ESTIMATING SETARs WITH MULTIVARIATE THRESHOLDS

Jan G. De Gooijer[1], and Gabriella Schoier[2]

[1] Department of Economic Statistics, University of Amsterdam
e-mail: j.g.degooijer@uva.nl

[2] Dipartimento di Scienze Economiche e Statistiche,
Università Degli Studi di Trieste, Italy
e-mail: gabriellas@econ.univ.trieste.it

Abstract

GRASP is a Greedy Randomised Adaptive Sampling Procedure that has been proposed to estimate parameters of self-exciting autoregressive threshold models (SETARs) with multivariate thresholds. We show that the GRASP procedure can often lead to an incorrect number of thresholds when estimating SETARs. Two simple modifications of the original GRASP procedure are suggested to overcome this problem.

Some key words: Combinatorial optimization; GRASP; Threshold time series models.

1 Introduction

Assume that a univariate and stationary time series $\{Y_t\}$ follows a so-called self-exciting threshold autoregressive (SETAR) process; see, e.g., Tong (1990). It is well-known that the dynamics of this process are controlled by a single threshold variable Y_{t-d} with $d > 0$ an integer value denoting the delay. Tiao and Tsay (1994) generalized the single threshold SETAR process to a model having two threshold variables, one for Y_{t-d} and another for $Y_{t-d} - Y_{t-d'}$ with $d \neq d'$. Thus, the relationship between Y_{t-d} and $Y_{t-d'}$ is linear and the parameters, here 1 and -1, are known. A more flexible SETAR formulation can be obtained by assuming the relationship between Y_{t-d} and $Y_{t-d'}$ still linear but the parameters are unknown. To put this into a general framework, consider an q-dimensional Euclidean space $\mathbb{R}^q$ and a point $\boldsymbol{x}$ in that space. Let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_q)'$ denote a q-dimensional unknown parameter vector. These parameters define a hyperplane as follows $\mathbb{H} = \{\boldsymbol{x} \in \mathbb{R}^q | \boldsymbol{\omega}'\boldsymbol{x} = \beta\}$, where β is a scalar parameter. The direction of $\boldsymbol{\omega}$ determines the orientation of the hyperplane whereas β represents the position of the hyperplane in terms of its distance from the origin. The hyperplane $\mathbb{H}$ induces a partition of the space into two regions defined by the half spaces $\mathbb{H}^- = \{\boldsymbol{x} \in \mathbb{R}^q | \boldsymbol{\omega}'\boldsymbol{x} \leq \beta\}$ and $\mathbb{H}^+ = \{\boldsymbol{x} \in \mathbb{R}^q | \boldsymbol{\omega}'\boldsymbol{x} > \beta\}$. In terms of the indicator function $I_{\boldsymbol{\omega},\beta}(\boldsymbol{x})$, the above partitioning can be written as $I_{\boldsymbol{\omega},\beta}(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in \mathbb{H}^-$, and zero otherwise.

Now, assume that an q-dimensional space is spanned by the vector of time series values $\tilde{\boldsymbol{Y}}_t = (Y_{t-1}, \ldots, Y_{t-q})'$. Further, suppose that there are ℓ functions $I_{\boldsymbol{\omega}_j,\beta_j}(\tilde{\boldsymbol{Y}}_t)$ $(j = 1, \ldots, \ell)$ where $\boldsymbol{\omega}_j = (\omega_{1j}, \ldots, \omega_{qj})'$ and β_j are real parameters. Thus, each of these functions defines a threshold. Then a SETAR model with $q > 1$ (multivariate) thresholds and order $(\ell; p, \ldots, p)$, denoted by SETAR$(\ell; p, \ldots, p)_q$, is defined as

$$\begin{aligned} Y_t &= \gamma_0 + \sum_{i=1}^{p} \gamma_i Y_{t-i} + \sum_{j=1}^{\ell} \Big\{ \lambda_0^{(j)} + \sum_{i=1}^{p} \lambda_{ij} Y_{t-i} \Big\} I_{\boldsymbol{\omega}_j,\beta_j}(\tilde{\boldsymbol{Y}}_t) + \varepsilon_t \\ &= \boldsymbol{\gamma}'\boldsymbol{Y}_t + \sum_{j=1}^{\ell} \boldsymbol{\lambda}_j'\boldsymbol{Y}_t I_{\boldsymbol{\omega}_j,\beta_j}(\tilde{\boldsymbol{Y}}_t) + \varepsilon_t \end{aligned} \tag{1}$$

