

Winker, Peter

**Working Paper**

## Identification of multivariate AR-models by threshold accepting

Diskussionsbeiträge - Serie II, No. 224

**Provided in Cooperation with:**

Department of Economics, University of Konstanz

*Suggested Citation:* Winker, Peter (1994) : Identification of multivariate AR-models by threshold accepting, Diskussionsbeiträge - Serie II, No. 224, Universität Konstanz, Sonderforschungsbereich 178 - Internationalisierung der Wirtschaft, Konstanz

This Version is available at:

<https://hdl.handle.net/10419/101463>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

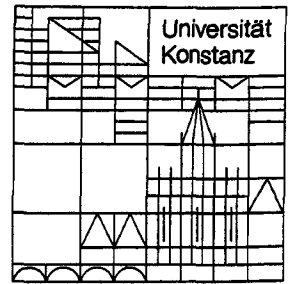
*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**Sonderforschungsbereich 178**  
**„Internationalisierung der Wirtschaft“**

Diskussionsbeiträge



Juristische  
Fakultät

Fakultät für Wirtschafts-  
wissenschaften und Statistik

---

Peter Winker

**Identification of Multivariate AR-Models  
by Threshold Accepting**

**Identification of Multivariate AR-Models  
by Threshold Accepting\***

**Peter Winker**

Serie II - Nr. 224

Mai 1994

Peter Winker  
Sonderforschungsbereich 178  
"Internationalisierung der Wirtschaft"  
Universität Konstanz  
Postfach 55 60

D-78434 Konstanz

e-mail: winker @ sonne.wiwi.uni-konstanz.de

- \* *Research was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 178 "Internationalisierung der Wirtschaft, University of Konstanz. I am indebted to B. Fitzenberger and W. Smolny for valuable comments on a first draft of this paper.*

### **Abstract**

In econometric modelling the choice of relevant variables is of crucial importance for the interpretation of the results. In many cases it is based on some a priori knowledge from economic theory and a rather heuristic procedure for determining other influential variables sometimes based on an information criterion.

This paper deals with an automatic method for the identification of relevant variables based solely on an information criterion. As an example, the identification of multivariate lag structures in AR-models is studied. This issue arises e.g. for large-scale econometric models, for Granger causality tests or the application of Johansen's test for cointegration. The procedure suggested in this paper allows the optimization of the lag structure over the whole set of possible multivariate lag structures with regard to a given information criterion, e.g. the Hannan-Quinn estimator or Akaike's final prediction error criterion.

The optimization is performed by the heuristic multiple purpose optimization algorithm Threshold Accepting which proved to be very successful for discrete optimization problems in economics and econometrics. The implementation of Threshold Accepting for subset identification in multivariate AR-models and some simulation results for a bivariate model are presented.

Keywords: AR-model; Identification; Optimization;  
Heuristics; Threshold Accepting

# 1 Introduction

In econometric modelling the choice of relevant variables is of crucial importance for the interpretation of the results. In many cases it is based on some a priori knowledge from economic theory and a rather heuristic procedure for determining other influential variables. Sometimes the number of variables and especially the number of lagged observations depends on some information criterion. Nevertheless, any modelling of this kind might be subject to the reproach of ad hoc data analysis.

This paper deals with an automatic method for the identification of relevant variables based solely on an information criterion. As an important example, the identification of multivariate lag structures in VAR-models will be studied. For large-scale econometric models as well as for Granger causality tests often quite complex lag structures in multivariate regressions must be estimated (see for example Pierce and Haugh (1976), p. 289, or Sims (1972)). In principle, Johansen's test procedure for cointegration (Johansen (1991)) also requires the identification of the lag structure in the underlying VAR-model. The procedure introduced in this paper allows the optimization of the lag structure over the whole set of possible multivariate lag structures with regard to a given information function such as the Hannan-Quinn estimator (HQ) or Akaike's final prediction error criterion (FPE).

