kolmogorov-Gabor (K-G) Polynomials via GMDH framework with polynomial resembling via perceptron and quadratic combinations (semi-parametric), Some of the authors describe that some nonlinear models involving basis functions are universal approximators such as Fourier series, splines, and others perceptron based like NN (White, 2006; Chen, 2006).

Ivakhnenko (1968), motivatedby the flexiblestructure of polynomial, emanated this new algorithm, called GMDH, by pursuing a heuristic and perceptron type framework. He tried to ensemble the K-G polynomial by utilizing sets of lower degreepolynomials for each duo input variates. He demonstrated that a 2nd degree polynomial - called Ivakhnenko polynomial given in equation (5) below-could rebuild the entire K-G polynomial via a repetitive perceptron style method. Introduced in the late sixties, gradually it became an alternative semiparametric method to nonlinear parametric modelling. As a contender to stochastic approximation method, pioneered by Ivakhnenko (1968), it gained rapid popularity all over the world due to its prediction accuracy in forecasting with ability to flexibly approximate underlying functional form of any degree, hence becoming a globally universal approximators. This method is a type of heuristic self-organizing black box framework entailing the concepts of connectionism of cognition theory and complete mathematical induction (Muller & Ivakhnenko, 1996). This framework provides improved accuracy owing to its perceptron style build up that permits the dichotomization of the observations into “beneficial” and “damaging.” It requires smaller data, hence reducing calculation time. Its initial progress was based on frequent computational trials and in similarity to the vindication of the Monte Carlo method, multiple reiterations of an investigational outcome make up its proof (Ivakhnenko, 1988).

3.2. GMDH

Owing to the deficient in apt mathematical basis apart from statistical postulations, the theory of GMDH grew as a branch of regression analysis initially (Stepashko & Yurachkovskiy, 1986). This theoretical deficiency has been criticized, and some of theorists have tried to justify some of the features like the convergence of multilayer algorithm. Further, a sequence of theorems was propoundedrelating to a broader-scope two levels forecasting with GMDH. These theorems could provide a base for a further generalized theory in relevant projections exercises (Ivakhnenko & Kocherga, 1983). Even though atheoretical several GMDH algorithms, with capability to approximate poorly-defined stuff with reasonable accuracy, have established its strong position as an apt nonlinear technique for pattern recognition, modelling, prediction and forecasting assignments. The broad scope of partial metaphors permits its utilization in various fields of modelling with edge over many other known statistical methods as well as generating several respective GMDH algorithms (Ivakhnenko & Kocherga, 1983). Many of its earlier utilizations were related to time series predictions.

K-G polynomial approximations to unknown nonlinear functional forms are uncommon in economic forecasting (Kock & Terasvirta, 2011). New spur in automated model selection have generated interest in such methods (see Krolzig & Hendry, 2001). Castle & Hendry (2006; 2010) utilized K-G polynomialsas a start for nonlinear model selection with the idea to approximate well-known nonlinear models such as the translog; a special case of K-G polynomials. Castle and Hendry have discussed linearity tests based on K-G polynomials[1] as these nest the linear model. Interestingly, the well-known translog production function is based on a K-G polynomial of order two.

3.2.1. GMDH-PNNs Structure

GMDH algorithm splits a model it into set of base functions called neurons and in each layer, diverse pairs of neurons are linked through a 2nd degree polynomial that generates new neurons in the subsequent layer. Such type of structure is applied in the model to map inputs to outputs. The recognition task is to find a function $\hat{f}$ that is very close approximation of actual, fso as to predict output,$\hat{y}$, for a given vector of inputs $X = (x_1, x_2, , \ldots, x_n)$ maximum possibly closest to its actual outputy. Hence, for given T observed data pairs of many inputs and single output, the tangible function is:

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}) (i = 1, 2, \ldots, T), \quad (1)$$

A GMDH-PNN can possibly be trained in order to predict the outputs$\hat{y}$, for the given inputs$X = (x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in})$, such that:

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in})(i=1,2,\ldots,T). \quad (2)$$

To resolve a GMDH-PNN such that the square of deviations between the observed and the predicted output becomes minimum that is:

$$\sum_{i=1}^{T} [\hat{f}(x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}) - \hat{y}_i]^2 \rightarrow \min. \quad (3)$$

For a multivariate specification of unknown model of n regressors, a general multivariate relationship between n inputs and an output variates can be approximated by an intricate discrete type of the Volterra series as under:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ijk} x_i x_j x_k + . \quad (4)$$

In GMDH for data mining, modelling, optimization, pattern recognition, and forecasting, the gradually complicated K-G polynomial given in (4) is used as the most popular base function. Like so, the given partial quadratic form can be recursively used in a NN of linked neurons to assemble the globalized mathematical relationship between inputs and outputs variates set in equation (5). If the number of lags, and hence the number of sums is finite, it is known as K-G polynomial. Further, complete structure of this mathematical depiction can be described by a system of only two variables (neurons) partial quadratic polynomials of the form:

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 {x_i}^2 + a_5 {x_j}^2 \quad (5)$$

The coefficients a_i's of (5) are estimated with regression method in a least-squares sense. The difficulty now lies in setting up a GMDH-PNN such that the square of deviations of the observed and the predicted output is minimized and, in turn, for every inputs pair of x_i, x_j is minimized. So that the deviation between observed output, y and predicted $\hat{y}$, is minimum i.e. (y-$\hat{y}$) is minimized. To fulfil this task for all such pairs of neurons, a hierarchy of polynomials is built using the quadratic form given in (5) to get the estimates of coefficients a_i's of every 2nd degree function $G(x_i, x_j)$. To fit the output optimally in the whole set of inputs-output observed vector we use:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T}(y_i - \hat{y})^2}{T}} \;\rightarrow\; min \quad (6)$$

Consequently, $\binom{n}{2} = \frac{n(n-1)}{2}$ neurons are constructed in the 1st hidden layer of the feed-forward NN from observed data $\{(y_i, x_{ip}, x_{iq})$; (i=1, 2,…, T)$\}$ $\forall p, q \in \{1,2,\ldots,n\}$ such that $p \neq q$. Now, it is feasible to build T data triples $\{(y_i, x_{ip}, x_{iq})$; (i=1, 2,…, T)$\}$ from observed values using all such $p, q \in \{1,2,\ldots,n\}$ in matrix form as: $\begin{bmatrix} x_{1p} & x_{1q} & \cdots & y_1 \\ x_{2p} & x_{2q} & \cdots & y_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{Tp} & x_{Tq} & \cdots & y_T \end{bmatrix}$.