where $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_p)'$, $\boldsymbol{\lambda}_j = (\lambda_{0j}, \ldots, \lambda_{pj})'$, $\boldsymbol{Y}_t = (1, Y_{t-1}, \ldots, Y_{t-p})'$, and where $\{\varepsilon_t\}$ is a white noise process with zero mean and finite variance σ^2. Note that (1) is not identified. For identification purpose the restrictions $\beta_1 \leq \ldots \leq \beta_\ell$ is imposed. Further, due to the fact that $I_{\boldsymbol{\omega}_j,\beta_j}(\cdot) = 1 - I_{-\boldsymbol{\omega}_j,-\beta_j}(\cdot)$, a convenient normalization condition is to set one element of $\boldsymbol{\omega}_j$ equal to unity, e.g. $\omega_{1j} = 1$. If needed, (1) can be further generalized by allowing for interactions between the ℓ functions $I_{\boldsymbol{\omega}_j,\beta_j}(\tilde{\boldsymbol{Y}}_t)$.

Estimation of (1) can be done by conditional least squares (CLS) or maximum likelihood once the parameters $\boldsymbol{\omega}_j$ and β_j have been determined. Often, however, $\boldsymbol{\omega}_j$ and β_j are unknown. Then the total number of possible hyperplanes to search is $n!/q!(n-q)!$, where n denotes the number observations on $\{Y_t\}$. Of course, for most practical problems a search over all possible combinations of hyperplanes is infeasible. To solve this problem, Medeiros, Veiga and Resende (2002) proposed a procedure based on a greedy randomized adaptive search procedure (GRASP). GRASP can be used to solve optimization problems which have a high number, but not infinite, of possible solutions; see Section 2 for some details.

In this paper, the potential of the GRASP algorithm in estimating SETARs with multivariate thresholds is investigated. We show that the algorithm often results in an incorrect number of thresholds when the time series under study are of moderate length. To overcome this problem, we propose two modifications of the GRASP algorithm. These modifications result in an improvement in model fit without increasing the computational burden.

This paper is organized as follows. In Section 2, we briefly review GRASP and its adaptation for estimating SETARs with multivariate thresholds. This section also contains two modifications of the basic GRASP algorithm. Section 3 shows results of a comparative simulation study. Section 4 concludes.

2 GRASP for SETARs

2.1 What is GRASP

GRASP is a multi-start iterative randomized sampling technique that can be used to quickly produce good quality solutions for a wide variety of optimization problems. The algorithm was originally proposed by Feo and Resende (1995) and Resende (1999). GRASP does not try all possible solutions and then finds the best one for the problem at hand. This approach would be far from good computing practice. In particular, when the number of possible solutions is high due to the fact that a large data set is analysed there is the risk that the processor cannot manage. In contrast, the GRASP algorithm is based on the idea that it is better to have a faster solution than one slower even if the solution is slightly bad. This basic characteristic makes the algorithm particularly useful for analysing various real life problems. Examples include designing efficient telecommunication networks, scheduling operations in a semiconductor manufacturing plant, locating strategic energy reserves, routing delivery vehicles, airline crew scheduling, and designing large experiments.

procedure (GRASP)
N = maximum number of grasp iterations
for $i = 1$ to N
procedure (MAKE RANDOM SUBSET)
procedure (CONSTRUCTION PHASE)
procedure (LOCAL SEARCH PHASE)
next i

Figure 1. Generic GRASP

Figure 1 shows, as a pseudocode, the generic GRASP algorithm used by Medeiros *et al.* (2002) for the estimation of (1). The procedure "make random subset" is only a preliminary one which selects the elements which could enter the solution set randomly; it is also used to reduce the computational burden of the algorithm. In the construction phase a feasible solution is iteratively constructed, one element at a time; see Figure 2 for the pseudo-code. The restricted candidate list (RCL) makes a partial screening of the elements which could enter the solution set. Thus, the GRASP algorithm does not choose an element randomly. Rather GRASP tries to select a solution near the best one. All candidate elements which could be in the solution set are classified with a greedy function. In the case of SETARs Medeiros *et al.* (2002) use the mean square error (MSE) as a greedy function to calculate the hyperplane or hyperplanes which constitute the solution set. Those elements which have a value of the greedy function below a prespecified threshold are selected as elements of the RCL set.

procedure (CONSTRUCTION PHASE)
do
Create an RCL set.
Add one RCL element to the solution.
loop until the solution is completed.