McClave (1975) proposed an algorithm for the identification of univariate AR-processes. It is based on a procedure by Pagano (1972) and Hocking and Leslie (1967) and Akaike's final prediction error criterion. In the example given by McClave (1975) this algorithm performed well, although the FPE criterion was proven to be inconsistent (cf. Shibata (1976)). In some simulations however, the performance of this algorithm was less satisfying. In particular, the algorithm checked almost half of the possible subsets before identifying the best one.

Unfortunately, already for a bivariate VAR-model with a maximum lag length of 20, there are more than  $10^{12}$  possible subsets of the lagged endogenous variables. Comparing all these possibilities would take lots of time even with the fastest computers. In cases with even more equations or lags to test for, the problem becomes intractable.

There are different ways out of this seemingly dead end. First, some a priori knowledge on the possible structure of the model might be used. However, a priori assumptions – especially if they cannot be based on economic theory – might be misleading, and even with some a priori restrictions the

test of the remaining combinations will still be intractable. Another way out for the lag structure of VAR-models is to restrict oneself to models which include all lagged endogenous variables up to a given lag. Of course, then only 20 models have to be compared for the example given above, but the chance of finding the best model just within these 20 models is extremely small.

In this paper, another strategy is proposed for the identification of the lag structure of VAR-models which could also be easily adapted to similar problems. It is based on the use of a discrete optimization heuristic. Over the last decade, many powerful multiple purpose optimization heuristics have been developed and applied for optimizing telephone nets, chip layouts, job shop schedules, bond portfolios, etc. (see for example Kirkpatrick et al. (1983); Dueck and Scheuer (1990); Dueck and Winker (1992)). "Heuristic" means that these algorithms do not give the global optimum with probability one, but solutions sufficiently near to the optimal value for practical purposes. The basic advantage of heuristics is their velocity which makes it possible to find approximative solutions even for problems of a very high degree of complexity, when deterministic algorithms cannot give any solution in reasonable computing time. In particular, for the example of lag structures a heuristic has to test only a very small part of the feasible subsets in order to identify an excellent approximation to the underlying model structure.

A famous heuristic is the classical Simulated Annealing approach (see Kirkpatrick et al. (1983); Aarts and Korst (1989)). In Dueck and Scheuer (1990), Dueck and Wirsching (1991), and Dueck and Winker (1992) an even more efficient form, the Threshold Accepting algorithm (TA) has been introduced. Althöfer and Koschnick (1991) proved that TA has the same asymptotic convergence properties as Simulated Annealing. TA is able to minimize almost every objective function out of almost every set of admissible solutions under almost arbitrary constraints.

Successful implementations of the TA algorithm in economics and econometrics exist for the Travelling Salesman problem (Dueck and Scheuer (1990); Winker (1994)) and the problem of Optimal Aggregation (Chipman and Winker (1992)). This paper adds an implementation for the problem of identifying the multivariate lag structure in VAR-models.

One problem for the identification of VAR-models is behind the scope of this paper. It concerns the question of how to choose the "loss function" in order to get the best estimate of the "true" model. After Akaike's final prediction error criterion (Akaike (1969, 1971, 1973)) had been proven

to be inconsistent, asymptotically consistent estimators were developed by G. Schwarz (1978) and E.J. Hannan and B.G. Quinn (1979). A discussion of the asymptotic properties of different estimators can be found in Odaki (1987). Unfortunately, the small sample properties of these estimators are not well known. Nevertheless, some quite promising computational results based on the Hannan–Quinn criterion (HQ) and a heuristic adaptation of Akaike’s final prediction error criterion with increasing weights (AFPE) can be reported.

The paper is organized as follows. The next section briefly describes the problem of identification in VAR-models. In section 3 some possible choices for the information criterion (loss function) are analyzed. Section 4 introduces the optimization heuristic Threshold Accepting and its implementation for identifying multivariate AR-processes. Section 5 presents some simulation results for a bivariate AR-model. In a conclusion the main results of the paper are presented and areas of further research are identified.

