

On multi-step SETAR predictors¹

I previsori multi-step dei modelli SETAR

Alessandra Amendola
DISES
Università di Salerno
alamendola@unisa.it

Marcella Niglio
DISES
Università di Salerno
mniglio@unisa.it

Riassunto: Nel presente lavoro sono analizzati i previsori del modello SETAR. La distribuzione analitica in corrispondenza di differenti *lead time* (ℓ) ed alcune proprietà sono esaminate, evidenziando, inoltre, come l'impiego di tali distribuzioni possa accrescere l'accuratezza delle previsioni. Infine, un'applicazione empirica all'indice Nikkey 300 mette a confronto diversi metodi di previsione.

Keywords: nonlinearity, multi-step forecasts, SETAR, mixture.

1. Introduction

The generation of forecasts from nonlinear time series models has raised a relevant attention in the last years. In particular their wide application in the analysis of financial time series, makes the generation of point forecasts not adequate for most users which are mainly interested to evaluate their portfolio risk. In this context the derivation of density forecasts can be useful to provide some indications about the forecast uncertainty (see Tay and Wallis (2000) for a recent survey about density forecasts). Granger and Teräsvirta (1993), further argue that, in the nonlinear domain, the analytic distribution of predictors can be useful for the generation of multi-step point forecasts too, obtaining the so called *closed form* forecasts. In this context Amendola and Niglio (2001) have investigated the multi-step ahead forecasts from the *Self Exciting Threshold Autoregressive* (SETAR) model (Tong and Lim, 1980) and have derived the analytic distribution of its predictors. In the present paper the forecast densities of the SETAR predictors are discussed in Section 2. In Section 3 some further properties of these predictors are presented and an application to the Nikkey 300 stock market index is reported in the forth section. Some concluding remarks are given in the final section.

2. SETAR model predictors

The SETAR($k; p_1, p_2, \dots, p_k$) model (Tong and Lim, 1980) can be represented as follows:

$$X_t = \sum_{i=1}^k (\phi_0^{(i)} + \sum_{j=1}^{p_i} \phi_j^{(i)} X_{t-j} + \epsilon_t^{(i)}) \cdot \mathbf{1}(r_{i-1} < X_{t-d} \leq r_i) \quad (1)$$

where $|\phi_j^{(i)}| < 1$, for $i = 1, 2, \dots, k$ and $j = 1, 2, \dots, p_i$; $p_1, p_2, \dots, p_k$ are the autoregressive orders of the k regimes, $\mathbf{1}(\cdot)$ is an indicator function, the $\epsilon_t^{(i)}$ are sequences of i.i.d.

¹This paper is partially supported by MURST 2000: *Modelli stocastici e metodi di simulazione per l'analisi di dati dipendenti*.

gaussian white noise, $\epsilon_t^{(i)} \sim N(0, \sigma_i^2)$, d is the threshold delay and the threshold values r_i are such that $-\infty = r_1 < r_2 < \dots < r_k = +\infty$.

The generation of the multi-step forecasts $\hat{X}_t(\ell)$ from the SETAR model have been differently examined (see Clements and Smith (1997) and Potter (1999) for a survey). Many authors have highlighted that when the lead time ℓ is greater than the threshold delay d , the estimation of the threshold variable increases the prediction uncertainty and, in order to reduce the forecast approximation, explicit expressions for $X_{t+\ell|t}$ are needed. In this context Amendola and Niglio (2001) have derived the analitic expression of the conditional distribution of $X_{t+\ell|t}$, for $\ell = d + 1$, which is briefly shown in the followings for the SETAR(2; p_1, p_2) case in order to avoid cumbersome formulas.

Given model (1), with $k = 2$, when $\ell \leq d$ the best predictor of $X_{t+\ell|t}$, in term of minimum mean square errors, is given by:

$$\hat{X}_t(\ell) = \sum_{i=1}^2 \left(\phi_0^{(i)} + \sum_{j=1}^{\ell-1} \phi_j^{(i)} \hat{X}_t(\ell-j) + \sum_{j=\ell}^{p_i} \phi_j^{(i)} X_{t+\ell-j} \right) \cdot \mathbf{1}(r_{i-1} < X_{t+\ell-d} \leq r_i) \quad (2)$$

with $-\infty = r_0 < r_1 < r_2 = +\infty$. Given the assumptions in model (1) and extending the results of Box and Jenkins (1970) for the autoregressive models, it is easy to demonstrate that for $\ell \leq d$, $X_{t+\ell|t}$ is normally distributed with mean $\hat{X}_t^{(i)}(\ell)$ and variance σ_i^2 , $i = 1, 2$. These results cannot be extended to the case with $\ell = d + 1$ where the $X_{t+\ell|t}$ distribution is a mixture of two normal density $f(x, \lambda_1, \lambda_2, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$, with $\lambda_1 + \lambda_2 = 1$.