A 2nd degree sub-expression as in eq. (5) is used for every row of T data triples to readily obtain the matrix equation as under:

$$A\boldsymbol{a} = Y^T \quad (7)$$

Where '$\boldsymbol{a}$' is the vector of coefficients to be estimated and is unknowns of the 2nd degree polynomial in eq. (5).

$$\boldsymbol{a} = \{a_0, a_1, a_2, a_3, a_4, a_5\} \quad (8)$$

And

$$Y = \{y_1, y_2, y_3, \ldots, y_T\}^T \quad (9)$$

is the vector of observed output's and A is the vector of observed output's takes the form:

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{Tp} & x_{Tq} & x_{Tp}x_{Tq} & x_{Tp}^2 & x_{Tq}^2 \end{bmatrix}. \quad (10)$$

By OLS method in matrix form the normal equations can easily be solved as under:

$$a = (A^T A)^{-1} A^T Y \qquad (11)$$

Where $A^T A$=

$$\begin{bmatrix} n & \sum x_{1p} & \sum x_{1q} & \sum x_{1p}x_{1q} & \sum x_{1p}^2 & \sum x_{1q}^2 \\ \sum x_{1p} & \sum x_{1p}x_{2p} & \sum x_{1p}x_{2q} & \sum x_{1p}x_{2p}x_{2q} & \sum x_{1p}x_{2p}^2 & \sum x_{1p}x_{2q}^2 \\ \sum x_{1q} & \sum x_{1q}x_{3p} & \sum x_{1q}x_{3q} & \sum x_{1q}x_{3p}x_{3q} & \sum x_{1q}x_{3p}^2 & \sum x_{1q}x_{3q}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum x_{1q}^2 & \sum x_{1q}^2 x_{Tp} & \sum x_{1q}^2 x_{Tq} & \sum x x_{1q}^2{}_{Tp} x_{Mq} & \sum x_{1q}^2 x_{Tp}^2 & \sum x_{1q}^2 x_{Tq}^2 \end{bmatrix}$$

Hence (11) provides solution for the vector of the best coefficients in (5) for the entire set of T observations triples. This process is replicated for every neuron of the later hidden layer depends on the connectivity topology of the NN, until final form is estimated.

Four quarter simple moving average $\bar{y}_t$is the most widely used procedural indicator which smoothes values of annual observed quantities utilizing average over time. For quantities observed on yearly basis, window of the time period n= 4 generally smoothes the seasonal variations. The shorter the time period, the more reactionary a moving average becomes. To account for seasonal variations, this study utilized n = 2 only due to fact that biannual times series are involved here in order to mimic the intrinsic inflation growth path and eliminate erratic short-term fluctuations, which perhaps have no link with long-run inflation growth. The average over the time period n against a time point 't' is calculated by:

$$\bar{y}_t(n) = \frac{1}{n} \sum_{i=0}^{n} B_{t-i}, \qquad (12)$$

where B_t corresponds to observed biannual quantity at time t.

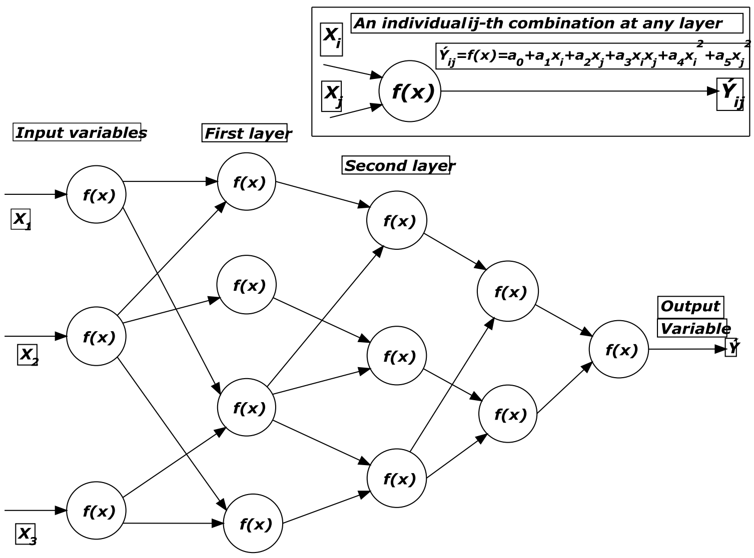

Figure 1. *GMDH-PNNs Structure with three inputs*

3.3.2. GMDH- Combinatorially Optimized (GMDH-CO)

Ivakhnenko demonstrated that a 2nd degree polynomial (called Ivakhnenko polynomials) given in (5) can reconstruct the complete K-G polynomial through an iterative perceptron or a multi-layer feed-forward combinations type procedure, which is its basic algorithm. In this method, an input observations series are considered as a matrix containing 'n' levels of observations over a set of 'm' variates. The 'n' observations are divided into two sets: learning and training or validation sets. Combinatorial model, linear in the parameter, is a truncated subset of terms of a polynomial function produced by a given set of variates. For modelling an output 'y'by three input variates x_1, x_2 and x_3, the usual quadratic polynomial function that will be optimized is as under:

$$y = a_0+a_1x_1+a_2x_2+a_3x_3+a_4x_1x_2+a_5x_2x_3+a_6x_1x_3+a_7x_1^2+a_8x_2^2+a_9x_3^2 \quad (13)$$

Another important way is to consider partial models i.e. a brute force combinatorial search, which consists of either truncated or complete combinatorial model. Although this model reduction approach has some advantages over PNN but computational task grows enormously and thus becomes less effective for objects with more than 30 inputs when full search is performed. Combinatorial GMDH selects an optimally complex model with a subset of terms of complete polynomial having least model error. At the stage of data pre-processing, different operators are allowed to apply to variates x_1 and x_2 e.g. exponent, sigmoid function, time series lags, moving averages, etc. yet the resulting model will be linear in the parameters. It is successful in outperforming linear regression approach for some positive noise level in the input data.

GMDH-COAlgorithms Combinatorial algorithms probe exhaustively among all candidate models. It produces models of all feasible input variate combinations and decides finally the best model from among the produced set of models consistent with a pre-specified choice criterion relevant to the optimum non-physical model. It employs complete mathematical induction method just to avoid missing any possible model. It sorts the models via progressively increasing the terms from 1 to m (i.e. the number of arguments) while a minimum value of an external criterion(a loss function type) in the plane of complexity indicates the optimum solution between models with the same complexity. It provides complex polynomial in independent variables. It selects the structure of the model itself without prior information about relationship in the form of the model in (4).

The technique involves fitting of quadratic equations for all pairs of independent variables and identifying a few best performers in terms of predictive ability (using appropriate statistics). Then converting entire set of independent variables (called zero generation variables) to new variables (first generation variables), which are obtained as predicted values from these selected quadratic equations (of zero generation variables). The process of fitting and identifying best quadratic equations is repeated using first generation variables and second generation variables are obtained. The whole process is repeated with every new generation of variables till appropriate model is obtained (using certain criteria). At final stage, one best quadratic equation is selected as the final model (Bahuguna & Chandrahas, 1992). Apposite to GMDH-PNNs this algorithm can't be halted at the specific level of complexity due to fact that a point of marginal enlargement of magnitude of the criterion can be a local minima rather than global minima. Steps involved in GMDH-combinatorial algorithms are:

1. Splits the observations into two sets: the learning and the testing subsets
2. Layers of partially described models with growing complexity

3. These partially described models for learning sub samples are estimated by OLS method.

4. Next magnitude of external criterion is computed based on testing subsample.

5. Choosing the best model/models obtained by minimal value of the external criterion.

All these steps are depicted in figures 1-3 when single layer is involved. For more than one layer, figure4 can help understand the algorithm.