## 2 Identifying a Multivariate AR-Process

For the purpose of this paper, a bivariate stochastic processes is considered as an example of multivariate AR-processes. However, it should be stressed that the method presented here is not constrained to bivariate models. A bivariate AR-proces  $\{x_t, y_t\}$  is given by

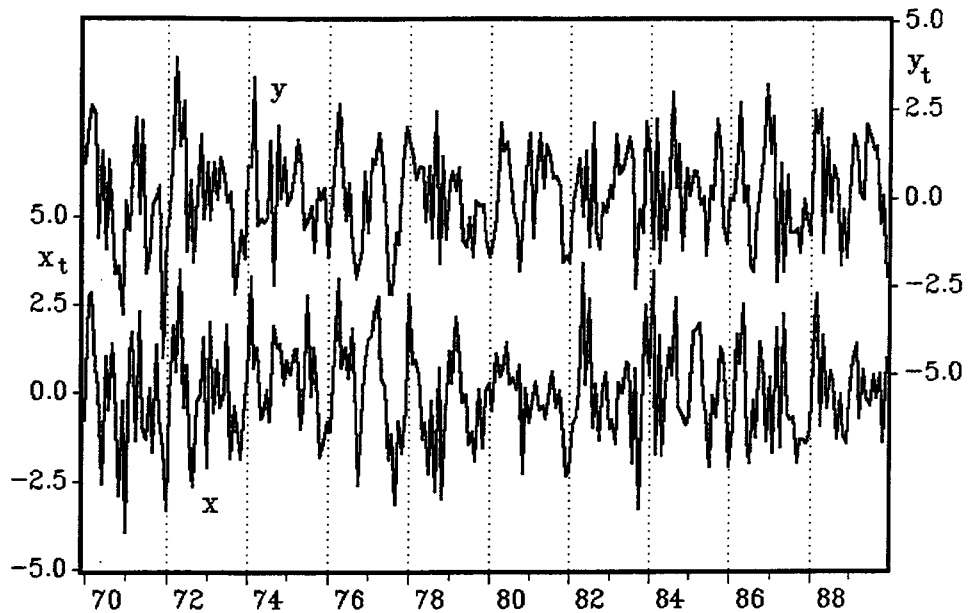
$$\begin{aligned} x_t &= \sum_{i=1}^K \alpha_i x_{t-i} + \sum_{i=1}^K \beta_i y_{t-i} + u_t \\ y_t &= \sum_{i=1}^K \gamma_i x_{t-i} + \sum_{i=1}^K \delta_i y_{t-i} + v_t, \end{aligned} \tag{1}$$

where  $u_t$  and  $v_t$  are *iid* as  $\mathcal{N}(0,1)$ .

For a given maximal lag length  $K' \geq K$  and a given observable realization, the problem of identification of the lag structures is to identify those coefficients  $\alpha_i, \beta_i, \gamma_i$  and  $\delta_i$  which are different from zero. Figure 1 shows the realization of a bivariate stochastic AR-process as defined in (1). The data were obtained starting the process with constant values for  $x_1 \dots x_{20}$  and  $y_1 \dots y_{20}$  using the normal random number generator of GAUSS and the coefficients given in (2) in section 5. The plot contains the realizations 241 to 480 for  $x_t$  and  $y_t$ .



Figure 1: Realization of a Bivariate AR-Process



When estimating such an observable realization none of the estimated coefficients will be zero. Therefore, the best estimation in terms of the sum of squared residuals will be that including all lagged values of  $x$  and  $y$  up to  $K$ . However, if only some coefficients of the “true” model are different from zero, this will not constitute a satisfactory identification of the underlying bivariate AR-process. In order to identify the model, an information criterion (loss function) has to be considered, i.e. a function which explicitly takes care of the number of estimated coefficients. The following section will briefly discuss a selection of possible loss functions.

### 3 The Loss Function

Possibly still the best known and most prominent rule for determining the dimension of an autoregressive process is the final prediction error criterion of Akaike (1969). It was first introduced as an ad hoc specification in order to explicitly take care of the number of included variables. Akaike (1973)

provides further arguments for the use of this loss function based on an information theoretic approach. He proposes to minimize the function

$$\text{FPE}_2 := |\Sigma_k| \left(1 + \frac{2k}{N}\right),$$

where  $k$  is the number of included lags,  $|\Sigma_k|$  the determinant of the sample variance-covariance matrix of the corresponding residuals, and  $N$  the number of observations.