Figure 2. Construction phase procedure

procedure (LOCAL SEARCH PHASE)
do
Select one solution in the neighbourhood.
If this solution has a larger SBIC value than the previous one reject it, else take this solution as the optimal one.
loop until the neighbourhood is fully analysed.

Figure 3. Local search phase procedure

In the local search phase a solution is replaced by a "better" one in the neighbourhood of the solution found in the construction phase; see Figure 3 for a possible implementation. For SETARs Medeiros *et al.* (2002) recommend the use of Schwarz' BIC which in this case is defined as

$$SBIC(\ell) = \ln(\hat{\sigma}^2) + n^{-1}\ln(n)\Big\{\ell(p+q+1)+p\Big\},$$

where $\hat{\sigma}^2$ is an estimate of σ^2, and n is the sample size. The GRASP procedure for SETARs is initialized with CLS estimates for the unknown parameters. At each iteration step, the GRASP procedure computes thresholds. In the next step it uses again CLS to estimate the parameters of the specified model.

2.2 Two modifications of GRASP

Medeiros *et al.* (2002, Section 6) showed in a Monte Carlo study that the basic GRASP algorithm for SETARs, using SBIC in the local search phase, has a strong tendency for selecting an incorrect number of thresholds when $n \leq 100$. This is confirmed by the results of the simulation study presented in the next Section. One modification to overcome this problem is to recalculate SBIC backward for every possible combination of hyperplanes estimated previously; see Figure 4. To illustrate this idea, consider the case with three possible hyperplanes. In the first step the modified GRASP algorithm calculates the value of SBIC with a SETAR model having only one hyperplane. Next, this is done with the second hyperplane. Then the value of SBIC is computed for the third, the first and the second, the first and the third, and finally all three hyperplanes. In this way it is unlikely that any threshold can be removed without seriously degrading the fit of the SETAR model. The resulting GRASP algorithm will be denoted by "basic+SBIC".

procedure (SBIC RECALCULATION)
Calculates the 2^n possible hyperplanes combinations.
for $i = 1$ to 2^n
Calculate SBIC for this combination.
next i

Figure 4. SBIC recalculation procedure

A second modification concerns the two-exchange local search, where a hyperplane $\mathbb{H}$ in the solution set is replaced by another hyperplane that is not in the solution set. To this end, the local search is divided two main blocks. The first block rotates each hyperplane in the solution set whereas the second one translates each hyperplane. For each hyperplane $\mathbb{H}$, the second block implies replacing β_j by the elements of the projection $\boldsymbol{\omega}_j'\boldsymbol{x}_j$ $(j = 1, \ldots, \ell)$. Clearly, the weakness of this approach lies in the fact that the search for optimal values of $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}_j$ is done separately rather than jointly. In fact, by proceeding as described above, it is unlikely that $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}_j$ are optimized. On the other hand, a procedure that numerically optimizes these parameters jointly is hard to implement and is computationally time consuming.

A simple procedure to improve the estimation of $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}_j$ is to start from the previous solution, say $\boldsymbol{\omega}_j^*$ and $\boldsymbol{\beta}_j^*$ with $\boldsymbol{\omega}_j^* = (\omega_{1j}^*, \ldots, \omega_{qj}^*)'$. Then the idea is to allow for a little change, by increasing and decreasing the values of $\omega_{2j}^*/\omega_{1j}^*, \ldots, \omega_{qj}^*/\omega_{1j}^*$ with some small number. Next, for each variation in the slope, $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}_j$ are re-estimated. The best combination of parameters is chosen as the one which has the smallest MSE. The resulting modified GRASP algorithm will be denoted by "basic+SBIC+slope". Both modifications of GRASP are shown, in pseudo-code, in Figure 5.