Observations	Y	X – Independent Variates			
For Learning/Training	Y_1	X_{11}	X_{12}	. . .	X_{1m}
	Y_2	X_{11}	X_{12}	. . .	X_{2m}
	.	.	.	. . .	.
	Y_{nt}	$X_{nt,1}$	$X_{nt,2}$	. . .	$X_{nt,m,}$
For Testing/Validation	.	.	.	. . .	.
	Y_n	X_{n1}	X_{n2}	. . .	X_{nm}

Figure 1. *Sample Splitting*

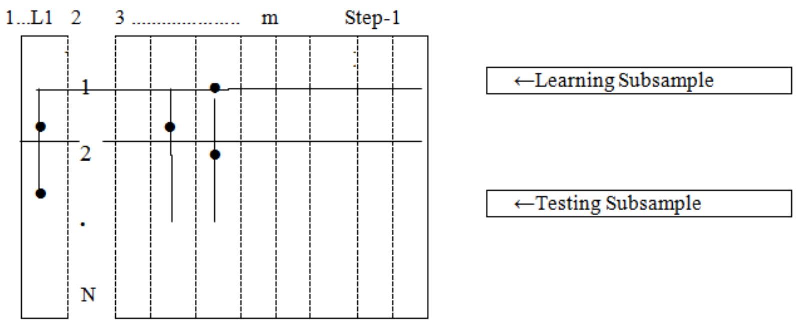

Figure 2. *Selection from 'm' arguments from learning sample.*

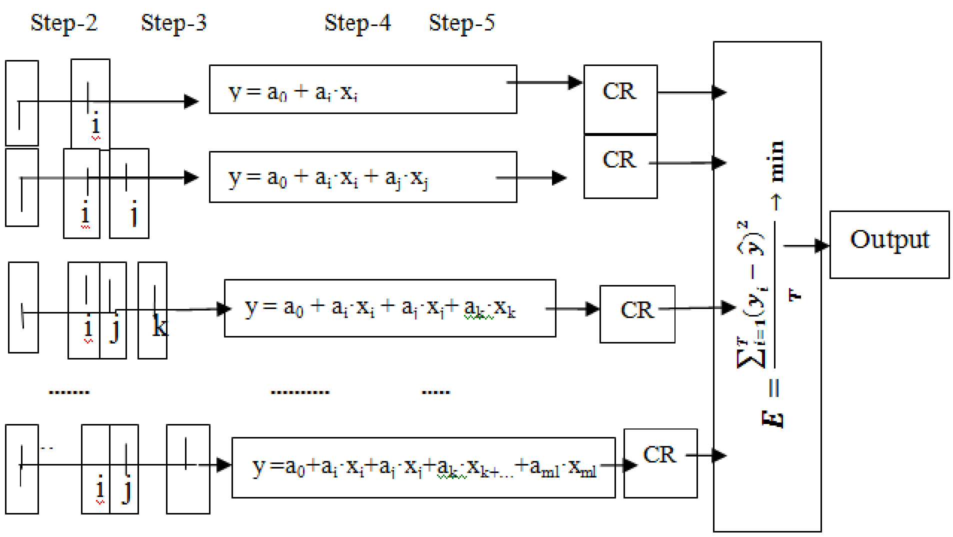

Figure 3. *Single layer of partial descriptions with gradual growth in complicity*

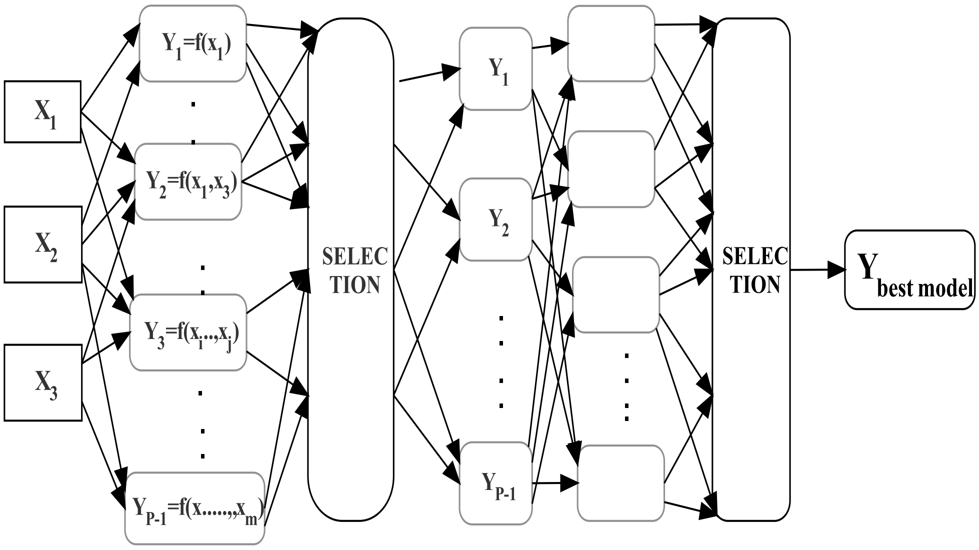

Figure 4. *Two layers of partial descriptions with gradual growth in complicity*

3.3. Combining Forecasts or Forecast Combination

For such complex structure forecasts, only one forecasting model, linear or nonlinear, cannot entirely capture inherent characteristics veiled in the observed series thus leads to inaccurate and futile forecasting exercise. It looks reasonable to apply many different multivariate models and from these models combine different sets of forecasts by simple or some efficient method to obtain improved and accurate forecasts that are benefitted from strength of great many estimation frameworks. Alternatively one fits several forecasting models and choose a model giving best performance in the in-sample period. The empirical practice has revealed that the best descriptive model might not be best forecaster of future values. Characteristically, time series are facing time varying state of affairs, or possibly facing regime switching out-rightly. This problem is further intensified by model misspecification and errors in parameter estimation.

A way out is to use various improved forecasting models and use forecasts combination to improve the forecast. That is why this study adopts forecasts combinations methodology, which combines and GMDH models to take advantage of the combined strengths of parametric and nonparametric models by optimal simple combining forecasting method. Combining forecasts frequently improve upon the individual forecasts and have a long been applied in econometrics (Timmermann, 2006). This method is less vulnerable to structural changes present in individual forecasting regressions as these, in effect, balance out intercept shifts (Hendry & Clements, 2004). In this technique, forecasts from multivariate models (each set up in a different set of predictors, lag lengths, or specifications) are combined. Though combining forecasts often outperform individual forecasts usually, but do not largely outperform factor-based forecasts, rather former are frequently a bit worse than later.

3.4. Forecasting Accuracy Criterions

This study uses the following forecasting accuracy criterions for measuring forecast accuracy for different specifications: The difference between the observed and the predicted values for the corresponding period is termed as forecast error i.e., $E_t=y_t-\hat{y}_t$, where E_t is the prediction error at period t, y_t is the observed value at period t, and $\hat{y}_t$ is the forecast for period t. The measures of aggregate error to be used are:

Forecast error = $e_t = y_t - \hat{y}_t$, =Actual – Forecast (14)

$$\text{Mean Absolute Error (MAE)} = \frac{\sum|Actual\ -Forecast\ |}{n} = \frac{\sum_{t=1}^{n}|y_t - \hat{y}_t|}{n} \quad (15)$$

$$\text{Mean Absolute Percent Error(MAPE)} = \frac{\sum|\frac{Actual\ -Forecast}{Actual}|}{n} = \frac{\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|}{n} \quad (16)$$

$$\text{Mean Squared Error (MSE)} = \frac{\sum(Actual\ -Forecast\ )^2}{n} = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n} \quad (17)$$

$$\text{Root mean square error (RMSE)} = \sqrt{\frac{\sum(\text{Actual} - \text{Forecast})^2}{n}} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \quad (18)$$

