

Rech, Gianluigi

Working Paper

Forecasting with artificial network models

SSE/EFI Working Paper Series in Economics and Finance, No. 491

Provided in Cooperation with:

EFI - The Economic Research Institute, Stockholm School of Economics

Suggested Citation: Rech, Gianluigi (2002) : Forecasting with artificial network models, SSE/EFI Working Paper Series in Economics and Finance, No. 491, Stockholm School of Economics, The Economic Research Institute (EFI), Stockholm

This Version is available at:

<https://hdl.handle.net/10419/56171>

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.

Forecasting with artificial neural network models

Gianluigi Rech

Quantitative Analysis - ELECTRABEL, Place de l'université, 16 B-1348
Louvain-la-Neuve BELGIUM;

Department of Economic Statistics, Stockholm School of Economics,
Box 6501, S-113 83, Stockholm, Sweden.

E-mail: gianluigi.rech@electrabel.com.

SSE/EFI Working Paper Series in Economics and Finance

No 491

January 2002

Abstract

This paper contains a forecasting exercise on 30 time series, ranging on several fields, from economy to ecology. The statistical approach to artificial neural networks modelling developed by the author is compared to linear modelling and to other three well-known neural network modelling procedures: Information Criterion Pruning (ICP), Cross-Validation Pruning (CVP) and Bayesian Regularization Pruning (BRP). The findings are that 1) the linear models outperform the artificial neural network models and 2) albeit selecting and estimating much more parsimonious models, the statistical approach stands up well in comparison to other more sophisticated ANN models.

JEL classification: C22, C45, C53

Acknowledgments: I am very grateful for the help I received from Marcelo Medeiros, who developed the programs which produced forecasting results and comparisons for the linear, ICP, CVP and BRP models. Financial support from the Tore Browaldh Foundation and the Stockholm School of Economics is gratefully acknowledged.

1 Introduction

Over recent years, several nonlinear time series models have been proposed in the literature, see f.i. Tong (1990), Granger and Teräsvirta (1993), van Dijk, Teräsvirta, and Franses (2001). One model that has found a large number of successful applications is the Artificial Neural Network model (hereafter ANN). The ANN model is a mathematical model inspired by the function of the human brain and its use is mainly motivated by its capability of approximating any Borel-measurable function to any degree of accuracy, as pointed out in Hornik, Stinchcombe, and White (1989).

In this paper the forecasting performance of a number of artificial neural network (ANN) models is compared with that of linear models. The idea is to find out whether or not different techniques of specifying ANN models, such as pruning and statistical techniques, lead to vastly different models. If they do, then the next question is whether this matters from a forecasting point of view. Values of 30 economic and other time series are predicted using models obtained by a number of different modelling techniques and the forecasts compared using a number of different criteria. The plan of the paper is as follows. Section 2 introduces the ANN models. Section 3 discusses the techniques used in this paper to build ANN models. Section 4 gives a brief discussion on linear ARMA models. The forecasting experiment is described in Section 5. Concluding remarks are made in Section 6.

2 Neural Network Model

The Artificial Neural Network (ANN) time series model is usually defined as

$$y_t = G(\mathbf{w}_t; \Psi) = \beta_0 + \sum_{j=1}^q \beta_j \psi(\gamma_j' \mathbf{w}_t) + u_t, \quad (1)$$

where $\beta = (\beta_1, \dots, \beta_q)'$, $\gamma_j = (\gamma_{j0}, \gamma_{j1}, \dots, \gamma_{j,k-1}, c_j)'$, $j = 1, \dots, q$; $\mathbf{w}_t = (w_{1t}, w_{2t}, \dots, w_{kt}, 1)'$ = $(y_{t-1}, y_{t-2}, \dots, y_{t-k}, 1)'$, u_t is *n.i.d.* and $\psi(\gamma_i' \mathbf{w}_t)$ is the sigmoid function defined as

$$\psi(\gamma_i' \mathbf{w}_t) = \frac{1}{1 + e^{-\gamma_i' \mathbf{w}_t}}. \quad (2)$$

In applications to economic time series it is normally useful to include a linear component in (1). The resulting model is defined as in Granger and Teräsvirta (1993)

$$y_t = G(\mathbf{w}_t; \Psi) = \alpha' \mathbf{w}_t + \sum_{i=1}^q \beta_i \psi(\gamma_i' \mathbf{w}_t) + u_t, \quad (3)$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_0)'$ is a vector of real coefficients and u_t is *n.i.d.*. One of the characteristic features of this model is that it is not globally identified. There are three characteristics of neural networks which cause the non-identifiability. The first one is due to the symmetries in the neural network architecture. The likelihood function of the model will be unchanged if one permutes the hidden units, resulting in $q!$ possibilities for each one of the coefficients of the model, and the same can occur because of the fact that $\psi(x) = 1 - \psi(-x)$ (see Hwang and Ding (1997)). The third reason is the mutual dependence of the parameters β_i and γ_i , $i = 1, \dots, q$. If $\beta_i = 0$, the corresponding γ_i can assume any value without affecting the value of the likelihood function. On the other hand, if $\gamma_i = 0$, then β_i can take any value. Nevertheless, under certain regularity conditions, the maximum likelihood estimates of the parameters of (3) are consistent and asymptotically normal; for a broader discussion, see Rech (2001).

3 Modelling

3.1 What is Modelling?

In this paper, I consider modelling and forecasting of time series with ANN models. Modelling in this context is taken to mean model specification, parameter estimation, and, finally, evaluation of the estimated model. The idea is to find out whether or not different modelling techniques lead to vastly different ANN models. If they do, the next question is whether this matters from a forecasting point of view. I compare well-known techniques in the neural network literature, such as, *early stopping*, *pruning*, and *regularization*, with a novel one based on statistical tests. It is one of the few techniques making use of in-sample evaluation. The other techniques considered in this work rely on out-of-sample evaluation, i.e., checking the forecasting performance of the specified and estimated model.