procedure (GRASP)
N = maximum number of grasp iterations
for $i = 1$ to N
procedure (MAKE RANDOM SUBSET)
procedure (CONSTRUCTION PHASE)
procedure (LOCAL SEARCH PHASE)
next i
procedure (SBIC RECALCULATION)
procedure (SLOPE RECALCULATION)

Figure 5. Generic GRASP modified

3 Simulated Examples

Example 1. Consider the SETAR model

$$Y_t = \begin{cases} -0.95Y_{t-1} + \varepsilon_t, & \text{if} \quad |Y_{t-1} - Y_{t-2}| \geq 0.7 \\ 0.95Y_{t-1} + \varepsilon_t, & \text{otherwise,} \end{cases} \tag{2}$$

where $\varepsilon_t \sim NID(0, \sigma^2)$. Model (2) is a SETAR$(2; 1, 1)_2$ model. It can be rewritten as follows

$$Y_t \;\; = \;\; 0.95Y_{t-1} - 1.9Y_{t-1}I_{\boldsymbol{\omega}_1,\beta_1}(\tilde{\boldsymbol{Y}}_t) + 1.9Y_{t-1}I_{\boldsymbol{\omega}_2,\beta_2}(\tilde{\boldsymbol{Y}}_t) + \varepsilon_t. \tag{3}$$

It is easy to see that, in terms of model (1), we have $\boldsymbol{\gamma} = (0,\ 0.95)'$, $\boldsymbol{\lambda}_1 = -\boldsymbol{\lambda}_2 = (0,\ 1.9)'$, $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2 = (1,\ -1)'$, $\tilde{\boldsymbol{Y}}_t = (Y_{t-1}, Y_{t-2})'$, and $\beta_1 = -\beta_2 = -0.7$.

Figure 6.a) shows $n = 500$ simulated observations from model (3) with $\varepsilon_t \sim NID(0, (0.25)^2)$. To avoid start-up problems 500 observations were deleted from the simulated series. Note that the dynamics of model (3) is controlled by two bidimensional thresholds. The first one when $Y_{t-1} - Y_{t-2} = -0.7$ and the second one when $Y_{t-1} - Y_{t-2} = 0.7$. The number of thresholds is 2, and the total number of hyperplanes is 124750. Using $q = 2$, Table 1 shows estimated values of $\boldsymbol{\omega}_j'$ and β_j $(j = 1, 2)$ for the best fitted model found for respectively $p = 1$, 2, and 3. The SBIC

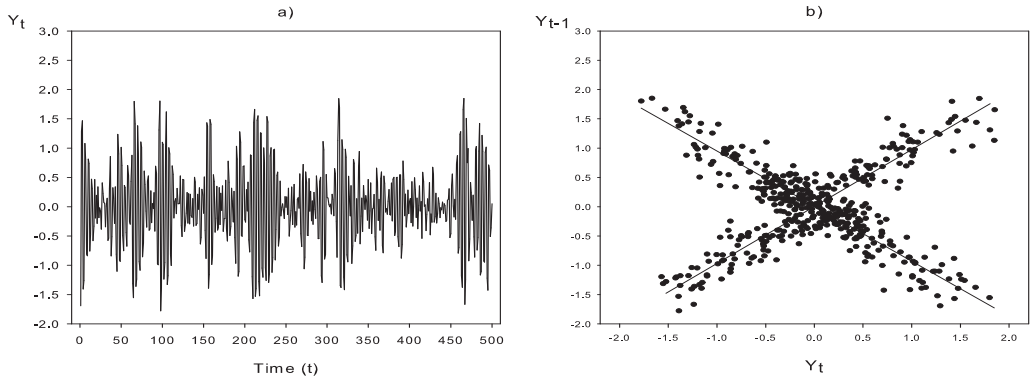

Figure 6: a) Time series generated by (3); b) Scatter plot of Y_t versus Y_{t-1} and estimated (solid lines) separating hyperplanes.

attains its lowest value for the estimated SETAR$(2;1,1)_2$ model

$$\begin{aligned} Y_t &= 0.009 + 0.956Y_{t-1} + (-0.082 + 1.937Y_{t-1})I_{\boldsymbol{\omega}_1,\beta_1}(\tilde{\boldsymbol{Y}}_t) \\ &\quad + (0.005 - 1.894Y_{t-1})I_{\boldsymbol{\omega}_2,\beta_2}(\tilde{\boldsymbol{Y}}_t) + \hat{\varepsilon}_t. \end{aligned}$$

The parameter values of γ_0, $\lambda_0^{(1)}$, and $\lambda_0^{(2)}$ are not statistically different from zero at the 5% level. Figure 6.b) shows the scatter plot of Y_t versus Y_{t-1} and the estimated hyperplanes. We see that the basic GRASP procedure has correctly identified the position of the separating hyperplanes.