3.6. Data and Variables Series Involved

For Pakistan mostly the research is done after the year of 1972 due separation of its eastern part. Thus, for five main variates are to be used to forecast inflation, the price level (GDP-IPD)denoted by P_t, are monetary aggregates SSMAs M2, DMI M2, real GDP, and IR(i.e. discount rate of SBP) denoted by G_t, R_t, S_t and D_t, respectively. The sources of data series including those used to construct Divisia Index of M2 are SBP (2010) and SBP-MSB (2011). The biannual series starting from middle of first half year of 1972 to middle of last half year of 2009 providing 76 observations i.e. from 1972I to 2009II based at 1999-00 prices have been used. The DMI constructed from SSMA M2 according to the procedure detailed in section 3.1 of Iqbal et al. (2015). The biannual ratios of GDPwere raised from quarterly estimates provided by Arby (2008). For the period, 2005/6 to 2009/10, where Arby's estimates quarterly estimates are not available, the estimates of biannual GDP were obtained by employing the average of 2000-01 to 2004-05 provided by Arby (2008) and raising the post 2004-05 quarterly ratios.

With a view to know and compare the contribution of the monetary aggregates DMI denoted by Dt and SSMA denoted by St in forecasting inflation, only one is to be used in each of the model along with two other variates G_t and R_t. Thus, the two specifications consisting of two sets of variables will be estimated with GMDH-PNN, and GMDH-CO methods as under:

$$P = f(G_t, R_t, D_t) \quad (19)$$

$$P = f(G_t, R_t, S_t) \quad (20)$$

4. Results Discussions and Comparisons

First of all the graphs of the variates modelled are shown in figures 5 and 6along with box-plot which better elaborates the descriptive statistics. The graph of series of price level P seems nonlinear in nature whereas graph of other series are commonly found in macroeconomic time series.

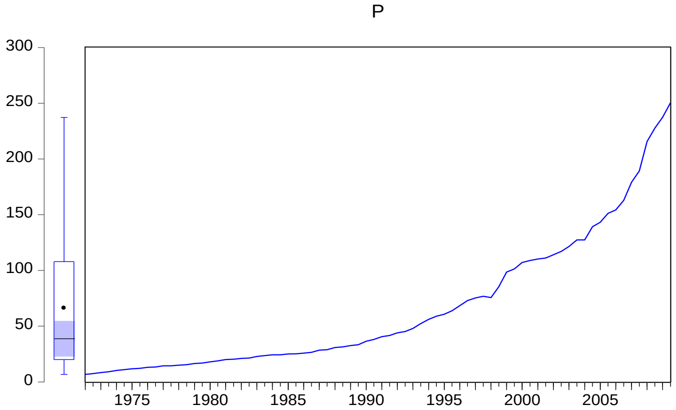

Figure 5. *Historigrams of price level*

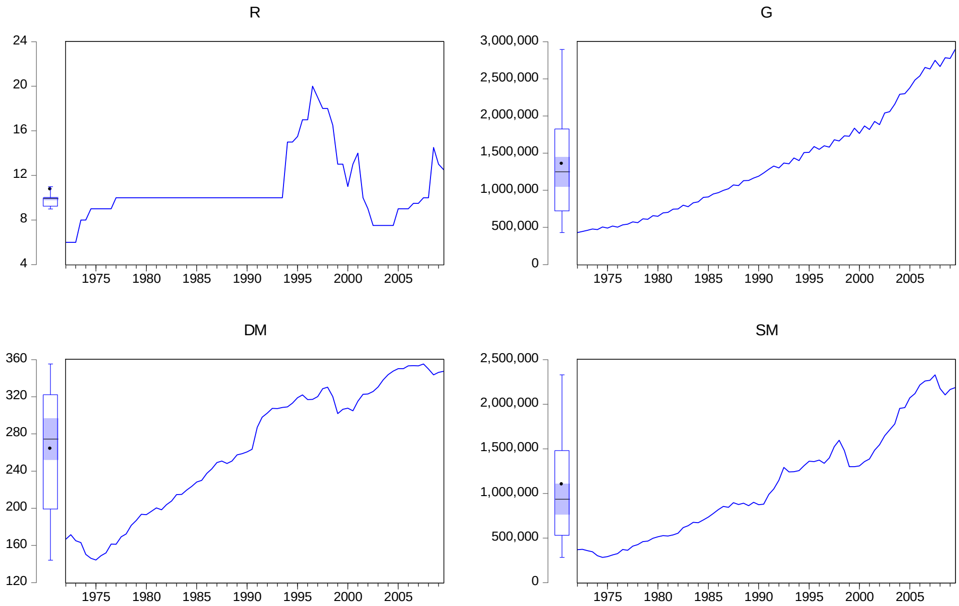

Figure 6. *Historigrams of R, G, D and S*

4.1. GMDH-Type NNs Model Forecast

The GMDH algorithms make it mathematically possible to formulate a model of optimal complexity having forecast optimization. In the adaptive or self-organizing manner, as complexity gradually increases, the computer finds by shifting the different models, the minimum of a selection criterion. For which the computer has been asked to look for to obtain a dynamic nonlinear model of optimum complexity for long-term prediction and forecasting of price level based on the observed biannual data.

The GMDH-PNN ensemble K-G polynomial of order 2 by including only those pairs of inputs combinations, which contribute sizably in predicting the output, and from these combinations constructs the universal approximators. Many transformations are likely to play role in improving the forecasting performance of the GMDH models, and these can be embedded in the observed series to incorporate smoothing or transmission effects for seasonal or lagged variation. To get seasonal smoothing in biannual data series, two periods moving average has been used along with 0 to 2 lags to enhance forecast accuracy. The validation strategy consisted of training and testing with a ratio of 9:1.The model complexity is limited to 11 terms only just to avoid overfitting in GMDH-CO and in GMDH-PNNs neurons, input was limited to two with model complexity limited to 11and two layers at maximum. The criterion, which corresponds to RMSE, values for the whole models ranges between .026 to .031; a reasonably low value in magnitude, showing that predictions and forecasts are quite close to actual values. The graphs of the forecasts show that GMDH-CO seems better in forecasting performance. The results are presented along with output details and plots for the residuals in dual sets of figures for each four models numbered7-14 in the appendix with model description between each set of figures. The -periods forecasts and corresponding error are depicted in red colour in the figures. The variates derived from monetary aggregation i.e. DMI and SSMA evidently emerged significant predictors due to their presence or usage in all the respective models. The input combinations are pruned by RMSE and only those combinations remain in the final model, which contributed sizably in the predictions. In GMDH-CO with DMI, the predictor R is missing and then R is missing again in GMDH-PNN with SSMA, which shows insignificance of predictor R with regard to forecasting inflation. However, numerical 7-periods forecasts are presented in tables 1 and 2below.

4.2. Forecasting Performances Measured by all the Four Criterions

The forecasting performance of GMDH-PNNs, and GMDH-CO models entailing DMIs with their two forecasts combinations: first comprising of forecasts from all the models and the second comprising of the GMDHs only is compared. To compare the performance of the models, four forecast accuracy criterions have been used: MSE, RMSE, MAE, and MAPE. The GMDH-CO is unanimously outperformed the competing models, with combination of GMDHs taking the second position and the GMDH-PNN performed third best.