In this paper, specification consists of selecting the variables of the ANN model from a set of candidates and determining the number of hidden units. This stage is of crucial importance in order to avoid overfitting and to find an adequate approximation of the true data generating process. In most neural network applications, it is customary to carry out the specification using some “rule of thumb”. A vast number of ANN models with different combinations of variables and number of hidden units are estimated, and the one with the best performance according to some known criterion is chosen as the final specification. The statistical approach is based on a sequence of hypothesis tests and gradually increasing the size of the ANN model until the test results suggest one to stop.

With the exception of methods of early stopping and regularization, estimation of parameters in ANN models is carried out by means of nonlinear least squares (NLS) or conditional maximum likelihood (ML). If the errors are normal, independent, and identically distributed, as assumed above, these two methods are equivalent. Hence, the parameter vector Ψ of (1) or (3) is estimated as

$$\hat{\Psi} = \underset{\Psi}{\operatorname{argmin}} Q_T(\Psi) = \underset{\Psi}{\operatorname{argmin}} \sum_{t=1}^T (y_t - G(\mathbf{w}_t; \Psi))^2 \quad (4)$$

In most applications a simple gradient descent algorithm (backpropagation) is used to estimate the parameters. However, the estimation of Ψ is usually not easy (*Hush (1999)*), and in general the optimization algorithm is very sensitive to the choice of the starting-values of the parameters. The use of algorithms such as the Broyden-Fletcher-Goldfarb-Shannon (BFGS) algorithm or the Levenberg-Marquardt are strongly recommended for the final estimation. See Bertsekas (1995) for details about optimization techniques.

Any estimated model must be evaluated. In econometrics, model misspecification tests play an important role in model evaluation. Most often, these tests concern the estimated residuals of the model. The residuals are used for testing assumed properties of the error process, such as, serial independence and normality. Testing parameter stability is another important model evaluation test. In typical ANN applications, out-of-sample forecasting appears to be the only evaluation tool available. This is, however, not surprising because the asymptotic properties of the parameter estimators are often unknown. The statistical ANN modelling procedure considered here (referred hereafter by SA) is intended to fill the gap between econometric model building and ANN modelling practices by proposing a number of misspecification tests for ANN model building.

3.2 ANN Modelling Based on a Statistical approach (SA)

The capability of single hidden-layer feedforward neural networks (hereafter NN) to approximating any Borel-measurable function to any degree of accuracy has been pointed out in Hornik, Stinchcombe, and White (1989). Nonlinear features in time series can then be successfully modelled applying statistical tools to the data of interest, since the connection between NN and statistics is generally well accepted. I call the following model an autoregressive NN model of order k with q hidden units and a linear component:

$$y_t = \boldsymbol{\alpha}' \mathbf{w}_t + \sum_{j=1}^q \beta_j \psi(\boldsymbol{\gamma}_j' \mathbf{w}_t) + u_t, \quad t = 1, \dots, T \quad (5)$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_0)'$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)'$, $\boldsymbol{\gamma}_j = (\gamma_{j0}, \gamma_{j1}, \dots, \gamma_{j,k-1}, c_j)'$, $j = 1, \dots, q$; $\mathbf{w}_t = (w_{1t}, w_{2t}, \dots, w_{kt}, 1)'$ and u_t is *n.i.d.*. Here, I assume for simplicity that all lags from 1 to k enter model (5).

Summing up, the overall device works as follows. The set of variables to be included in the model is selected as in Rech, Teräsvirta, and Tschernig (2001). The hypothesis of no nonlinear hidden units (linear model) is tested at a given significance level α . If rejected, a model with a linear part and one hidden unit is estimated and the approximate t-values of the parameters computed, approximating the covariance matrix of the parameters by the outer product of the gradient matrix. The lags with low t-values are removed and the model re-estimated. The whole procedure is redone until the hidden unit contains only significant estimates. Subsequently, the hypothesis of no additional hidden units is tested at the significance level $\alpha/2$. If rejected, a model with two hidden units is estimated and the dimension of the model reduced by checking the t-values of its estimates as above. The procedure continues halving the significance level again to $\alpha/4$, $\alpha/8$, ..., stopping the procedure at the first acceptance of the null hypothesis of no additional hidden units. Letting the significance level converge to zero as $q \rightarrow \infty$ keeps the dimension of the model under control.

Evaluating a model requires, as in Eitrheim and Teräsvirta (1996), to develop specific LM tests for the hypothesis of no error autocorrelation and parameter constancy, while additional nonlinearity is already checked when I choose the number of hidden units. The test for error autocorrelation is based on model (5), where the residuals u_t follow an autoregressive process of order r , $u_t = \sum_{j=1}^r a_j u_{t-j} + \varepsilon_t$, $\varepsilon_t \sim n.i.d.(0, \sigma^2)$. The corresponding LM test for the hypothesis $H_0 : \mathbf{a} = 0$ can be carried out in 3 steps as in testing for q against $q + 1$ hidden units. As to parameter constancy, I generalize model (5) assuming that the hidden units have constant parameters whereas both β s and α s may change smoothly over time. Therefore $\boldsymbol{\alpha} = \boldsymbol{\alpha}(t) = \alpha_0 + \boldsymbol{\lambda}_2 F_1(t, \gamma_1, c_1)$ and $\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{\lambda}_1 F_j(t, \gamma_j, c_j)$, where the $F_j(\cdot)$ s are transitional functions monotonically increasing and bounded between zero and one. The null hypothesis of parameter constancy implies that $F_j(t, \gamma_1, c_1) \equiv \text{constant}$ for any t . For further details on the modelling procedure, see Rech (2001).

3.3 Early Stopping

In this paper, I compare forecasts from models obtained applying the modelling cycle of Section 3.2 to forecasts from ANN models of the form of (1) obtained by other means. The first alternative method to be considered is the so-called *early stopping*. It is perhaps one of the most simple and vastly used modelling technique in the ANN framework. The key idea is to split the available data into three subsets. The first subset is used to estimate the

parameters. The second subset is called the validation set. The error on the validation set is monitored during the estimation process. When the network begins to overfit the data, the error on the validation set typically begins to rise. When the validation error increases for a specified number of iterations, the estimation process is discontinued, and the parameters estimated at the minimum of the validation error serve as final estimates.