Table 1: Estimated parameters $\boldsymbol{\omega}_j'$ and β_j $(j = 1, 2)$ for three SETAR$(2;p,p)_2$ models fitted to a simulated series of length $n = 500$, and the corresponding SBIC values.

p	$\hat{\boldsymbol{\omega}}_1'$	$\hat{\boldsymbol{\omega}}_2'$	$\hat{\beta}_1$	$\hat{\beta}_2$	SBIC
1	(1.000, -0.993)	(1.000, -1.013)	0.657	-0.701	-2.575
2	(1.000, -1.006)	(1.000, -1.013)	0.660	-0.701	-2.524
3	(1.000, -0.967)	(1.000, -1.192)	0.746	-0.705	-2.556

To investigate whether the above results for (3) still hold for smaller sample sizes, we generated series of length $n = 100$. Table 2 shows the mean and standard deviations of the estimated model parameters using the basic, basic+SBIC, and basic+SBIC+slope algorithms. For the basic algorithm the estimation results for $\sigma = 0.5$ ($\sigma = 1$) are based on 88 (97) replications. For the basic+SBIC and basic+SBIC+slope algorithms the number of replications is 89 (102) for $\sigma = 0.5$ ($\sigma = 1$). Note that in terms of bias and accuracy, as measured by respectively the mean

Table 2: Mean and standard deviations (in parentheses) of the parameter estimates of model (2) with $\sigma = 0.5$ and 1 using three GRASP algorithms; $n = 100$.

True parameters	$\sigma = 0.5$			$\sigma = 1$		
	Basic	Basic+SBIC	Basic+SBIC + slope	Basic	Basic+SBIC	Basic+SBIC + slope
$\gamma_0 = 0$	-0.028 (0.214)	-0.030 (0.214)	-0.031 (0.210)	0.124 (1.742)	0.111 (1.705)	-0.097 (0.339)
$\gamma_1 = 0.95$	0.953 (0.126)	0.953 (0.126)	0.954 (0.123)	0.943 (0.258)	0.939 (0.255)	0.892 (0.119)
$\beta_1 = 0.7$	0.712 (0.116)	0.714 (0.117)	0.708 (0.103)	0.749 (0.206)	0.752 (0.206)	0.688 (0.249)
$\beta_2 = -0.7$	-0.752 (0.257)	-0.751 (0.256)	-0.749 (0.256)	-0.720 (0.154)	-0.721 (0.153)	-0.717 (0.162)
$\lambda_{01} = 0$	0.002 (0.236)	0.005 (0.237)	0.011 (0.229)	-0.082 (1.682)	-0.083 (1.646)	0.180 (0.417)
$\lambda_{11} = 1.9$	1.9243 (0.282)	1.923 (0.281)	1.930 (0.272)	1.832 (0.485)	1.848 (0.509)	1.690 (0.658)
$\lambda_{02} = 0$	0.088 (0.245)	0.088 (0.244)	0.083 (0.251)	-0.009 (0.477)	-0.004 (0.468)	-0.029 (0.500)
$\lambda_{12} = -1.9$	-1.837 (0.338)	-1.835 (0.336)	-1.847 (0.330)	1.750 (0.486)	1.766 (0.504)	1.676 (0.669)
$\omega_{12} = -1$	-1.007 (0.072)	-1.007 (0.072)	-1.004 (0.069)	-1.001 (0.141)	-0.999 (0.138)	-1.013 (0.136)
$\omega_{22} = -1$	-0.984 (0.220)	-0.984 (0.218)	-0.982 (0.219)	-0.989 (0.071)	-0.988 (0.070)	-0.998 (0.073)

and the standard deviation of the estimated parameters, the basic+SBIC algorithm improves over the basic GRASP algorithm. For larger values of n we observed that both algorithms give more or less the same estimation results. We also see from Table 2 that the improvement of the basic+SBIC+slope algorithm over the previous two algorithms is limited.