Table1. *Fitted Model (using DMAs)* with Four Forecast Accuracy Measures

DMA			
ACTUALS	COMB	PNN	Combined
162.84	162.75	162.1	162.42
179.2	178.04	177.77	177.9
189.21	189.88	189.54	189.71
215.61	212.53	213.03	212.78
227.66	226.88	228.03	227.45
237.35	235.7	237.28	236.49
250.61	253.45	256.88	255.16
MSE	3.245	6.982	4.526
RMSE	1.801	2.642	2.127
MAE	1.467	1.685	1.524
MAPE	0.007	0.008	0.007

Table 2. *Fitted Model (using SSMAs) with Four Forecast Accuracy Measures*

SSMA			
ACTUAL	COMB	PNN	Combined
162.84	171.54	159.88	165.71
179.2	181.25	170.87	176.06
189.21	200.57	187.88	194.23
215.61	212.08	200.3	206.19
227.66	218.59	225.85	222.22
237.35	233.97	242.47	238.22
250.61	247.87	252.75	250.31
MSE	46.107	49.804	23.221
RMSE	6.79	7.057	4.819
MAE	4.956	5.288	3.867
MAPE	0.024	0.026	0.019

The forecasting performance of GMDH-PNNs, and GMDH-CO models entailing SSMAs and their two forecasts combinations: first comprising of forecasts from all the models and the second comprising of the GMDHs is compared. The forecasts combination of GMDHs is unanimously outperformed the competing models, with GMDH-CO taking the second position and the GMDH-PNN performed third best. When the forecasts accuracy with regard to methods of aggregation is considered, the models involving DMIs have outperformed their respective models entailing SSMAs by substantial margins. A substantial gain in accuracy can be achieved by using Divisia measures of monetary aggregation instead of Simple summation. If central banks attempt forecast inflation usingthe DMI, the evidence from this study supports that they will gain significantly in forecast accuracy.

4.3. Final results and discussions

1) The forecasts using DMIs are better than the forecasts with SSMAs. Hence, with regard to method of aggregation, DMI dominated and outperformed the SSMAs w.r.t. all forecast accuracy criterions used. 2) The GMDH–CO forecastsproved dominant methodology w.r.t. all forecast accuracy criterions used in both methods of aggregation. 3)For the both ofaggregationmethods, the monetary aggregates play significant role in predicting inflation as their coefficients are significant in parametric models and their coefficients are present in the final prediction model and these have not been pruned out due insignificant part in the forecasts.

4.4. Findings and policy implications on projections

1*)* To the question of ongoing debate whether the monetary variables still play a significantrole in the predicting and forecasting of inflation, this study concludes that monetary aggregate play dominant role in the task, since both monetary aggregate are significantly present in inflation forecasting models and not been pruned out. 2) Further, parametric models are too much restrictive and cannot apprehend a complex nonlinear structure unless the functional form is fully known or assumed in advance. However, the complex functional structure is generally unknown to researcher and hence parametric models poses difficulties in modelling and further these further these models cannot forecast accurately the unknown complex nonlinear structures. 3) The shape of the price level graph shows essentially nonlinear pattern and needs to be dealt accordingly in forecasting excises especially by SBP in its routine inflation forecasts. 4) The task of forecasting can better be handled by non- or semi-parametric models, and among them, the universal approximators are the best. Most of the universal approximators require more observation for learning and fail to apprehend the true form in small samples; however, GMDH-CO can perform more accurately even in small samples. GMDH-PNN require more observations to perform more accurately. 5) Nonlinear parametric models a priori require the knowledge about the functional form of the model to be estimated which is seldom known and most of the time is unknown to the researcher. Hence they fail to qualify on merit for complex nonlinear structures. Forecasting models are more concerned about the exact underlying functional form than then the mere descriptive model. The non-parametric nonlinear model, universal approximators can better handle the task of forecasting under conditions of unknown functional form and limited number of observations. 6) Method of monetary aggregation is an outstanding question economics. In this regard, we conclude here that in all models incorporating DMI has outperformed models entailing SSMA in forecasting inflation. Hence, this study recommends the use of DMI instead of SSMAs in routine forecasting inflation exercises of SBP.

Notes

[1] Polynomials functions are usefully utilized in econometrics due to Weierstrass's (1885) function approximation theorem that states "any continuous function on a closed and bounded interval can be approximated by polynomials", i.e. if x ∈ [a, b], for any ϙ> 0 there exists a polynomial p (x) ∈ [a, b] such that |f (x) − P (x)| <ϙ∀x ∈ [a, b].

Appendixes

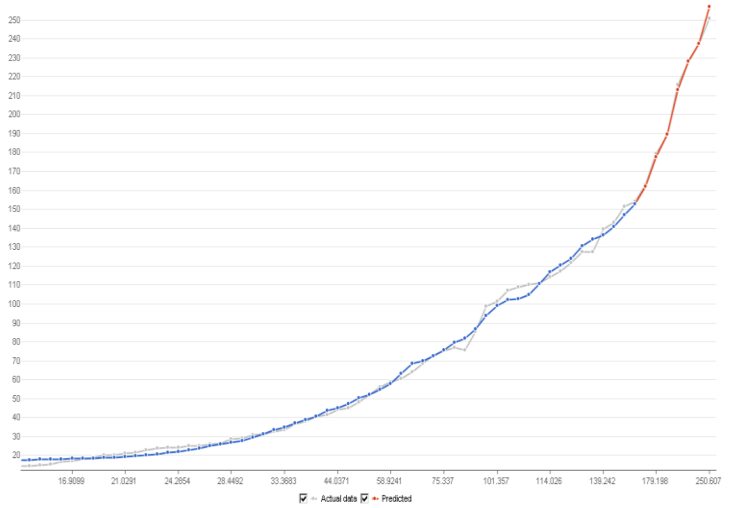

Figure 7. *GMDH- Neural Networks* Fitted Model (using DMAs) *and Forecasts*

GMDH-Type Neural Networks Fitted Model (DMA)
Y= 29.67 + "G@1, ȳt(2)"*(-5.65e-05) + "G@1, ȳt(2)"*"DMI@1, ȳt(2)"*1.58e-07 + "G@1, ȳt(2)"^2*4.47e-11 + "DMI@1, ȳt(2)"^2*(-0.0003)
Criterion-Value =0.0282