The test set is not used for estimation, but it is saved for comparing different models. A large number of different specifications are estimated and compared by means of the out-of-sample performance. The model with the best forecasting performance is chosen as the final specification.

3.4 Pruning

Pruning is another popular technique. The objective of pruning is to find the smallest network that fits the data well and produces good forecasts. I consider three pruning techniques. For a general survey on pruning see Reed (1993). The first two, information criterion pruning and cross-validation pruning, are described in detail in Anders and Korn (1999). The third one, interactive pruning, is based on the results of Kaashoek and van Dijk (1998), Siestma and Dow (1991), and Siestma and Dow (1988).

Although there are many other pruning algorithms, the three to be considered here are the one most frequently used in applications.

To estimate the parameters, I used the BFGS optimization procedure with the starting-values given by the Nguyen-Widrow algorithm (Nguyen and Widrow (1990)).

3.4.1 Information Criterion Pruning

The idea of the Information Criterion Pruning (ICP) is to choose between two models with different degrees of complexity. This is done using information criterion (IC) such as, AIC Akaike (1974) or SBIC (Rissanen (1978); Schwarz (1978)). As pointed out in Anders and Korn (1999), such information criteria are derived based on the assumption of asymptotic normality of the maximum likelihood estimators in globally identified models. Hence, their use is not theoretically justified in the neural network case when the ANN model may be over-parametrized and thus unidentified. To circumvent this problem, Anders and Korn (1999) suggested the following methodology to estimated ANN models based on IC measures.

First, estimate the ANN model with only one hidden unit. Next, compute the residuals and regress them on a third-order Taylor expansion of an additional hidden unit. Calculate the value of information criterion and accept the additional hidden unit if the value of the IC is less than the IC of a white noise. Continue adding hidden units until there is no further improvement. Note that this resembles the strategy presented in Section 3.2. However, in determining the number of hidden units, information criteria are used in place of statistical tests.

The second step of the procedure starts from the fully connected network obtained in the first step and tries to detect irrelevant input connections. One thus proceeds in the opposite direction. All submodels with one of the input connections removed are estimated and compared with the full network by means of the IC. If the full network turns out to have the lowest IC value, this model is accepted as the final one. Otherwise the lowest IC submodel is chosen to serve as the baseline for the next round of the specification search. Again all submodels having one input connection less than the baseline model are compared with the latter using an appropriate IC.

It is also possible to start with a network with a large number of hidden units and proceed by removing units one at a time, if necessary. This “general to specific” alternative may involve estimating a considerable amount of unidentified or near-unidentified models and is not applied in this work.

3.4.2 Cross-Validation Pruning

The idea of the cross-validation pruning (CVP), or, more specifically, leave- k -out cross-validation, is the following. In order to find an optimal number of hidden units and the correct set of input variables, it is appealing to compare the mean squared prediction errors (MSPE) of different model specifications. Such prediction errors are obtained by dividing the sample into M subsets S_m , $m = 1, \dots, M$, which contain k observations each. The ANN model is repeatedly re-estimated, leaving out a different subset each time. The average MSPE on the M subsets that have been left out defines the cross-validation error, CV:

$$CV = \frac{1}{M} \sum_{m=1}^M MSPE_m, \quad (6)$$

where

$$MSPE_m = \frac{1}{k} \sum_{i=1}^k (\hat{y}_i - y_i)^2. \quad (7)$$

The model with the lowest cross-validation error is finally chosen. The CVP is independent of probabilistic assumptions. The main disadvantage of the CVP method is that re-estimating the models for each one of M subsets can be time-consuming. In principle, all combinations of input variables and hidden units can be compared which requires plenty of computational resources.

The CVP is carried out in two stages. The first one consists of estimating the number of hidden units and the second one of removing any irrelevant input variables. The first step is summarized as follows. Start fitting a model with only one hidden unit and compute the value of CV . Add another hidden unit to the model and compute the CV again. Compare the new CV value with the previous one. If the former is smaller than the latter, add another hidden unit. Otherwise discontinue the process and proceed to the second stage, which is the following. Compute the value of the cross-validation error for all submodels containing one input connection less than the original model. If at least one submodel turns out to have an smaller CV than the full model, remove that variable and continue without it. Otherwise terminate the process. The finally chosen ANN model is such that no variable can be removed without increasing the cross-validation error.

As in the ICP method, it is also possible to start with a network with a large number of hidden units and proceed by removing one unit at each round. Again, this is not done in this paper. Several papers have pointed out arguments for and against the CVP.

For example, Stone (1977) presented an insightful discussion of the limitations of cross-validation. See also Zhu and Rohwer (1996), Goutte (1997), and Rivals and Personnaz (1999).

3.4.3 Interactive Pruning

Another way of identifying irrelevant units and/or inputs is the so called interactive pruning (IP). The main idea is to start by estimating an ANN model with a large number of both

hidden units and variables. After estimating the parameters of the model, the incremental contribution of each hidden unit is considered. This is done in several ways. Siestma and Dow (1991) and Siestma and Dow (1988) tackled the problem by removing hidden units with variance close to zero ($< 10^{-8}$) and combining units with highly correlated responses ($\rho = 0.98$, for example). Kaashoek and van Dijk (1998) suggested a different approach based on the correlation coefficient of the fitted values \hat{y}_t , defined as

$$R^2 = 1 - \frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{\sum_{t=1}^T (y_t - \bar{y})^2}, \quad (8)$$

where $\bar{y} = (1/T) \sum_{t=1}^T y_t$. The idea is to remove the i^{th} hidden unit, $i = 1, \dots, q$, re-estimate the model, and compute the value of the percentage difference in R^2 , defined as

$$\Delta R^2 = \frac{R^2 - R_i^2}{R^2}, \quad (9)$$

where R_i^2 is the R^2 without the i^{th} hidden unit. However, re-estimation means that only the vector of ‘‘connection strengths’’, β in (1) is re-estimated each time. If ΔR^2 is small ($\leq 1\%$), I remove the unit from the model. After removing the hidden units deemed unimportant, the procedure is continued by removing one input variable each time and comparing the correlation coefficients. Irrelevant variables are removed from the model following the same procedure as before, based on ΔR^2 . Again, only the connection strength vector β is re-estimated each time.