Applications based on this loss function proved to be quite successful (Akaike (1973)). Pagano (1972) and McClave (1975) provide an application on subset autoregression. Although the theoretical introduction in Akaike (1973) is rather convincing and the applications gave useful results, the choice of the constant 2 in the criterion has no theoretical justification. It can be shown (Shibata (1976)) that identification based on the  $\text{FPE}_2$  overestimates the number of included lags.

Schwarz (1978) and Hannan and Quinn (1979) introduce alternative loss functions which can be shown to give consistent estimators of the dimension of the model. For these loss functions, the constant 2 is replaced by  $\ln N/2$  and  $2c \ln \ln N$  with  $c > 1$ , respectively. Unfortunately, not much is known about the small sample performance of these estimators.

The Hannan-Quinn criterion is given by:

$$\ln |\Sigma_k| + \frac{2kc}{N} \ln \ln N,$$

where  $|\Sigma_k|$  is the variance of the residuals in the univariate case and the determinant of the variance-covariance matrix in the multivariate case,  $k$  the number of included lags,  $N$  the number of observations and  $c$  a constant greater than 1.

For the Threshold Accepting implementation presented in this paper, the Hannan-Quinn estimator (HQ) with different values for the constant  $c$  and an adaptation of Akaike's final prediction error criterion (AFPE) have been chosen. For this adaptation, the constant in FPE increases during the execution of the optimization algorithm described in the following section. The constant increases from 2 in integer steps. A sharp increase in the corresponding values of  $\text{FPE}_2$  indicates that the number of estimated coefficients became too small. It should be noted that this method constitutes an ad hoc approximation to the asymptotically consistent estimators depending on the number of observations  $N$ . Nevertheless, the results presented in section 5 show that with both, the HQ estimator and the AFPE, a good approximation of the "true" model based on the observed data is possible.

## 4 Threshold Accepting

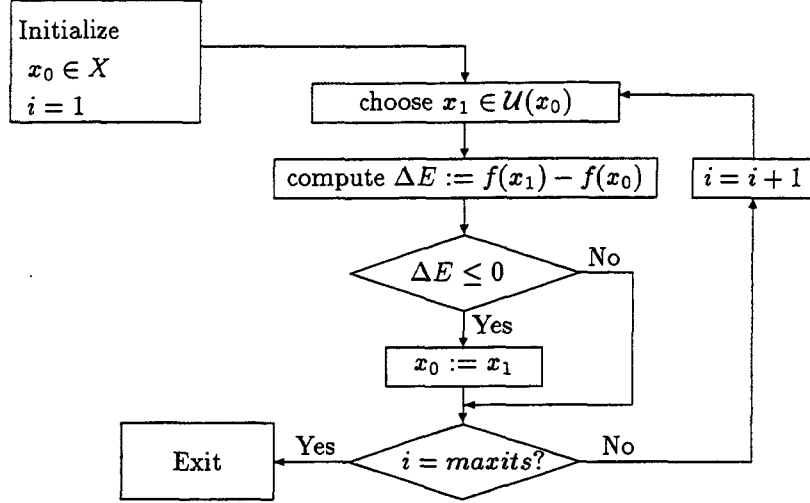
Given an observable realization of a multivariate AR-process as shown in figure 1, how can the underlying lag structure be identified using one of the criteria introduced in the previous section? A trivial method would consist of calculating the loss function for all possible lag structures and choosing the one with a minimal value for the loss function. Unfortunately, this trivial algorithm requires tremendous computing resources, since already a bivariate model with up to 20 lags allows for more than  $10^{12}$  different lag structures. For even more equations, or an increased number of lags to be considered, the problem becomes intractable.