Variable	Usage
G@1, ȳt(2)	2
DMI@1, ȳt(2)	1
R@1, ȳt(2)	1

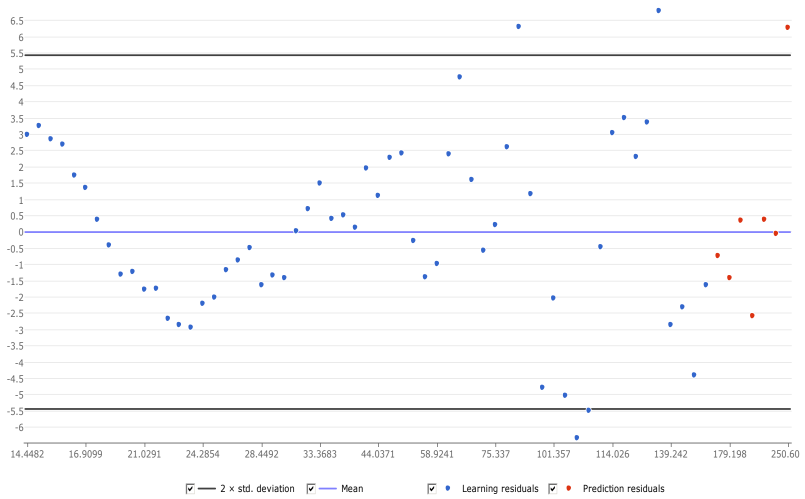

Figure 8. *GMDH-Type Neural Networks*Fitted Model's (using DMAs)*Residuals*

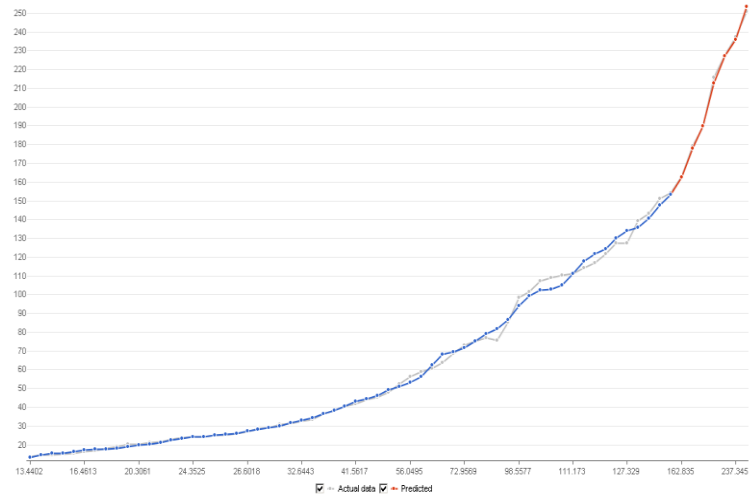
Figure 9. *GMDH-Combinatorially* Fitted Model (using DMAs) *and Forecasts*

GMDH- Combinatorially Fitted Model (DMA)
Y="G@1, ȳt(2)"*0.0005 + "G@1, ȳt(2)"*"DMI@1, ȳt(2)"^(-1)*(-0.10) + "G@1, ȳt(2)"^(-1)*(-2.08+08) + "G@1, ȳt(2)"^(-1)*"DMI@1, ȳt(2)"*1.04e+06 + "DMI@1, ȳt(2)"*(-1.48) + "DMI@1, ȳt(2)"^(-1)*62297.51
Criterion-Value =0.0264

Term	Usage
DMI@1, ȳt(2)	3
DMI@1, ȳt(2)^-1	3
G@1, ȳt(2)	3
G@1, ȳt(2)*DMI@1, ȳt(2)^-1	3
G@1, ȳt(2)^-1	3
G@1, ȳt(2)^-1*DMI@1, ȳt(2)	3
(constant term)	1
G@1, ȳt(2)*DMI@1, ȳt(2)	1

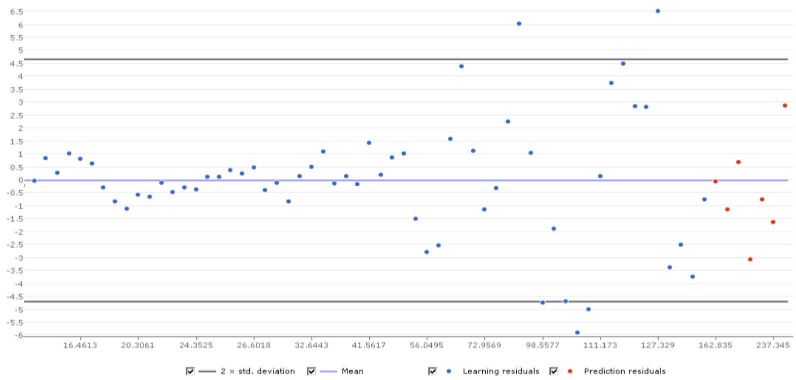
Figure 10. *GMDH-Combinatorially* Fitted Model's (using DMAs) *Residuals*

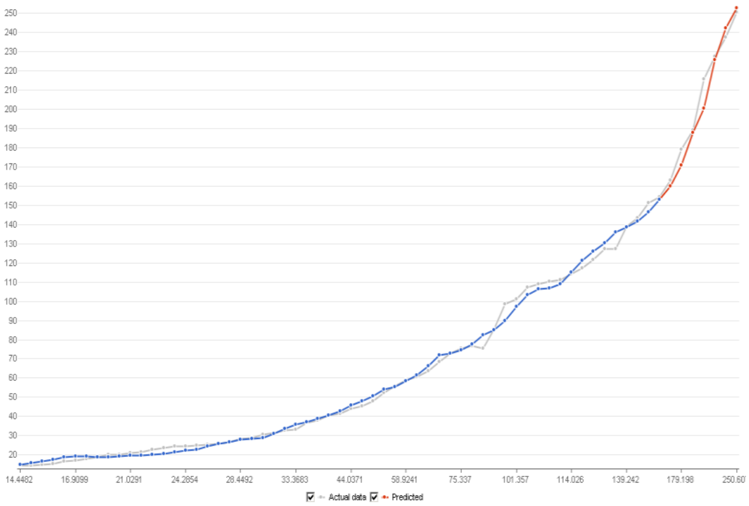

Figure 11. GMDH- Neural Networks Fitted Model (using SSMAs)and Forecasts

GMDH-Type Neural Networks Fitted Model (SSMA)
Y= 19.26 + "G@2, ȳt(2)"*"SSMA@2, ȳt(2)"*1.08e-10 + "SSMA@2, ȳt(2)"*(-4.10e-05) + "SSMA@2, ȳt(2)"^2*(-4.78e-11)
Criterion-Value =0.0299

Term	Usage
G@2, ȳt(2)	1
SSMA@2, ȳt(2)	1

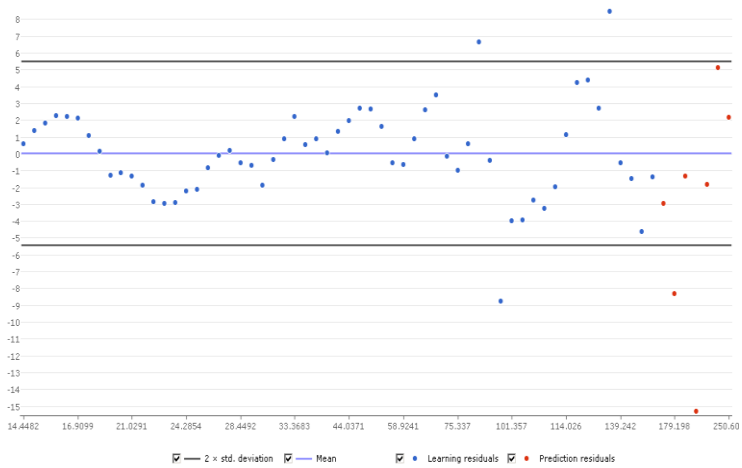

Figure 12. *GMDH- Neural Networks* Fitted Model's (using SSMAs) *Residuals*

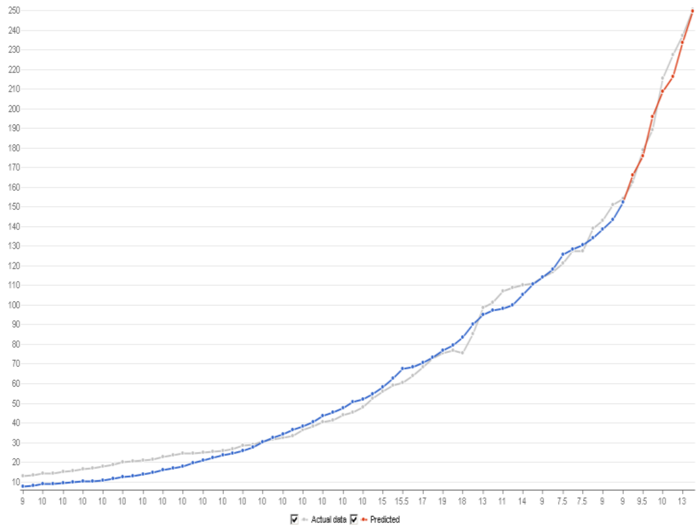

Figure 13. *GMDH-Combinatorially* Fitted Model (using SSMAs)*and Forecasts*

GMDH-Combinatorially Fitted Model (SSMA)
y ="G, ȳt(2)"^2*3.96e-11
Criterion-Value =0.031