The main drawback of this method is that estimating an ANN model with a large number of inputs and hidden units is very difficult and quite often the estimation algorithm does not converge.

3.5 Regularization

The last modelling technique considered here is regularization. The idea, familiar from statistics, is to find a balance between the number of parameters and goodness of fit by penalizing large models. The objective function is modified in such a way that the estimation algorithm effectively prunes the network by driving irrelevant parameter estimates to zero during the estimation process. The parameter vector Ψ is estimated as

$$\hat{\Psi} = \underset{\Psi}{\operatorname{argmin}} \tilde{Q}_T(\Psi) = \underset{\Psi}{\operatorname{argmin}} (Q_T(\Psi) + \eta Q_T^*(\Psi)), \quad (10)$$

where $Q_T(\Psi) = \sum_{t=1}^T (y_t - G(\mathbf{w}_t; \Psi))^2$ as in (4), $Q_T^*(\Psi)$ is the *regularization* or *penalty term*, and $\eta > 0$ is often called the *decay constant*.

To be in agreement with the terminology often used in the ANN literature, in this section ‘‘weight’’ means parameter and ‘‘generalization’’ is a synonym to out-of-sample forecasting.

3.5.1 Weight Decay

The usual penalty is the sum of squared parameters times a constant. In a linear model, this form of regularization, called *weight decay*, is equivalent to ridge regression. For the ANN model (1), the parameters are estimated as

$$\hat{\Psi} = \underset{\Psi}{\operatorname{argmin}} \tilde{Q}_T(\Psi) = \underset{\Psi}{\operatorname{argmin}} \left(\sum_{t=1}^T (y_t - G(\mathbf{w}_t; \Psi))^2 + \eta \sum_{j=1}^n \Psi_j^2 \right), \quad (11)$$

where Ψ_i is a typical element of the parameter vector Ψ and n is the number of parameters of the model.

The penalty term in (11), by definition, penalizes large estimates. There exist other regularization methods where the penalty term involves not only the parameter estimates but various partial derivatives of $G(\mathbf{w}_t; \Psi)$ as well.

The weight decay penalty term makes the parameter estimates smaller in absolute value than they otherwise would be. Large estimates can affect the forecasting performance in two different ways. Very large parameters estimates of γ_{ij} , $i = 1, \dots, q$, $j = 0, \dots, p$, in (1) or (3) lead to hidden units that cause $G(\mathbf{w}_t; \Psi)$ to be too rough, possibly with near discontinuities. Excessively large estimates of β can cause outputs far beyond the range of the data, if the output activation function is not bounded to the same range as the data. To put it another way, large parameter estimates can cause excessive variance of the output (Geman, Bienenstock, and Doursat (1992)). According to Bartlett (1997), the size (L_1 norm) of the parameter estimates is more important than the number of parameters in determining the out-of-sample performance of the estimated ANN model.

Other penalty terms besides the sum of squared weights are sometimes used.

3.5.2 Weight Elimination

Weigend, Rumelhart, and Huberman (1991) proposed the *weight elimination* where

$$\hat{\Psi} = \underset{\Psi}{\operatorname{argmin}} \tilde{Q}_T(\Psi) = \underset{\Psi}{\operatorname{argmin}} \left(\sum_{t=1}^T (y_t - G(\mathbf{w}_t; \Psi))^2 + \eta \left(\frac{\sum_{j=1}^n \Psi_j^2}{1 + \sum_{j=1}^n \Psi_j^2} \right) \right). \quad (12)$$

While the penalty term using the sum of squared weights tends to shrink the large estimates more than the small ones, weight elimination tends to shrink the small coefficients more, and is therefore more useful for suggesting subset models (pruning).

The forecasting ability of the ANN model can depend crucially on the decay constant η , especially with small sets of data used for estimation. If η is too small, the network may still overfit, and if it is too large, the ANN model does not have an adequate fit in the estimation period. Usually, different types of parameters in the ANN model will usually require different decay constants for good forecasting ability.

One approach to choosing the decay constant is to estimate several networks with different values of η and choose the one that minimizes the sum of squared out-of-sample residuals. Weigend, Rumelhart, and Huberman (1991) iteratively update the decay constant during estimation. Adjusting all these decay constants to produce the model with the best forecasting ability often requires vast amounts of computation.

Fortunately, there is a superior alternative to weight decay and weight elimination: the *Bayesian regularization*. Bayesian regularization makes it possible to estimate efficiently numerous decay constants.

3.5.3 Bayesian Regularization

One approach to determining the optimal regularization parameter η is the Bayesian framework of MacKay (1992). In this context the coefficients of the network are assumed to be random variables with well-specified distributions. The regularization parameters are related to the unknown variances associated with these distributions and can be estimated with statistical techniques. Foresee and Hagan (1997) gave a detailed discussion of the

use of Bayesian regularization in combination with the Levenberg-Marquardt optimization algorithm. The function to minimize is

$$\operatorname{argmin}_{\Psi} \left(\zeta \sum_{t=1}^T (y_t - G(\mathbf{w}_t; \Psi))^2 + \eta \left(\frac{\sum_{j=1}^n \Psi_j^2}{1 + \sum_{j=1}^n \Psi_j^2} \right) \right) = \operatorname{argmin}_{\Psi} (\zeta Q_1(\Psi) + \eta Q_2(\Psi))$$

The steps the authors follow to minimize the objective function are:

- set $\zeta = 1$, $\eta = 0$ and use the Nguyen-Widrow method of initializing the weights;
- Take one step of the Levenberg-Marquardt algorithm to minimize $Q_1(\Psi)$;
- compute the effective number of parameters $\hat{\delta} = \dim(\Psi) - 2\eta \cdot \operatorname{tr}(\hat{\mathbf{H}}^{-1})$, where $\hat{\mathbf{H}}$ is the Gauss-Newton approximation to the hessian;
- compute the new estimates for ζ and η , $\hat{\zeta} = \frac{T - \hat{\delta}}{2Q_1(\hat{\Psi})}$ and $\hat{\eta} = \frac{\hat{\delta}}{2Q_2(\hat{\Psi})}$, where T is the number of observations;
- iterate step 2 to 4 until convergence.

