



Nonlinear and Nonparametric Models for Forecasting the US Gross National Product

Siddharth Arora^{1,2}, Max A. Little^{2,3} and Patrick E. McSharry^{1,2}

¹Smith School of Enterprise and the Environment, University of Oxford, UK

²Oxford Centre for Industrial and Applied Mathematics (OCIAM), Mathematical Institute, University of Oxford, Oxford, UK

³Oxford Centre for Integrative Systems Biology, Department of Physics, University of Oxford, Oxford, UK

Email: (S. Arora) arora@maths.ox.ac.uk, (M.A. Little) littlem@physics.ox.ac.uk,
(P.E. McSharry) patrick.mcsharry@smithschool.ox.ac.uk

Abstract– Nonlinear models, like the self-exciting threshold autoregressive (SETAR) model and Markov-switching autoregressive (MS-AR) model have been proposed for modeling the Gross National Product (GNP) time series. Both SETAR and MS-AR, however, require estimation of a large number of parameters relative to the small amount of GNP observations. While modeling the training data reasonably well, these models tend to overfit and perform poorly in terms of out-of-sample forecasting. The aim here is to investigate the efficacy of a novel parsimonious nonparametric and nonlinear model, which can outperform SETAR and MS-AR in terms of out-of-sample GNP forecasting accuracy. As it is important to quantify forecast uncertainty, leading to well informed policy-making, we generate both point and density forecasts. We evaluate point forecasts using the root mean square error (RMSE) and mean absolute error (MAE), while density forecasts are evaluated using the continuous ranked probability score (CRPS).

1. Introduction

The GNP provides a measure of the economic wealth of a country. It is one of the most commonly used macroeconomic indicators and is reported quarterly. The need for accurate and timely forecasts of the GNP time series stems from the requirements of efficient decision and policy-making. The complex dynamics underlying the GNP time series, along with limited number of post-war observations, however, make the task of generating reliable forecasts quite challenging [1].

The last two decades have witnessed a surge in nonlinear modeling techniques for characterizing economic time series [2, 3, 4, 5]. Deviations from the Gauss-Markov assumptions (linearity, homogeneity and independence) are indications of the presence of nonlinearity in the data generating process. The GNP time series is shown to follow an asymmetric cyclical process, giving rise to different regimes (growth and recession), with growth periods being much longer than recessionary periods [2]. Nonlinear models like the SETAR model [4,

5], and MS-AR models [3] have been employed for forecasting the GNP time series. The rationale behind these nonlinear models (SETAR and MS-AR) is to characterize regimes of growth and recession separately rather than treat them as one, as opposed to the linear models. Hence, SETAR and MS-AR are also known as regime-switching models.

Due to the complex model structure, the parameter estimation procedure for regime-switching models is not straightforward. Also, it has been shown that the forecast performance of SETAR is highly sensitive to the uncertainty in parameter estimates [6]. The major limitation of regime-switching models, however, lies in the fact that there is no clear consensus if these models (SETAR and MS-AR) are better than their linear counterpart (AR models), from the perspective of out-of-sample forecasting [7].

Motivated by Occam's razor and the impact of accurate GNP forecasts on policy-making and in turn on our financial markets, we seek parsimonious nonlinear models which are based on simple assumptions and can generate accurate point and density forecasts of US GNP. In this paper, we propose a nonlinear model that can be viewed as a hybrid between a nearest neighbour method [8] and the random analogue prediction (RAP) model [9]. The proposed model does not make prior assumptions regarding the true functional form generating the data, hence the model is nonparametric. We compare the competitiveness of the proposed model with classical linear and nonlinear models (AR, SETAR and MS-AR) in generating out-of-sample point and density forecasts.

The paper is arranged as follows. Section 2 gives details of the employed dataset, and discusses the classical models used previously for forecasting US GNP. This section also proposes a novel nonlinear and nonparametric model. Section 3 provides out-of-sample forecast comparison between different models and the random walk benchmark, while section 4 offers conclusion.

2. Linear and Nonlinear Modeling Approaches

2.1. Data

The GNP time series comprises seasonally adjusted quarterly prices of real US GNP (in billions of dollars). The time series contains US GNP recordings from 1947Q1 to 2008Q3 at 2000 prices. The recordings were obtained from the Federal Reserve Economic Data II (FRED II) affiliated with the Federal Reserve Bank of St. Louis. The original data was transformed to give quarterly percentage growth rate, and all the models were estimated using the transformed time series.