Table 3 gives another view on the performance of the basic– and basic+SBIC algorithms for $\sigma = 0.25$, 0.5 and 1, with $n = 100$. Here the number of thresholds selected by these two algorithms are reported. Since in all cases the number of thresholds selected by the basic+SBIC+slope algorithm were identical to those obtained by the basic+SBIC algorithm we only included the latter results. Clearly, for increasing values of σ the basic+SBIC algorithm does a much better job in identifying the correct number of thresholds than the basic algorithm.

Example 2. Consider the SETAR$(2; 2, 2)_2$ process

$$\begin{aligned} Y_t &= 1.5 + 0.5Y_{t-1} - 0.8Y_{t-2} + (4 - 0.4Y_{t-1} + 0.2Y_{t-2})I_{\boldsymbol{\omega}_1,\beta_1}(\tilde{\boldsymbol{Y}}_t) \\ &\quad + (3 + 0.5Y_{t-1} - 0.2Y_{t-2})I_{\boldsymbol{\omega}_2,\beta_2}(\tilde{\boldsymbol{Y}}_t) + \varepsilon_t \end{aligned} \tag{4}$$

Table 3: Thresholds selected by two GRASP algorithms for $\sigma = 0.25$, 0.5 and 1; $n = 100$, 112 replications.

Thresholds	$\sigma = 0.25$		$\sigma = 0.5$		$\sigma = 1$	
	Basic	Basic+SBIC	Basic	Basic+SBIC	Basic	Basic+SBIC
0	20	20	11	11	3	3
1	16	16	6	6	2	2
2	40	40	88	89	97	102
3	36	36	7	6	5	3
4	0	0	0	0	5	2

with parameters $\boldsymbol{\gamma} = (1.5,\ 0.5,\ -0.8)'$, $\boldsymbol{\lambda}_1 = (4,\ -0.4,\ 0.2)', \boldsymbol{\lambda}_2 = (3,\ 0.5,\ -0.2)'$, $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2 = (-1,\ 0)'$, $\tilde{\boldsymbol{Y}}_t = (Y_{t-1}, Y_{t-2})'$, and $\beta_1 = 0$, $\beta_2 = -0.7$. In other words, the dynamics of model (4) is controlled by two bidimensional thresholds. Table 4 shows the mean and standard deviations of the estimated model parameters using the basic, basic+SBIC, and basic+SBIC+slope algorithms. For the basic algorithm the estimation results for $\sigma = 0.5$ ($\sigma = 1$) are based on 56 (89) replications. For the basic+SBIC and the basic+SBIC+slope algorithms the number of replications is 59 (91) for $\sigma = 0.5$ ($\sigma = 1$). Note that in terms of bias and accuracy, as measured by respectively the mean and the standard deviation of the estimated parameters, the basic+SBIC algorithm improves over the basic algorithm. For larger values of n we observed that both algorithms give more or less the same estimation results. We also see from Table 4 that the improvement of the basic+SBIC+slope algorithm over the previous two algorithms is limited. Table 5 shows the performance of the basic– and basic+SBIC algorithms for $\sigma = 0.25$, 0.5 and 1, with $n = 100$. The results are similar to those reported in Table 3, i.e. for increasing values of σ the basic+SBIC algorithm outperforms the basic algorithm. In terms of computational effort, the basic+SBIC+slope algorithm takes about 1.5 more time than the other two algorithms. In contrast, the increase in accuracy of the basic+SBIC algorithm over the basic algorithm does not require extra computational time.

4 Conclusions

The basic+SBIC GRASP algorithm, as a modification of the basic GRASP algorithm, is recommended for estimating SETARs with multivariate thresholds. In particular, the modified GRASP algorithm, is useful when the sample size is relatively small.

Table 4: Mean and standard deviations (in parentheses) of the parameter estimates of model (4) with $\sigma = 0.5$ and 1 using three GRASP algorithms; $n = 100$.