The main advantage of this method is that even if the ANN model is over-parametrized, the irrelevant parameter estimates are likely to be close to zero and the model behaves like a small network.

4 AR approximation to Box & Jenkins' ARMA modelling

The ARMA(p_1, p_2) model is defined as

$$(1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_{p_1} L^{p_1}) y_t = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_{p_2} L^{p_2}) u_t, \quad (13)$$

where $\rho_1, \dots, \rho_{p_1}, \theta_1, \dots, \theta_{p_2}$ are real parameters, L is the backshift operator defined as $L^i y_t = y_{t-i}$, p_1 is the autoregressive order, p_2 is the moving-average order, and $u_t \sim \text{NID}(0, \sigma_u^2)$. Stationarity and invertibility are basic assumptions. Under such conditions, we may write (13) in the autoregressive representation, that is, an $AR(p = \infty)$ process:

$$(1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_{p_1} L^{p_1})(1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_{p_2} L^{p_2})^{-1} y_t = u_t, \quad (14)$$

or, equivalently:

$$(1 - a_1 L - a_2 L^2 - \dots) y_t = u_t, \quad (15)$$

A linear model of the type (15), with a finite number of lags, can approximate (13) satisfactorily if a sufficient number of lags is encompassed. I choose the maximum lag $p = 25$ because several series are monthly and their need to be investigated at least 2 years back in time. The estimation can be performed by ordinary least squares (hereafter OLS). Hereafter, I call such modelling procedure "AR" and I shall refer to it by this abbreviation.

The AR model building procedure is summarized as follows:

1. Specification

- (a) If necessary, take logarithm or power transformations.

- (b) If necessary, take short and/or long (seasonal) differences of y_t to achieve stationarity in the mean.
 - (c) Determine the order of the AR model, i. e., choose the value of p . This is done by examining the partial autocorrelation (PACF) functions of the series after taking the differences and transforming the data (if it is necessary).
2. Estimate the parameters of model (15) by OLS.
 3. Model evaluation.
 - (a) Test whether the errors are white noise (Ljung-Box statistic).
 - (b) Test whether the errors are normally distributed (skewness, kurtosis and the Jarque-Bera test).
 4. Reduce the dimension of the estimated model using some information criterion like SBIC or AIC.

5 The Experiment

ANN models are mainly used for forecasting. To assess the practical usefulness of different methods of building an ANN model and the linear ARIMA model I conduct a forecasting experiment. The forecasts made by each estimated model are compared according to several statistics. I estimated ANN models for 30 different time series using the following model building procedures: statistical methods, information criterion pruning, cross-validation pruning, and Bayesian regularization. I also tried to estimate ANN models based on interactive pruning. However, in most of the cases considered here I was unable to estimate a satisfactory model. Thus, I discarded these results from the present paper. To compare the performance of the ANN models with the linear model, I also estimate, for each series, a linear autoregressive model.

5.1 Forecasting

The forecasts were made according to the following procedure.

1. Split the sample into two subsamples: the estimation set ($t = 1, \dots, t_0$) and the forecasting set ($t = t_0 + 1, \dots, T$).
2. Estimate the parameters of each model using only the estimation set.
3. For $t = t_0, \dots, T - 12$, compute the out-of-sample forecasts of one to 12-step-ahead, $\hat{y}_t(h)$, and the associated forecast errors denoted by $\hat{u}_t(h)$ where h is the forecasting horizon.
4. For each forecasting horizon, compute the following performance measures:
 - (a) Normalized root mean squared error (nRMSE):

$$nRMSE(h) = \sqrt{\frac{1}{T - t_0 - 11} \frac{\sum_{t=t_0}^{T-12} \hat{u}_t^2(h)}{\hat{\sigma}_y^2}}, \quad (16)$$

where $\hat{\sigma}_y^2$ is the estimated in-sample unconditional variance of the series. The $nRMSE(h)$ is a measure of the relative forecasting efficiency of the estimated model in relation to $\hat{y}_t(h) = \bar{y}$, where \bar{y} is the in-sample mean of the series.

(b) Mean Absolute Error (MAE):

$$MAE(h) = \frac{1}{T - t_0 - 11} \sum_{t=t_0}^{T-12} |\hat{u}_t(h)| \quad (17)$$

(c) Median Absolute Deviation (MAD):

$$MAD(h) = \text{median}(|\hat{u}_t(h) - \text{median}(\hat{u}_t(h))|). \quad (18)$$

Reporting the MAD was suggested by van Dijk (1999) and can be interpreted as a measure that is robust to outliers.

Another interesting way of comparing the forecasting performance of different models is to compare the sign of $(y_{t+h} - y_t)$ with the sign of $(\hat{y}_{t+h} - y_t)$. This can be done with a 2×2 contingency table, known as the *confusion matrix* and defined as

$$\begin{array}{cc} & \text{actual} \\ & \begin{array}{cc} up & down \end{array} \\ \begin{array}{c} predicted \\ \\ \end{array} & \begin{array}{|cc|} \hline up & \begin{array}{|c|} \hline a_{11} & a_{12} \\ \hline \end{array} \\ \hline down & \begin{array}{|c|} \hline a_{21} & a_{22} \\ \hline \end{array} \\ \hline \end{array} \end{array} \quad (19)$$