2.2. Autoregressive Model

An autoregressive (AR) model is a linear recursive model that belongs to the Box and Jenkins [10] class of models. The AR model characterizes the time series assuming linear relationship between observations. The number of parameters incorporated in the model corresponds to the order of the model. An AR model of order p is denoted as AR (p) and can be represented using the following form:-

$$x_t = \alpha_0 + \sum_{i=1}^p \alpha_i x_{t-i} + \varepsilon_t \quad (1)$$

where x_t is an observation in the time series at time t , α_0 is a constant, α_i denotes AR model parameters for $i = 1, 2, \dots, p$, while ε_t is an independent and identically distributed (IID) process with mean zero and variance σ^2 . In this paper, we use the Akaike Information Criterion (AIC) proposed in [11] for selecting the AR model order. The AR model parameters are estimated using ordinary least squares (OLS).

2.3. Self-exciting Threshold Autoregressive Model

Threshold Autoregressive (TAR) models were proposed in [12]. The TAR model divides the original time series into a number of distinct non-overlapping regimes, and models every regime as a separate linear process. SETAR models are a subclass of TAR models whereby the threshold variable is forced to be endogenous, i.e. the threshold is chosen from an observation in the time series.

A SETAR model composed of N_r number of distinct regimes, for a time series x_t , is denoted as SETAR ($N_r; p_1, p_2, \dots, p_{N_r}$), and is represented as:-

$$x_t = \alpha_{0j} + \sum_{i=1}^{p_j} \alpha_{ij} x_{t-i} + \varepsilon_{ij} \quad (2)$$

where α_{ij} are the autoregressive coefficients for a given regime index j that obeys $r_{j-1} \leq x_{t-d} < r_j$, while r_j for $j = 1,$

$2, \dots, N_r$ are the thresholds that divide the time series into different regimes, p_j is model order for the corresponding regime, whereas ε_{ij} is an IID process with mean zero and variance σ_j^2 . We estimate a bi-regime SETAR model for US GNP as employed previously in [4, 5]. To estimate the model parameters, we varied r_j and d over a grid, and select a particular set of parameter values that minimizes the overall residual sum of squares (RSS). For a given threshold and delay order, the AR model order and model coefficients were estimated by applying OLS in each regime separately.

2.4. Markov-switching Autoregressive Model

The bi-regime Markov-switching model (as proposed in [3]) with AR order p can be denoted as MS (2)-AR (p), and is defined as:-

$$x_t = \mu(s_t) + \sum_{i=1}^p \alpha_i (x_{t-i} - \mu(s_{t-i})) + \varepsilon_t \quad (3)$$

where x_t is an observation in the time series, α_i are the AR coefficients for $i = 1, 2, \dots, p$, s_t is a regime variable such that, $s_t = 1$ corresponds to growth and $s_t = 2$ corresponds to recession. The mean of the process $\mu(s_t)$ switches between the two regimes, such that $\mu(s_t)$ is positive if $s_t = 1$, and negative otherwise.

The transition between the two regimes depends on the variable s_t , governed by a first order Markov process with transition probabilities $P_{ij} = P(s_t = j \mid s_{t-1} = i)$. The parameter vector that needs to be estimated for MS-AR model is $\theta = \{\mu(1), \mu(2), \alpha_1, \alpha_2, \dots, \alpha_p, \sigma, P_{11}, P_{22}\}$. The parameter vector θ is chosen so as to maximize the likelihood of the observations. The estimation of this parameter vector was based on maximizing the likelihood using expectation maximization (EM), for details see [13].

2.5. fraction-Nearest Neighbor Model

The nearest neighbor method [8] is a nonparametric nonlinear approach which relies on the assumption that neighboring states have similar future outcomes (see for example [9, 14]). The fraction-Nearest Neighbor (f -NN) is an adaptive nearest neighbor model that defines a neighborhood size based on the fraction of total points in the training set. Denoting neighborhood in terms of a fraction has the advantage that the neighborhood size adapts to changes in local data density and total number of observations, as opposed to the case when neighborhood size is defined in terms of a fixed radius, or number of neighbors.

The rationale behind the f -NN model lies in estimating an optimum neighborhood fraction, and using the collective behavior of the selected neighbors in generating a forecast. This model requires estimation of parameter f ,

which denotes the size of the optimum neighborhood. Given an optimum neighborhood size, one can quantify the collective behavior of neighboring states using for example, the mean, or build a local linear model, and use the estimated local model for forecasting the time series.