Term	Usage
R, ȳt(2)	2
DMI, ȳt(2)	1
G, ȳt(2)	1

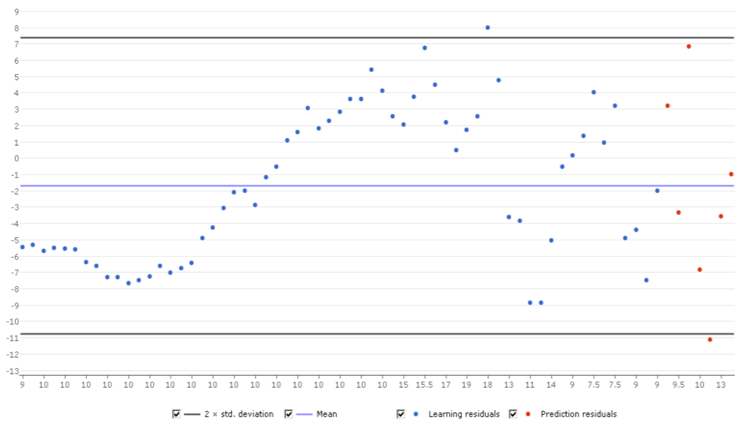

Figure 14. *GMDH-Combinatorially* Fitted Model's (using SSMAs)*Residuals*

References

Abdullah, M., & Kalim, R. (2009). *Determinants of food price inflation in Pakistan*. Paper presented in the conference of University of Management Sciences, 1-21.

Abidemi, O.I., & Malik, S.A.A. (2010). Analysis of Inflation and its determinant in Nigeria. *Pakistan Journal of Social Sciences*, 7(2), 97-100. doi. 10.3923/pjssci.2010.97.100

Ahangar, R.G., Yahyazadehfar, M., & Pournaghshband, H. (2010). The comparison of methods artificial neural network with linear regression using specific variables for prediction stock price in Tehran stock exchange, *International Journal of Computer Science and Information Security*, 7(2), 38-46.

Ang, A., Bekaert, G., & Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54(4), 1163-1212. doi. 10.1016/j.jmoneco.2006.04.006

Bahuguna, G.N., & Chandrahas, (1992). *Models for forecasting aphid pest of mustard crop*. IASRI Publication.

Bashir, F., Nawaz, S., Yasin, K., Khursheed, U., Khan, J., & Qureshi, M.J. (2011). Determinants of inflation in Pakistan: An econometric analysis using johansen co-integration approach. *Australian Journal of Business and Management Research*, 1(5), 71-82.

Bates, J.M. & Granger C.W.J. (1969). Combination of forecasts. *Operational Research Quarterly*, 20(4), 451-468. doi. 10.2307/3008764

Berger, H., & Österholm, P. (2011). Does money growth granger cause inflation in the Euro area? Evidence from out-of-sample forecasts using Bayesian VARs. *Economic Record*, 87(276), 45-60. doi. 10.1111/j.1475-4932.2010.00657.x

Binner, J.M., Gazely, A.M., Chen, S.H. & Chie, B.T. (2004). Financial innovation and Divisia money in Taiwan: comparative evidence from neural networks and vector error correction forecasting models, *Contemporary Economic Policy*, 22(2), 213-24. doi. 10.1093/cep/byh015

Binner, J., Tino, P., Tepper, J., Andersen, R., Jones, B., & Kendall, G. (2010). Does money matter in inflation forecasting? *Physica A:Statistical Mechanics and its Applications*, 389(21), 4793-4808. doi. 10.1016/j.physa.2010.06.015

BLS, (1997) *US bureau of labor statistics: Measurement issues in the consumer price index*. [Retrieved from].

Bokil, M., & Schimmelpfennig, A. (2005), Three attempts at Inflation Forecasting in Pakistan, *IMF Working Paper*, No. 05/105.

Canova, F. (2007). G-7 Inflation forecasts: Random walk, Phillips curve, or what else? *Macroeconomic Dynamics*, 11(1), 1-30. doi. 10.1017/S136510050705033X

Castle, J.L., & Hendry, D.F. (2006). Extending the boundaries of PcGets: Nonlinear models, *Working Paper*, Department of Economics, Oxford University.

Castle, J.L., & Hendry, D.F. (2010). A low-dimension portmanteau test for non-linearity, Journal of Econometrics, 158(2), 231-245. doi. 10.1016/j.jeconom.2010.01.006

Chaudhary, M.A., & Chaudhary M.A.S. (2006). Why the state bank of Pakistan should not adopt inflation targeting. *SBP Research Bulletin*, 2(1), 195-209.

Chaudhuri, R. (2012). *A combinatorial GMDH approach to identification, modelling and prediction of money market state variables*, International Conference on Technology and Business Management, March 26-28.

Clark, T.E., & McCracken, M.W. (2010). Averaging forecasts from VARs with uncertain instabilities. *Journal of Applied Econometrics*, 25(1), 5-29. doi. 10.1002/jae.1127

Clemen, R.T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4), 559-583. doi. 10.1016/0169-2070(89)90012-5

Diebold, F.X. & Lopez, J.A. (1996). Forecast evaluation and combination. in Maddala & Rao, eds., *Handbook of Statistics*, Elsevier: Amsterdam.

Dorsey, R. (2000). Neural networks with Divisia money: better forecasts of future inflation? in *Divisia Monetary Aggregates; Theory and Practice*, edited by M.T. Belongia, & J.M. Binner. Basingstoke, pp. 28-43. UK: Palgrave Publishers.

Drake, L. & Mills, T. (2005). A new empirically weighted monetary aggregate for the United States. *Economic Enquiry*, 43(1), 138-157. doi. 10.1093/ei/cbi010

ECB, (2003). *The European Central Bank (ECB) ECB's monetary policy strategy after the evaluation and clarification of May 2003*. Speech by Jean-Claude Trichet, 20 November.

Elger, C.T., Jones, B.E., & Nilsson, B. (2006). Forecasting with monetary aggregates: Recent evidence for the United States. *Journal of Economics and Business*, 58(5-6), 428-446. doi. 10.1016/j.jeconbus.2006.06.004

Haider, A. & Hanif, M.N. (2009). Inflation forecasting in Pakistan using artificial neural networks. *Pakistan Economic and Social Review*, 47(1), 123-138.