Hence, an implementation of the Threshold Accepting algorithm (TA) introduced by Dueck and Scheuer (1990) is used to accomplish the task of identification. Like the more widely used – but at least for the Travelling Salesman problem less efficient – Simulated Annealing algorithm, TA belongs to the class of refined local search algorithms for discrete optimization problems. The basic idea of a local search algorithm is to compare a given element  $x_0$  in the set of possible solutions  $X$  with other elements in a neighbourhood  $x_0 \in \mathcal{U}(x_0) \subset X$  with regard to a loss function  $f$ . As the meaning of “neighbourhoods” is not given in a standard manner as e.g. for euclidian spaces ( $\varepsilon$ -spheres in  $\mathbb{R}^n$ ), it must be defined explicitly. It should be defined such that elements of a neighbourhood  $\mathcal{U}(x_0)$  of  $x_0$  are “close” to  $x_0$ . The choice of neighbourhoods for the identification problem under study will be discussed below.

The basic structure of any local search algorithm can be described as follows. In each iteration a new current solution  $x_1$  is searched in a neighbourhood  $\mathcal{U}(x_0)$  of the current solution  $x_0$ . If  $x_1$  can be accepted according to a given criterion it will replace  $x_0$  for the next step. These exchange steps can be iterated either for a limited number of iterations (*maxits*) or until no further improvement can be achieved. Figure 2 provides a flow chart for such a local search algorithm for a maximum of *maxits* iterations.

The central idea of the refined local search algorithms is related to the acceptance criterion for  $x_1$ . In order to escape “bad” local minima, a temporary worsening of the loss function  $f$  has to be admitted during the optimization process. Consequently, instead of accepting  $x_1$  if and only if  $\Delta E := f(x_1) - f(x_0) \leq 0$ , TA accepts  $x_1$  if and only if  $\Delta E \leq T$  for a given threshold value  $T > 0$ . To put it differently, Threshold Accepting accepts every new element  $x_1$  if the corresponding value of the loss function is **not much worse** than the old one, i.e. the new element is accepted if its value of

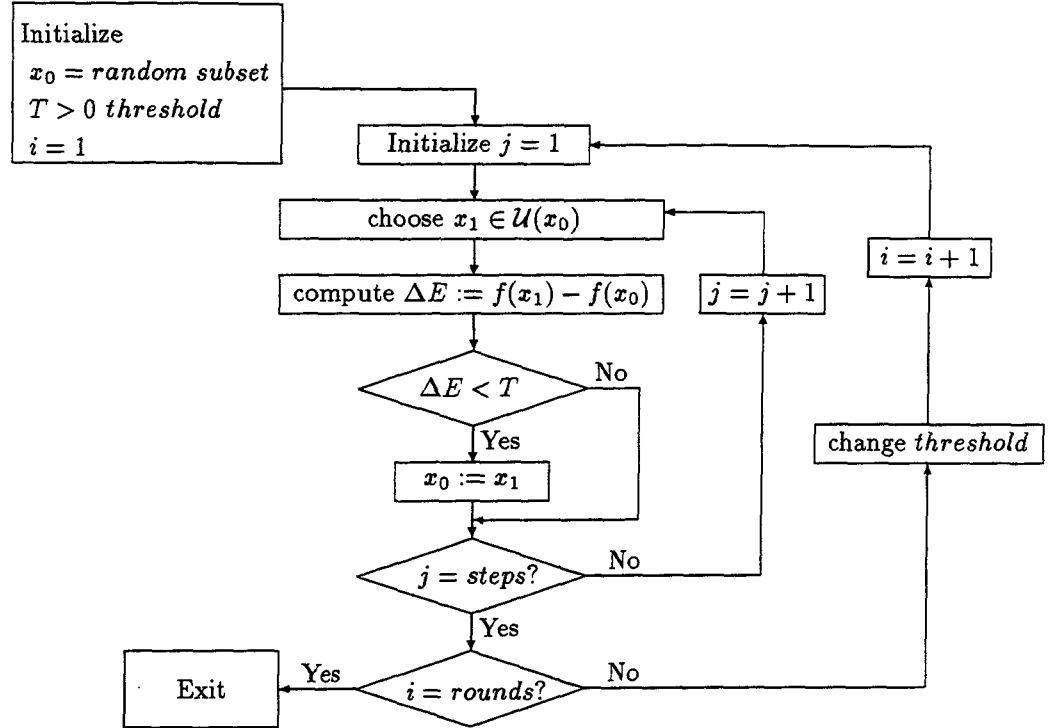
Figure 2: Flow Chart for Standard Local Search Algorithm