To estimate the optimum neighborhood size, we first create a set of delay vectors using the time series for GNP quarterly growth rate. The dimension of the delay vector can be viewed as being similar to the order of the model. Having created a set of delay vectors, we compute the distance (Euclidean distance) of the current delay vector, with previous delay vectors.

For estimating the model parameter, we divide the in-sample data into two parts, such that the first half is employed for training and the second-half for validation. The value of f is varied from a minimum, so as to include only the closest delay vector to the current state, to a maximum, so as to include all states within the training set. For a given f , a set of delay vectors that lie within the neighborhood are selected from the training set. The future observations of these delay vectors are randomly sampled and issued as a density forecast. The value of f that minimizes the density forecast error (quantified using CRPS) on the validation set is chosen for generating out-of-sample forecasts.

3. Results and Observations

3.1. Performance Scores

The point forecast performance for a particular forecast horizon h is evaluated using the root mean square error (RMSE) and mean absolute error (MAE), given by:-

$$RMSE_h = \sqrt{\frac{1}{N-T-h+1} \sum_{i=T}^{N-h} (x_{i+h} - \hat{x}_{i+h|i})^2} \quad (4)$$

$$MAE_h = \frac{1}{N-T-h+1} \sum_{i=T}^{N-h} |x_{i+h} - \hat{x}_{i+h|i}| \quad (5)$$

where $RMSE_h$ (MAE_h) is the RMSE (MAE) at horizon h , x_{i+h} is the actual observation, $\hat{x}_{i+h|i}$ is the h -step ahead forecast, T is the forecast origin and N is the length of the time series. In order to evaluate the density forecast performance, we use the empirical form of continuous ranked probability score (CRPS). The CRPS is defined as:-

$$CRPS = E_F |X - x| - \frac{1}{2} E_F |X - X'| \quad (6)$$

where X and X' are independent samples being drawn from the forecasts density function, each having the same distribution F , E_F is the expectation with respect to the

distribution F , while x is the actual observation. For details regarding CRPS, please see [15].

3.2. Forecasting Scheme

A series of forecasts for horizons ranging from one to four quarters (one year) ahead is generated based on a rolling forecast scheme, as previously employed for GNP time series by [6]. The time series for GNP quarterly growth rate dating from 1947Q2-2008Q3, is divided into in-sample data from 1947Q2-1996Q4 for training and out-of-sample data 1997Q1-2008Q3 for testing. Using the training set, the specifications of the optimum models were estimated as- AR (4), SETAR (2; 1, 1), MS (2)-AR (5), and $f=0.13$.

All models were estimated using only the training set, and the estimated parameters were then held 'fixed' while making out-of-sample forecasts. Hence, each model had access to precisely the same amount of information as every other model during the forecasting stage. We generated density forecasts using the proposed and classical based on Monte Carlo simulations (size 10,000). In order to generate point forecasts, we issued the mean of the forecast distribution at each forecast horizon as the point forecast.

3.3. Forecast Comparison

We evaluate AR, SETAR and MS-AR and the proposed model (f -NN) on out-of-sample GNP data (1997Q1-2008Q3), using three different performance scores (RMSE, MAE and CRPS). We employ the random walk (RW) benchmark, whereby the current value in the time series is issued as a forecast for the next step ($h=1$). For multistep ahead forecast ($h>1$), the current observation is issued as a h -step n. Note that we can generate only point forecast using this benchmark, hence there are no RW statistics for CRPS. We generate and evaluate forecasts from one quarter ($h=1$) up to four quarters ($h=4$) ahead.

The evaluation of point forecasts is presented in Table I and II. When evaluation is based on RMSE (Table I), we find that all models outperform the RW benchmark at all forecast horizons. This signifies that with respect to the benchmark, all models have some skill in forecasting the GNP time series. On comparing different models, we find that f -NN is one of the best models, producing most accurate forecasts for all horizons, except for $h=2$, where SETAR is better. When forecast evaluation is based on MAE (Table II), we find that f -NN consistently outperforms all the models on all forecast horizons. SETAR generates more accurate point forecasts than AR and MS-AR.

The superior performance of the proposed model (f -NN) is further highlighted when forecast evaluation is based on CRPS. It is evident from Table III, that density

forecasts from f -NN are superior to any of the classical models. For point forecasts, SETAR is the most accurate model compared to AR and MS-AR, while MS-AR is found to be superior to both SETAR and AR in generating density forecasts on the out-of-sample data.