Assuming the coefficients of model (1) known and $k = 2$, the proportions λ_i , $i = 1, 2$, can be estimated using the number of the observations which belong to each regime. In particular given the indicator function $\mathbf{1}_t(\cdot)$ such that $\mathbf{1}_t(\cdot) = 1$, if $X_{t-d} \leq r_1$, and $\mathbf{1}_t(\cdot) = 0$ if $X_{t-d} > r_1$ (for $t = 1, \dots, n$), the proportion of the observations in the first regime is given as:

$$p_n = \frac{\sum_{i=1}^n \mathbf{1}_i(x)}{n} \quad \text{or equivalently} \quad p_n = \frac{[\#\{x_i \leq r_1\} | n]}{n}$$

with p_n the empirical cumulative distribution function of $\mathbf{1}_t(x)$ and n the series length. The independence among the $\mathbf{1}_i(x)$, for $i = 1, 2, \dots, n$, implies the convergence of p_n to the mixture proportion λ_1 , as n grows up:

$$\lim_{n \rightarrow \infty} p_n \rightarrow \lambda_1$$

where, given the assumption on p_n , the restrictions on λ_1 ($0 < \lambda_1 < 1$) are satisfied. From these results, the density function of $X_{t+\ell|t}$, for $\ell = d + 1$, is given as:

$$g(X_{t+\ell|t}, \theta) = \lambda \cdot \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot e^{\left\{-\frac{(X_{t+\ell|t} - \hat{X}_t^{(1)}(\ell))^2}{2\sigma_1^2}\right\}} + (1 - \lambda) \cdot \frac{1}{\sqrt{2\pi\sigma_2^2}} \cdot e^{\left\{-\frac{(X_{t+\ell|t} - \hat{X}_t^{(2)}(\ell))^2}{2\sigma_2^2}\right\}} \quad (3)$$

where θ is the vector of the estimated parameters and $\hat{X}_t^{(i)}(\ell)$ and σ_i^2 are the conditional mean and the variance of regime i (for $i = 1, 2$).

Following Granger and Teräsvirta (1993) the distribution in (3) can be used to obtain *closed form* point forecasts, given as:

$$\hat{X}_t^c(\ell) = \int_{-\infty}^{\infty} X_{t+\ell|t} \cdot g(X_{t+\ell|t} | \Omega_t) dX_{t+\ell|t} \quad (4)$$

which, in the SETAR case, gives more accurate forecasts than the *naïve* and *Monte Carlo* predictions (Amendola and Niglio, 2001).

3. Properties of the SETAR predictors

Given the results in the previous section, it is interesting to investigate some properties of the predictors in (2) and (4). Following Box and Jenkins (1970), we have shown that the predictors in (2), for $\ell \leq d$, are unbiased while the biasedness of $\hat{X}_t(\ell)$ in (4) depends from the regimes coefficients, $\hat{X}_t(1)$ and the errors variance.

The variance of the predictors (2) and (4), differently from the linear domain, is not always a monotonic nondecreasing function. In particular, for $\ell \leq d$, and assuming for simplicity that $p_1 = p_2 = p$ in model (1) with $k = 2$, the variance of the predictor (2) is given as:

$$\begin{aligned}\ell = 1, \quad \hat{\sigma}^2(1) &= \sigma_i^2; \\ \ell = 2, \quad \hat{\sigma}^2(2) &= (\phi_1^{(i)})^2 \hat{\sigma}^2(1) + \sigma_i^2; \\ \ell = 3, \quad \hat{\sigma}^2(3) &= (\phi_1^{(i)})^2 \hat{\sigma}^2(2) + (\phi_2^{(i)})^2 \hat{\sigma}^2(1) + \sigma_i^2, \quad \text{and so on,}\end{aligned}$$

which are monotonic nondecreasing, with $i = 1$ if $X_{t+\ell-d} \leq r_1$ and $i = 2$ otherwise. When $\ell = d + 1$ the variance of the predictor (4) depends from the second order moments and from the values of the model coefficients. In particular, for the sake of simplicity, we consider the SETAR(2;1,1) model with $d = 1$, where the predictor variance $\hat{\sigma}^2(\ell)$, with $\ell = d + 1 = 2$, is given as:

$$\hat{\sigma}^2(2) = (\phi_1^{(1)} \cdot \lambda)^2 \cdot E[\gamma^2] + (\phi_1^{(2)} \cdot (1 - \lambda))^2 \cdot E[\delta^2] + 2\phi_1^{(1)} \cdot \phi_1^{(2)} \cdot \lambda(1 - \lambda) \cdot E[\gamma \cdot \delta] \quad (5)$$

where $\gamma = -\frac{\hat{\sigma}(1)}{\sqrt{2\pi}} \cdot \exp(-\frac{1}{2}(\frac{r_1 - \hat{X}_t(1)}{\hat{\sigma}(1)})^2)$ and $\delta = \frac{\hat{\sigma}(1)}{\sqrt{2\pi}} \cdot \exp(-\frac{1}{2}(\frac{r_1 - \hat{X}_t(1)}{\hat{\sigma}(1)})^2)$.

In (5) the inequality $\hat{\sigma}^2(\ell) \geq \hat{\sigma}^2(\ell - 1)$ is obtained under some restrictions on the moments and on the model parameters.

These results contrast the common belief that the further ahead we forecasts the less reliable is the generated prediction. Yao and Tong (1994) argue that the nonmonotonicity of the conditional variance for nonlinear models implies the possibility that the error of an $(\ell + 1)$ -step prediction can be smaller than the ℓ -step prediction from the same initial values, and this reflects that the prediction accuracy is influenced by whether we are currently. For example in a financial context, to be on a "peak" or in a "trough" of a business cycle impacts differently the accuracy of the forecasts.

4. An application to real data

The performance of the *naïve*, *monte carlo*, *closed form* (Granger and Teräsvirta, 1993) and the *normal forecasting errors* (NFE) (Moeanaddin, 1991) forecasts of the SETAR model are evaluated using the Nikkey 300 returns from January 1, 1996 to December 31, 2000. After fitting a SETAR(2;6,1) model with threshold delay $d = 3$ to the data, the multi step ex-post forecasts have been generated for the first 4 ($\ell \in [1, 4]$) working days of January 2001 using the listed methods.

The results given in table 1 show that the closed form forecasts outperform the other three methods in terms of RMSE and MAE with the exception of the NFE method which gives slightly better results in term of MAE even if, observing the generated values, the sign is better predicted from the closed form forecasts. Furthermore the bimodality of $X_{t+4|t}$ penalizes heavily the forecasts obtained from the monte carlo procedure which give the less accurate results.

Table 1: *RMSE and MAE of the first 4 steps ahead Nikkey 300 forecasts*

procedure	RMSE $\times 10^{-2}$	MAE $\times 10^{-2}$
<i>naïve</i>	0.3009	0.2586
<i>monte carlo</i>	1.3000	1.2341
<i>closed form</i>	0.2705	0.2403
<i>NFE</i>	0.2847	0.2370

The results obtained in the present paper and in Amendola and Niglio (2001) suggest a guidance in the choice of the SETAR predictor. In fact, from the properties given in section 3 and assumed the gaussianity of the errors in (1), for $\ell \leq d$ *naïve* forecasts are as accurate as the *closed form* forecasts but when $\ell = d + 1$ the latter predictor should be preferred.

5. Concluding remarks

The analytic distribution of the multi-step predictors of a SETAR model with gaussian errors has been derived. Using these results the closed form forecasts have been generated and their prediction performances have been empirically compared with *naïve*, *monte carlo* and *NFE* forecasts. The results show the good performances of the closed form forecasts with respect to the other methods both in term of prediction accuracy and of correct sign prediction. The latter result is not negligible in a financial context where the sign of the returns assumes a remarkable relevance for the evaluation of the market uncertainty.

References

- Amendola, A. and Niglio, M. (2001). Predictors distribution and forecast accuracy of threshold models, in C. Provasi (ed.), *Modelli complessi e metodi computazionali intensivi per la stima e la previsione*, CLEUP, pp. 247–252.
- Box, G. and Jenkins, G. (1970). *Time series analysis, forecasting and control*, Holden-Day, San Francisco.
- Clements, M. and Smith, J. (1997). The performances of alternative forecasting methods for setar models, *International Journal of Forecasting* **13**, 463–475.
- Granger, G. and Teräsvirta, T. (1993). *Modelling nonlinear relationships*, Oxford University Press, Oxford.
- Potter, S. (1999). Nonlinear time series modelling: an introduction, *Staff reports 87*, Federal Reserve Bank of New York.
- Tay, A. and Wallis, K. (2000). Density forecasting: a survey, *Journal of Forecasting* **19**, 235–254.
- Yao, Q. and Tong, H. (1994). Quantifying the influence of initial values on nonlinear prediction, *Journal of Royal Statistical Society (B)* **56**, 701–725.