the loss function is not much higher than the value for  $x_0$ . Thus, it becomes possible to escape local minima with regard to the given neighbourhoods. It is due to this feature of the refined local search algorithms that they outperform any standard local search algorithm in complex discrete optimization problems. The threshold values  $T$  must decrease to zero during the optimization. Then, it can be guaranteed that the algorithm stops after a finite number of iterations. In general, the algorithm stops at a local minimum with respect to the chosen neighbourhood definition. In the best case, this local minimum is also a global minimum. Althöfer and Koschnick (1991) proved that with the number of iterations going to infinity TA will end up at the global optimum.

For the application in VAR-models different cases can be admitted for the set of possible solutions. It might contain either all possible lag structures for all equations, which is the most general case, or only all subsets of all considered lags for all variables, which will be the case used for the example in this paper. In the first case, the set of possible solutions  $X$  is equal to  $\{1, \dots, K'\}^d$ , where  $K'$  is the maximal number of lags to be considered and  $d$  is the dimension of the VAR, i.e. the number of included variables. For

Figure 4: Flow Chart for Threshold Accepting Algorithm



their distributions. Nevertheless, the method proposed makes it possible to automatically identify a lag structure minimizing a given loss function or at least to find a lag structure very close to the best one.

## 5 Results of a Simulation

A first application of the procedure introduced in the previous section was conducted with a bivariate AR-process as shown in figure 1. The data were

generated by the bivariate AR-process

$$\begin{aligned}
x_t &= 0.12x_{t-1} - 0.30x_{t-4} + 0.30x_{t-12} \\
&\quad + 0.46y_{t-1} - 0.20y_{t-2} + u_t \\
y_t &= 0.45x_{t-1} - 0.4x_{t-2} \\
&\quad + 0.3y_{t-1} - 0.12y_{t-3} + 0.16y_{t-12} - 0.17y_{t-18} + v_t,
\end{aligned} \tag{2}$$

where  $u_t$  and  $v_t$  were independently drawn from a standard normal distribution. The standard normal random numbers were generated by the random number generator of GAUSS. For all simulations the AR-process was started with constant values for  $x_1 \dots x_{20}$  and  $y_1 \dots y_{20}$ . The samples used for the identification began with  $t = 241$ . All samples had length 240 corresponding to 20 years of monthly data.

For the AFPE the algorithm ran 25 rounds with 100 iterations each while the constant for the FPE criterion increased from 2 to 11. Of course, a number of only 2500 iterations is quite small in order to find a good approximation to the optimum. Further, the choice of the sequence for the FPE criterion was nearly arbitrary. Nevertheless, the algorithm succeeded in identifying the model as described above missing only the term for  $y_{t-3}$  which was excluded in one of the last iterations.

In order to gain insights on the power of identifying procedures based on the TA implementation, a small simulation experiment was undertaken. The results are presented in tables 1 and 2. For the loss functions HQ with  $c = 2.0$  and for the AFPE 10, and for HQ with  $c = 1.01$ ,  $c = 1.1$ ,  $c = 1.2$  and  $1.3$  40 samples of the AR-process were generated. Only 1000 iterations of the TA algorithm were performed on each sample. The table indicates the frequency of the different lags in the model identified by the algorithm. Thus, a value of 100% for  $x_1, x_2, x_{12}$  and  $y_1$  shows, that these lags were always identified by the algorithm. It should be noted that the specification of the set of feasible solutions  $X$  for this application does not allow discrimination between lagged variables included in the equation for  $x_t$  and  $y_t$ , respectively. However, this more general case can also be studied using the same method.