The columns in (19) correspond to actual moves, up or down, while the rows correspond to predicted moves. In this way, the diagonal cells correspond to correct directional predictions, while the off-diagonal cells correspond to incorrect predictions. I measure the overall performance in terms of the *confusion rate*, defined as

$$CR = \frac{a_{12} + a_{21}}{a_{11} + a_{12} + a_{21} + a_{22}}. \quad (20)$$

In order to test the hypothesis of equal accuracy in forecast performance between two estimated models, I use the Diebold-Mariano statistic Diebold and Mariano (1995) with the correction proposed by Harvey, Leybourne, and Newbold (1997a). The Diebold-Mariano test works as follows. Consider two competing models m_1 and m_2 which have produced a vector of h -step ahead estimated forecast errors $\hat{\varepsilon}_{m_i}(h) = (\hat{\varepsilon}_{1,m_i}(h), \dots, \hat{\varepsilon}_{t_f,m_i}(h))'$. Consider the difference

$$d_t = g(\hat{\varepsilon}_{t,m_1}(h)) - g(\hat{\varepsilon}_{t,m_2}(h)), t = 1, \dots, t_f$$

where $g(\cdot)$ is a generic loss function. Here I base $g(\hat{\varepsilon}_{t,m_i}(h))$ on the $nRMSE(h)$ and on $MAE(h)$, so that the null hypothesis of equal performance can be defined as

$$E[nRMSE(\hat{\varepsilon}_{t,m_1}(h)) - nRMSE(\hat{\varepsilon}_{t,m_2}(h))] = 0$$

or

$$E[MAE(\hat{\varepsilon}_{t,m_1}(h)) - MAE(\hat{\varepsilon}_{t,m_2}(h))] = 0$$

for all t and h . Defining $d_t = \widehat{\varepsilon}_{t,m_1}^2(h) - \widehat{\varepsilon}_{t,m_2}^2(h)$, I denote its mean and variance as follows

$$\bar{d} = \frac{1}{t_f} \sum_{t=1}^{t_f} d_t, \hat{V}(\bar{d}) = \frac{1}{t_f} \left(\hat{\gamma}_0 + 2 \sum_{k=1}^{h-1} \hat{\gamma}_k \right)$$

where $\hat{\gamma}_k$ is the estimated k -th autocovariance and $\hat{\gamma}_0$ the estimate of the variance. The Diebold-Mariano statistic is

$$S = \left[\hat{V}(\bar{d}) \right]^{-1/2} \bar{d}$$

which is asymptotically distributed as a $N(0,1)$ under the null hypothesis. The test is oversized in small samples; Harvey, Leybourne, and Newbold (1997b) suggested to use:

$$S^* = S \cdot \left(\frac{t_f + 1 - 2h + h(h-1)/t_f}{t_f} \right)^{1/2}$$

where the statistic S^* is compared with the critical values of a student t distribution with $n-1$ degrees of freedom since $\hat{V}(\bar{d})$ has to be estimated.

5.2 The Data

The time series used in the forecasting experiment are summarized in Table 1. For each of the 30 series, the applied transformation, the period of interest, the total number of observations T and the number of observations t_0 utilized for estimating the models are given. The first twelve series were obtained from *Economagic* (www.economagic.com). The rest of the series were obtained from Rob J. Hyndman's Time Series Data Library home page (www-personal.buseco.monash.edu.au/hyndman/TSDL/mhts/). Logarithmic transformations were applied to some of the series to stabilize the variance, and first or twelve-month differences were taken to achieve stationarity and/or to remove seasonal variation. The idea is to perform an investigation on a broad set of data with a different sample size, related to topics from macro and financial economics, to physics, to ecology and health. As to macro-economic data, I included the M1, M2 and M3 series for US money stock, the unemployment rate, the industrial production, the consumer and the producer price index for the US. The set of financial series includes 3 daily exchange rates (Canadian \$/US \$, German mark/US \$, Japanese yen/US \$) and 2 monthly stock indexes (S&P 500 and Dow Jones). All these series are monthly, seasonally adjusted, and transformed by taking the first difference of logarithms, such that I model their approximated growth rates. One series represents the volcanic dust veil index, 4 series relate to river flows, 2 to blow-flies population, 2 to ozone concentration, 3 to temperatures (1 monthly and 2 yearly), and 3 to cases of chickenpox, measles and mumps. Two classical benchmarks for nonlinear models, the lynx and the annual sunspot series, are included too.

5.3 Estimation Results

Table 2 shows the specification results. The set of selected lags widely differ depending on the different method employed. The SA approach is the most parsimonious both in terms of number of hidden units included in the estimated models and number of selected lags. It is also the only one which allows the detection of a linear model. Note that it leads to the detection of a linear model in 17 cases over 30. If I compare such linear models to the AR models estimated by OLS, I can notice that they are much more parsimonious. The

AR approach detected even 10 lags in one case, while SA never detected more than 4 lags. Likewise, if I compare the nonlinear models estimated by the SA approach to the ICP and CVP methods (the BRP's specification is fixed), it is seen from table 2 that SA is the most parsimonious. The question now is if this fact will influence the performances of the SA approach either in the sample fit (table 3) or in the forecasting performances (tables 4–11) or in both of them.

As to the in-sample fit, table 3 shows the $nRMSE$ of the estimated models. SA models always have the largest $nRMSE$, but this is expected because the corresponding estimated models are the most parsimonious. As a whole, the BRP approach performs better than the others in 11 cases according to the $nRMSE$ criterion, but this is not surprising because it is the less parsimonious one, with 12 lags and 4 hidden units in each estimated model: including more regressors gives, on the average, a better fit, but does not guarantee better forecasts. The ICP has the lowest residual mean square error 7 times, and the CVP and the AR follow, both with 6. Whether or not the larger models forecast better than the more parsimonious ones will be discussed in the next paragraph.