Hale, G., & Jorda, O. (2007). *Do monetary aggregates help forecast inflation?* FRBSF Economic Letter No. 2007-10.

Hashem, S., Yih, Y., & Schmeiser B. (1993). An efficient model for product allocation using optimal combinations of neural networks. In *Intelligent Engineering Systems through Artificial Neural Networks*, C. Dagli, L.I. Burke, B.R. Fernandez, & J. Ghosh, (eds.), 3:669-674, ASME Press.

Hashem, S. (1997). Optimal linear combinations of neural networks. *Neural Networks*, 10(4), 599–614. doi. 10.1016/S0893-6080(96)00098-6

Hendry, D.F. & Clements, M.P. (2004), Pooling of forecasts, *Econometrics Journal*, 7(1), 1–31. doi. 10.1111/j.1368-423X.2004.00119.x

Inoue, A. & Kilian, L. (2008). How useful is bagging in forecasting economic time Series? A case study of U.S. CPI inflation. *Journal of the American Statistical Association*, 103(482), 511-522. doi. 10.1198/016214507000000473

Iqbal. S., Sial, M.H., & ul Hassan, N. (2015). Inflation welfare cost analysis for Pakistan: An ARDL approach, *Pakistan Journal of Commerce and Social Sciences Johar Educational*, 9(2), 380-417.

Ivakhnenko, A.G. (1968). The group method of data handling - a rival of the method of stochastic approximation. *Soviet Automatic Control* c/c of *Avtomatika*, 1(3), 43-55.

Ivakhnenko, A.G. (1988). Sorting methods for modelling and clusterization (survey of the GMDH papers for the years 1983-1988). The present stage of GMDH development. *Soviet Journal of Automation and Information Sciences* c/c of *Avtomatika*. 21(4), 1-13.

Ivakhnenko, A.G. & Kocherga, Y.L. (1983). Theory of two-level GMDH algorithms for long-range quantitative prediction. *Soviet Automatic Contr ol* c/c of *Avtomatika* 16(6), 7-12.

Jaditz, T., Riddick, L.A., & Sayers, L. (1998). Multivariate nonlinear forecasting. *Macroeconomic Dynamics*, 2(3), 369-382.

Kock, A.B. (2009). *Forecasting with universal approximators and a learning algorithm,* Working paper, CREATES, Aarhus University.

Kock, A.B., & Terasvirta, T. (2011). Forecasting with nonlinear time series models, in M. P. Clements & D. F. Hendry (eds), Oxford *Handbook of Economic Forecasting*, Oxford University Press, Oxford, pp. 61-87.

Kovanen, A. (2011). *Does Money Matter for Inflation in Ghana?* IMF Working Papers 11/274.

Krolzig, H.-M. & Hendry, D.F. (2001). Computer automation of general-to-specific model selection procedures. *Journal of Economic Dynamics and Control* 25(6-7), 831-866. doi. 10.1016/S0165-1889(00)00058-0

Leung, M.T., Chen, A.S. & Daouk, H. (2000). Forecasting exchange rates using general regression neural networks, *Computers and Operations Research*, 27(11), 1093-1110. doi. 10.1016/S0305-0548(99)00144-6

Marcellino, M. (2002). *Instability and non-linearity in the EMU*, Discussion Paper No. 3312, Centre for Economic Policy Research.

Marcellino, M. (2004). Forecasting EMU macroeconomic variables, *International Journal of Forecasting*, 20(2), 359-372. doi. 10.1016/j.ijforecast.2003.09.003

Muller, J.A., & Ivakhnenko, I.G. (1996). Self-*organizing modelling in analysis and prediction of stock market*. In Proceedings of the Second International Conference on Application of Fuzzy Systems and Soft Computing – ICAFS'96, Siegen, Germany, pp.491-500.

Samsudin, R., Saad, P., & Shabri, A. (2011). River flow time series using least squares support vector machines. *Hydroogy and Earth System Sciences*, 15, 1835–1852. doi. 10.5194/hess-15-1835-2011

SBP, (2010). *Handbook of Statistics on Pakistan Economy 2010.*The State Bank of Pakistan..

SBP-MSB, (2011). Monthly Statistical Bulletin. SBP, Karachi January, and various previous Issues.

Schunk, D.L. (2001). The Relative Forecasting Performance of the Divisia and Simple Sum Monetary Aggregates. *Journal of Money, Credit and Banking*, 33(2), 272-83. doi. 10.2307/2673885

Stepashko, V.S., & Yurachkovskiy, Y.P. (1986). The present state of the theory of the group method of data handling. *Soviet Journal of Automation and Information Sciences* c/c of *Avtomatika*, 19(4), 36-46.

Stock, J.H.. & Watson, M.W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2), 293-335. doi. 10.1016/S0304-3932(99)00027-6

Stock, J.H. & Watson M.W. (2003). *Has the Business Cycle Changed and Why?* NBER Working Paper No. 9127. doi. 10.3386/w9127

Terasvirta, T., Dijk, D.V., & Medeiros, M.C. (2005). Smooth transition autoregressions, neural networks, and linear models in forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, 21(4), 755-774. doi. 10.1016/j.ijforecast.2005.04.010

Timmermann, A. (2006). Forecast combination. In Elliott, G., Granger, C.W.J. & Timmermann, A. (eds), *Handbook of Economic Forecasts*, Amsterdam: Elsevier.

Urwatul-Wutsqa, D., Subanar, Guritno, S., & Sujuti, Z. (2006). *Forecasting Performance of VAR-NN and VARMA Models*. Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Their Applications School of Mathematical Sciences, Universiti Sains Malaysia, Penang, June 13-15.

Valle e Azevedo, J., & Pereira, A. (2010a). *Forecasting Inflation with Monetary Aggregates, Economic Bulletin and Financial Stability*. Report Articles, issue (Autumn): 151-168.

Varahrami, V. (2012). Good Prediction of Gas Price between MLFF and GMDH Neural Network.*International Journal of Finance and Accounting*, 1(3), 23-27. doi. 10.5923/j.ijfa.20120103.01

Weierstrass's, K. (1885). Über die analytische Darstellbarkeit sogenannter willkürlicher Funktionen einer reellen Veränderlichen, Sitzungsberichte der Akademie zu Berlin, 1885: 633-639 and 789-805.

White, H. (2006). Approximate nonlinear forecasting methods, in G. Elliott, C. W. J. Granger & A. Timmermann (eds), *Handbook of Economic Forecasting*, Elsevier, Amsterdam, 1: 459-512.

Yang, X.H., Wang, F.M., Huang, J.F., Wang, J.W., Wang, R..C., Shen, Z.Q. & Wang, X.Z. (2009). Comparison between radial basis function neural network and regression model for estimation of rice biophysical parameters using remote sensing. *Pedosphere*. 19(2), 176–188. doi. 10.1016/S1002-0160(09)60107-7

Yao, Y., & Ni, Q. (2009). *Oil price forecasting based on self-organizing data mining*. In Conference on Grey Systems and Intelligent Services, 10-12 Nov. 2009. GSIS 2009. IEEE International, 1386-1390.

Zheng, A., Liu, W., & Zhao, F. (2010). *Double trends time series forecasting using a combined ARIMA and GMDH model*. In Control and Decision Conference (CCDC), 26-28 May, 2010 Chinese: 1820-1824.