Although tables 1 and 2 show that the TA algorithm was not always able to identify the "true" model correctly, the values of the loss function for the identified models were always equal to or less than the value for the "true" model, given the observed data sample. Therefore, one might conclude that this TA implementation is able to identify AR-processes automatically. This is done by the use of raw computing power instead of human resources which are better used for the interpretation of the results.

Table 1: Simulation Results for  $x$

LAG	HQ					AFPE
	c=1.01	c=1.1	c=1.2	c=1.3	c=2.0	
$x_1$	100%	100%	100%	100%	100%	100%
$x_2$	100%	100%	100%	100%	100%	100%
$x_3$	7.5%	5%	2.5%	5%	0%	10%
$x_4$	100%	100%	100%	100%	100%	100%
$x_5$	12.5%	12.5%	10%	7.5%	0%	0%
$x_6$	10%	7.5%	5%	5%	0%	0%
$x_7$	7.5%	5%	5%	5%	0%	0%
$x_8$	10%	10%	10%	7.5%	10%	20%
$x_9$	5%	2.5%	0%	2.5%	0%	0%
$x_{10}$	15%	15%	15%	12.5%	10%	10%
$x_{11}$	10%	7.5%	5%	7.5%	10%	10%
$x_{12}$	100%	100%	100%	100%	100%	100%
$x_{13}$	12.5%	10%	5%	5%	10%	10%
$x_{14}$	15%	15%	12.5%	12.5%	10%	10%
$x_{15}$	7.5%	7.5%	5%	2.5%	0%	0%
$x_{16}$	5%	2.5%	7.5%	7.5%	10%	20%
$x_{17}$	7.5%	7.5%	7.5%	10%	0%	0%
$x_{18}$	5%	5%	5%	2.5%	0%	0%
$x_{19}$	12.5%	15%	15%	15%	20%	10%
$x_{20}$	10%	10%	10%	7.5%	0%	0%

Table 2: Simulation Results for  $y$ 

	HQ					AFPE
LAG	c=1.01	c=1.1	c=1.2	c=1.3	c=2.0	
$y_1$	100%	100%	100%	100%	100%	100%
$y_2$	80%	80%	80%	80%	40%	60%
$y_3$	50%	45%	42.5%	37.5%	40%	30%
$y_4$	7.5%	7.5%	5%	5%	0%	0%
$y_5$	2.5%	2.5%	2.5%	2.5%	0%	0%
$y_6$	10%	7.5%	7.5%	7.5%	0%	0%
$y_7$	5%	5%	2.5%	2.5%	0%	0%
$y_8$	2.5%	2.5%	0%	2.5%	0%	0%
$y_9$	12.5%	12.5%	10%	10%	0%	0%
$y_{10}$	7.5%	7.5%	7.5%	7.5%	0%	0%
$y_{11}$	7.5%	5%	2.5%	2.5%	10%	10%
$y_{12}$	47.5%	50%	50%	45%	40%	40%
$y_{13}$	2.5%	2.5%	2.5%	2.5%	0%	0%
$y_{14}$	0%	0%	0%	0%	0%	0%
$y_{15}$	12.5%	10%	10%	7.5%	0%	20%
$y_{16}$	2.5%	2.5%	2.5%	2.5%	0%	0%
$y_{17}$	10%	7.5%	5%	5%	10%	10%
$y_{18}$	85%	85%	82.5%	80%	60%	60%
$y_{19}$	12.5%	10%	10%	5%	0%	0%
$y_{20}$	5%	2.5%	2.5%	2.5%	0%	0%
Number of included lags (mean)	10.15	9.825	9.45	9.225	7.5	8.3



## 6 Conclusion

This paper contributes to the discussion on how to choose lag structures in large-scale econometric models, Granger causality tests, or Johansen's procedure by proposing an automatic procedure to identify an optimal lag structure with regard to a given information criterion. The procedure is based on a heuristic discrete optimization algorithm which in the past was successfully implemented for other problems in economics and econometrics.

A simulation study for a bivariate AR-process shows that this method produces a good identification of the underlying lag structure limited mainly by the finite number of available observations. Thus, the identification and choice of lag structures can be achieved through the mechanical use of the presented Threshold Accepting implementation. The econometric model builder is freed from the task of testing hundreds of different lag structures, the resulting subjective factor is avoided, and, because lag structures in general will have many "holes", the number of degrees of freedom is increased.