True parameter	$\sigma = 0.5$			$\sigma = 1$		
	Basic	Basic+SBIC	Basic+SBIC + slope	Basic	Basic+SBIC	Basic+SBIC + slope
$\gamma_0 = 1.5$	1.760 (0.745)	1.623 (0.714)	1.543 (0.585)	1.448 (0.793)	1.452 (0.798)	1.528 (0.616)
$\gamma_1 = 0.5$	0.553 (0.132)	0.528 (0.128)	0.513 (0.104)	0.480 (0.131)	0.487 (0.130)	0.496 (0.108)
$\gamma_2 = -0.8$	-0.814 (0.066)	-0.801 (0.064)	-0.800 (0.054)	-0.805 (0.074)	-0.803 (0.072)	-0.809 (0.048)
$\beta_1 = 0$	1.070 (1.405)	0.782 (1.504)	0.727 (1.438)	0.589 (1.237)	0.309 (1.222)	0.045 (0.951)
$\beta_2 = -0.7$	-0.531 (0.104)	-0.614 (0.360)	-0.614 (0.360)	-0.767 (0.502)	-0.745 (0.500)	-0.708 (0.5666)
$\lambda_{01} = 4$	3.811 (0.663)	3.916 (0.676)	3.982 (0.584)	4.022 (0.563)	4.008 (0.523)	3.860 (0.566)
$\lambda_{11} = -0.4$	-0.359 (0.044)	-0.367 (0.062)	-0.366 (0.063)	-0.435 (0.104)	-0.434 (0.103)	-0.431 (0.113)
$\lambda_{12} = 0.2$	0.113 (0.145)	0.141 (0.161)	0.155 (0.145)	0.2512 (0.139)	0.241 (0.132)	0.230 (0.119)
$\lambda_{02} = 3$	2.898 (0.330)	2.983 (0.291)	2.998 (0.258)	3.107 (0.358)	3.105 (0.350)	3.122 (0.361)
$\lambda_{21} = 0.5$	0.485 (0.030)	0.486 (0.039)	0.480 (0.039)	0.522 (0.105)	0.501 (0.102)	0.517 (0.087)
$\lambda_{22} = -0.2$	-0.182 (0.029)	-0.191 (0.037)	-0.191 (0.037)	-0.208 (0.066)	-0.207 (0.065)	-0.208 (0.060)
$\omega_{12} = -1$	-0.895 (0.139)	-0.929 (0.142)	-0.928 (0.144)	-0.961 (0.122)	-0.960 (0.121)	-1.011 (0.097)
$\omega_{22} = 0$	0.016 (0.028)	0.032 (0.053)	0.032 (0.053)	0.038 (0.100)	0.038 (0.099)	0.003 (0.112)

Table 5: Thresholds selected by two GRASP algorithms for $\sigma = 0.25$, 0.5 and 1, and with $n = 100$; 112 replications.

Thresholds	$\sigma = 0.25$		$\sigma = 0.5$		$\sigma = 1$	
	Basic	Basic+SBIC	Basic	Basic+SBIC	Basic	Basic+SBIC
0	0	0	0	0	0	0
1	2	2	17	17	18	18
2	40	50	56	59	89	91
3	68	59	39	36	5	3
4	2	1	0	0	0	0

Acknowledgement: We are grateful to Marcelo Medeiros for allowing us to use the MATLAB code for estimating SETARs with multivariate thresholds using GRASP. Also we would like to thank Tommaso Poli for helping us to modify the GRASP algorithm.

References

Feo, T.A. and Resende, M.G.C. (1989), "A probabilistic heuristic for a computationally difficult set covering problem", *Operations Research Letters*, 8, 67-71.

Feo, T.A. and Resende, M.G.C. (1995), "Greedy randomized adaptive search procedures", *Journal of Global Optimization*, 6, 109-133.

Medeiros, M.C., Veiga, A. and Resende, M.G.C. (2002), "A combinatorial approach to piecewise linear time series analysis", *Journal of Computational and Graphical Statistics*, 11, 236-258.

Resende, M.G.C. (1999), "Greedy randomized adaptive search procedures (GRASP)". In *Encyclopedia of Optimization* (Floudas, C. and Paradols, P., eds.), Kluwer Academic Publishers.

Tiao, G.C. and Tsay, R.S. (1994), "Some advances in non-linear and adaptive modelling in time series analysis", *Journal of Forecasting*, 13, 109-131.

Tong, H. (1990), *Non-Linear Time Series: A Dynamical System Approach.* Oxford: Oxford University Press.