Table I

Out-of-sample point forecast evaluation using the **RMSE**, for forecast horizon (h) ranging from 1 to 4 quarters ahead. Least RMSE (most accurate) value at each horizon is depicted in **bold**.

H	RW	AR	$SETAR$	$MS-AR$	$f-NN$
1	0.780	0.588	0.579	0.592	0.550
2	0.648	0.572	0.552	0.586	0.560
3	0.791	0.602	0.598	0.628	0.568
4	0.712	0.606	0.607	0.632	0.577

Table II

Out-of-sample point forecast evaluation using the **MAE**. Least MAE (most accurate) value at each horizon is depicted in **bold**.

H	RW	AR	$SETAR$	$MS-AR$	$f-NN$
1	0.670	0.493	0.484	0.498	0.446
2	0.499	0.471	0.448	0.481	0.446
3	0.637	0.487	0.472	0.509	0.455
4	0.569	0.489	0.480	0.511	0.461

Table III

Out-of-sample density forecast evaluation using the **CRPS**. Least CRPS (most accurate) value at each horizon is depicted in **bold**.

h	AR	$SETAR$	$MS-AR$	$f-NN$
1	0.361	0.357	0.348	0.320
2	0.357	0.352	0.345	0.325
3	0.373	0.370	0.369	0.330
4	0.375	0.370	0.369	0.333

4. Conclusion

We proposed a simple nonlinear and nonparametric model that convincingly outperformed AR, SETAR and MS-AR models on multiple forecast horizons, when evaluated using different performance scores. These results point towards classical parametric models (linear and nonlinear) over-fitting the in-sample data, due to which they fail to generalize on the out-of-sample dataset. Given the need for quantifying uncertainty in forecasts for informed decision and policy-making, we emphasize the

need for evaluating models based on their ability to generate accurate density forecasts, as quantified by CRPS.

References

- [1] D. Orell and P.E. McSharry, "System economics: Overcoming the pitfalls of forecasting models via a multidisciplinary approach," *International Journal of Forecasting*, vol. 25, pp. 734–743, 2009.
- [2] T. Terasvirta and H.M. Anderson, "Characterizing nonlinearities in business cycles using smooth transition autoregressive models," *Journal of Applied Econometrics*, vol. 7, pp. 119–139, 1992.
- [3] J.D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, vol. 57, pp. 357–384, 1989.
- [4] S.M. Potter, "A nonlinear approach to US GNP," *Journal of Applied Econometrics*, vol. 10, pp. 109–125, 1995.
- [5] G.C. Tiao and R.S. Tsay, "Some advances in nonlinear and adaptive modeling in time-series," *Journal of Forecasting*, vol. 13, pp. 109–131, 1994.
- [6] M.P. Clements and H.M. Krolzig, "A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP," *Econometrics Journal*, vol. 1, pp. C47–C75, 1998.
- [7] F.X. Diebold and J.A. Nason, "Nonparametric exchange rate prediction," *Journal of International Economics*, vol. 28, pp. 315–332, 1990.
- [8] T.M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, pp. 21–27, 1967..
- [9] F. Paparella, A. Provenzale, L.A. Smith, C. Taricco and R. Vio, "Local random analogue prediction of nonlinear processes," *Physics Letters A*, vol. 235, pp. 233–240, 1997.
- [10] G. E. P. Box and G.M. Jenkins, "Time series analysis, forecasting and control," San Francisco: Holden-Day, 1970.
- [11] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Second International Symposium on Information Theory*, Budapest: Akademiai Kiado, pp. 267–281, 1973.
- [12] H. Tong, "On a threshold model," *Pattern Recognition and Signal Processing*, Amsterdam: Sijhoff and Noordoff, pp. 101–141, 1978.
- [13] J.D. Hamilton, "Analysis of time series subject to changes in regime," *Journal of Econometrics*, vol. 45, pp. 39–70, 1990.
- [14] F. Fernandez-Rodriguez, S. Sosvilla-Rivero, and J. Andrada-Felix, "Exchange-rate forecasts with simultaneous nearest-neighbor methods: evidence from the EMS," *International Journal of Forecasting*, vol. 15, pp. 383–392, 1999.
- [15] T. Gneiting, F. Balabdaoui, and A.E. Raftery, "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society*, vol. 69, pp. 243–268, 2007.