The method introduced here for subset identification with respect to a given information criterion is not constrained to pure AR-models. Likewise, MA terms or dummy variables can be considered or exogenous variables might be included which are known to influence the process a priori as a result of economic theory.

## References

- Aarts, E. and J. Korst, *Simulated Annealing and Boltzmann Machines* (John Wiley & Sons, Chichester, 1989).
- Akaike, H., Fitting autoregressive models for prediction, *Annals of the Institute of Statistical Mathematics*, **21** (1969) 243–247.
- Akaike, H., Information theory and an extension of the maximum likelihood principle, in: B. N. Petrov and F. Csaki (Eds.), *Second International Symposium on Information Theory* (Budapest, 1973) 267–281.
- Akaike, H., Autoregressive model fitting for control, *Annals of the Institute of Statistical Mathematics*, **23** (1971) 163–180.
- Althöfer, I. and K.-U. Koschnick, On the convergence of "Threshold Accepting", *Applied Mathematics and Optimization*, **24** (1991) 183–195.

- Chipman, J.S. and P. Winker, Optimal aggregation by Threshold Accepting, *Working Paper No. 180* (Sonderforschungsbereich 178, Universität Konstanz, 1992).
- Dueck, G. and T. Scheuer, Threshold Accepting: A general purpose algorithm appearing superior to Simulated Annealing, *Journal of Computational Physics*, **90** (1990) 161–175.
- Dueck, G. and P. Winker, New concepts and algorithms for portfolio choice, *Applied Stochastic Models and Data Analysis*, **8** (1992) 159–178.
- Dueck, G. and J. Wirsching, Threshold Accepting algorithms for 0–1 knapsack problems, in: H. Wacker and W. Zulehner (Eds.), *Proceedings of the Fourth European Conference on Mathematics in Industry* (B. G. Teubner, Stuttgart, and Kluwer Academic Publishers, 1991) 255–262.
- Hannan, E. J. and B. G. Quinn, The determination of the order of an autoregression, *Journal of the Royal Statistical Society*, **B 41**(2) (1979) 190–195.
- Hocking, R. R. and R. N. Leslie, Selection of the best subset in regression analysis, *Technometrics*, **9**(4) (1967) 531–540.
- Hsiao, C., Autoregressive modelling and money–income causality detection, *Journal of Monetary Economics*, **7** (1981) 85–106.
- Johansen, S., Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, **59**(6) (Nov. 1991) 1551–1580.
- Kirkpatrick, S., C. Gelatt and M. Vecchi, Optimization by Simulated Annealing, *Science*, **220** (1983) 671–680.
- McClave, J., Subset autoregression, *Technometrics*, **17**(2) (1975) 213–220.
- Odaki, M., Tests of Granger causality by the selection of the orders of a bivariate autoregressive model, *Economics Letters*, **22** (1987) 223–227.
- Pagano, M., An algorithm for fitting autoregressive schemes, *Journal of the Royal Statistic Society (Series C) Applied Statistics*, **21** (1972) 274–281.
- Pierce, D.A. and L. D. Haugh, Causality in temporal systems, *Journal of Econometrics*, **5** (1977) 265–293.

- Schwarz, G., Estimating the dimension of a model, *The Annals of Statistics*, **6**(2) (1978) 461–464.
- Sen, D.L. and D. A. Dickey, Symmetric test for second differencing in univariate time series, *Journal of Business & Economic Statistics*, **5**(4) (1987) 463–473.
- Shibata, R., Selection of the order of an autoregressive model by Akaike's information criterion, *Biometrika*, **63** (1976) 117–126.
- Sims, C. A., Money, income, and causality, *American Economic Review*, (1972), 540–552.
- Winker, P., The Tuning of the Threshold Accepting Heuristic for Traveling Salesman Problems, Mimeo. (Sonderforschungsbereich 178, Universität Konstanz, 1994).