5.4 Forecasting Results

Forecasting with nonlinear models is far more complicated than in the linear case, especially in the case of multi-steps ahead forecasts. Techniques such as Monte-Carlo and Bootstrap should be used. See Granger and Teräsvirta (1993), p. 130-135 for a discussion. In this paper, I use the Monte-Carlo approach with 500 replications to produce multi-step forecasts from the nonlinear estimated models. Let $\mathbf{y}_t = (y_t, y_{t-1}, \dots, y_{t-k})'$ the set of lags entering the estimated ANN function $g(\cdot)$ at time $t + 1$, and \hat{y}_{t+1} the forecast one steps ahead, its optimal value is computed as follows:

$$\hat{y}_{t+1} = E[y_{t+1}|I_t] = g(\mathbf{y}_t)$$

For $h > 1$, it still holds that $E[y_{t+2}|I_t] = g(\mathbf{y}_{t+1}), \dots, E[y_{t+h}|I_t] = g(\mathbf{y}_{t+h-1})$, but we still base our forecast on the information set at time t , I_t . It follows that f.i. in vector \mathbf{y}_{t+1} , we do not know y_{t+1} . So forth so on, in vector \mathbf{y}_{t+h-1} we do not know $y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1}$, and they have to be generated by taking their variances into account. The most sensible assumption is that $y_{t+1}, \dots, y_{t+h-1}$ follow the same process than y_t , which we assume to be normal: $y_t \sim N(\mu, \sigma^2)$. It follows that we can generate forecasts at time $h > 1$ by using random numbers drawn by the same distribution, $z^{(j)} \sim N(0, \sigma^2)$. Therefore,

$$\begin{aligned} \hat{y}_{t+2} &= g(\mathbf{y}_{t+1}) = \frac{1}{500} \sum_{j=1}^{500} g\left(g(\mathbf{y}_t) + z_{t+2}^{(j)}\right) \\ \hat{y}_{t+3} &= g(\mathbf{y}_{t+2}) = \frac{1}{500} \sum_{j=1}^{500} g\left(g\left(g(\mathbf{y}_t) + z_{t+2}^{(j)}\right) + z_{t+3}^{(j)}\right) \\ &\dots \\ \hat{y}_{t+h} &= g(\mathbf{y}_{t+h-1}) = \frac{1}{500} \sum_{j=1}^{500} g\left(g\left(\dots\left(g(\mathbf{y}_t) + z_{t+2}^{(j)}\right)\dots\right) + z_{t+h}^{(j)}\right) \end{aligned}$$

where the sequences of random numbers $\left\{z_{t+2}^{(j)}, \dots, z_{t+h}^{(j)}\right\}$, $j = 1, \dots, 500$, are drawn from the normal distribution, $z_{t+i}^{(j)} \sim N(0, \sigma^2)$, and are mutually independent for all $i = 2, \dots, h$;

$j = 1, \dots, 500$. Results are reported in tables 4–15 and summarized in tables 4–10. In order to examine the forecasting performances for the 5 different models at the 3 horizons $h = 1, 6$ and 12 , first I compute a set of indexes which allows to make a comparison in absolute value for nRMSE, MAE (table 12) and MAD and the confusion rate CR (table 13). Albeit this approach does not constitute any sort of valid statistical test, it helps to shed light on the relative merits of the models. Results from tables 12 and 13 are summarized in tables 4 and 5, where the number of times each model performs best are demonstrated for each h and each of the 4 criteria.

The hypothesis of equal accuracy of forecasting performances measured by the nRMSE criterion as a function of the forecasting error is tested, by means of the modified Diebold-Mariano test statistic S^* . Results are reported in table 14. The number of times each modelling strategy performs better versus another one is reported in table 6, 7 and 8 for $h = 1, 6$ and 12 respectively. Likewise, in table 15, I use S^* based on the MAE criterion. Tables 9, 10 and 11 summarize the results for the 3 different forecasting horizons.

Let's take the first series, the first difference of the logarithms of the monthly US M1 money stock, seasonally adjusted, for the period 1959.1-1999.7. As can be seen in table 2, the sets of selected lags widely differ depending on the different models. The number of hidden units is 1 for all the NN models but BRP which keeps them fixed to 4 in advance. From tables 12 and 13, first line for model 1 corresponding to $h = 1$, the AR model ranks first as to the nRMSE criterion, MAE and MAD are the same for all the models, while BRP performs best as to the CR criterion. As to $h=6$ (second line), the BRP performs best as to the nRMSE, MAE and MAD are still the same for all the models, and CVP is the best according to CR. At $h=12$ (third line), AR has the lowest nRMSE, MAE and MAD the same for all the models, BRP has the lowest CR. Summarizing, for series 1 the BRP modelling strategy and the AR are the best for $h=1$, BRP and CVP for $h=6$, and BRP and AR at $h = 12$.

In order to perform the analysis by using a statistically tool, the corrected Diebold-Mariano test statistic S^* , I look into table 14 first (nRMSE). For $h=1$, AR and BRP are still the models performing best. AR turns out to be significantly better than ICP at the 5% level and it is better than both SA and CVP slightly over the 10% level, while it cannot be considered more accurate than BRP. BRP is better than SA at the 10% level, than ICP at the 1%, than CVP at the 5%. Now, performing the same analysis for $h = 1$ in table 15, the corresponding outcomes for MAE corroborates results in table 14. Proceeding now to the forecasting horizon $h = 6$, BRP outperforms 2 models and so does AR (nRMSE), while (MAE) BRP turns out to be better over 2 models and AR over 1. At $h = 12$, neither in table 14 or 15 is there a clear indication of the models performing best. Such outcome is expected, because forecasting accuracy obviously diminishes as h increases, and at $h = 12$ (one year ahead in the M1 series) it becomes difficult, if not impossible, to discriminate among the forecasting performances.

Such analysis is performed for each of the 30 series. A synthetic picture of the forecasting performances can be viewed in table 4 and 5. In the former, the AR methodology outperforms the other 5 ones both for $h = 1, 6$ and 12 , both for the nRMSE criterion (11 ranks best for $h = 1$; 10 for $h = 6$; 9 for $h = 12$) and for the MAE (11 times for $h = 1$; 10 for $h = 6$; 11 for $h = 12$). In the same table, neither of the 4 NN clearly outperforms any of the other. In the latter, the AR methodology loses its predominance both as to the MAD and CR criteria, but no clear dominant model can be observed.

If I now take tables 6-12, which summarize tables 14 and 15, I can find the confirmation of the outcomes of table 4. Looking at the column "sum wins - sum losses", for table 6

(nRMSE criterion, $h = 1$), I can see that the AR model performs best 24 times versus the other models as a whole. In table 9 (MAE, $h = 1$), the corresponding figure is 20. The AR methodology thus seems superior to the other 2 ones at the forecasting horizon $h=1$. As to other models, the ICP performs worse both in table 6 (-22) and table 9 (-18), while SA (+2 and +6) ranks second. As to $h=6$ (tables 7 and 10), results are less clear-cut. BRP is ranked first as to the nRMSE criterion (+16 vs. +15 for AR) but results change for the MAE (+5 vs +15 for AR). ICP is still the worst with -20 and -16, while SA performs quite badly too (-10 and -6). At $h=12$, it is still clear that AR is the best (+15 for nRMSE and +19 for MAE), but results for the other models are quite difficult to interpret, being all negative or slightly positive (ICP = -2 and -7; CVP = -14 and -14; BRP = -1 and +1; SA = -5 and +1). Therefore, from the analysis of table 6-11 I can conclude that, according to the nRMSE and MAE criteria and on the forecasting comparison based on the DM test on nRMSE and MAE, the AR methodology turns out to outperform all the other ones. Among those methods for NN modeling, none can be considered the best.

Going to table 5, where results for MAD (median absolute deviation) and CR (confusion rate, based on the comparison of the signs of $(y_{t+h} - y_t)$ with the sign of $(\hat{y}_{t+h} - y_t)$) are summarized, I can notice that the predominance of the AR models disappears. This may be due to the fact that both MAD and the CR criteria are less sensitive to outliers than nRMSE and MAE. At $h = 1$, AR ranks last as to MAD together with BRP, and last together with SA as to CR. AR turns out to be the best as to MAD and around average as to CR at $h = 6$, the second worst in MAD and the second best in CR at $h = 12$. As to SA, at $h = 1$ it ranks second best as to MAD and the worst as to CR. Because the CR criterion is a measure for the accuracy in forecasting the sign, and MAD is less sensitive to large forecasting errors than nRMSE and MAE, based on the median of the data and not on the mean, such 2 criteria are more useful to assess the goodness in forecasting the direction of the value of a stochastic process. In such view, the NN models clearly outperform the AR ones in forecasting one period ahead.

It is also of interest to see how the statistical methodology SA performs in the case of financial series. In this forecasting exercise, I included 3 monthly series of exchange rates: Canadian \$ / US \$ (series 9), German mark / US \$ (series 10), and Japanese yen / US \$ (series 11). For such series, the SA approach leads to the acceptance of linearity (no neural network model for any). The corresponding linear autoregressive model selected by the SBIC criterion is based on the first lag (see section 3.2 in this essay or section 3 in essay I for an extensive discussion). All the other approaches produce models which are much less parsimonious for each of the 3 series (see table 2). Furthermore, ranking the models according to their in-sample fit measured by the nRMSE criterion (see table 3), SA is ranked last for series 9,10 and 11. Now, if I look at the comparison for the forecasting performances for series 9, $h = 1$, corresponding to one month ahead, (table 14 and 15), it is clear that the hypothesis of equal accuracy between SA and ICP cannot be rejected either for nRMSE or MAE. This result holds for SA vs CVP, SA vs BRP, and SA vs AR as well. Series 10 and 11 perform similarly at $h = 1$. In this exercise, the SA methodology leads to parsimonious but very efficient models for financial forecasting.

6 Conclusion

In order to draw some conclusion on the overall exercise, I summarize the analysis of the 3 monthly financial series first. As to those 3 special cases, I can conclude that 1) the well-known property of parsimonious models of being the best forecasters is confirmed, 2) the

SA approach is able to detect a parsimonious model which performs better than the more complicated ones at the forecasting horizon $h = 1$ and 3) even when SA does not reject linearity and leads to the estimation of a linear model, such model is still competitive to the one estimated by the the AR approximation to Box & Jenkins' methodology. More generally, results from table 4 and 5 do not disagree with Swanson and White in (Swanson and White (1997)). In their analysis, the authors found out that each model performs best in at least one case and that there is no clear "winner". Likewise, what happens in this forecasting exercise is the following: while the linear model can be considered the best as the nRMSE and MAE criteria (table 4), it cannot definitely outperform the other models as to the MAD and CR criterion.

An overall conclusion about the results of the SA modelling procedure in this forecasting exercise is that, albeit it does not perform better than the other approaches, it produces similar forecasting performances with the advantage of leading to the estimation of more parsimonious models.

References

- AKAIKE, H. (1974): "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19, 716–723.
- ANDERS, U., AND O. KORN (1999): "Model Selection in Neural Networks," *Neural Networks*, 12, 309–323.
- BARTLETT, P. L. (1997): "For Valid Generalization, The Size of The Weights is More Important Than The Size of The Network," in *Advances in Neural Information Processing Systems 9*, ed. by M. C. Mozer, M. I. Jordan, and T. Petsche, pp. 134–140. The MIT Press.
- BERTSEKAS, D. P. (1995): *Nonlinear Programming*. Athena Scientific, Belmont, MA.
- DIEBOLD, F. X., AND R. S. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.
- EITRHEIM, Ø., AND T. TERÄSVIRTA (1996): "Testing the Adequacy of Smooth Transition Autoregressive Models," *Journal of Econometrics*, 74, 59–75.
- FORESEE, F. D., AND M. . T. HAGAN (1997): "Gauss-Newton Approximation to Bayesian Regularization," in *IEEE International Conference on Neural Networks (Vol. 3)*, pp. 1930–1935, New York. IEEE.
- GEMAN, S., E. BIENENSTOCK, AND R. DOURSAT (1992): "Neural Networks and the Bias/Variance Dilemma," *Neural Computation*, 4, 1–58.
- GOUTTE, C. (1997): "Note on Free Lunches and Cross-Validation," *Neural Computation*, 9, 1211–1215.
- GRANGER, C. W. J., AND T. TERÄSVIRTA (1993): *Modelling Nonlinear Economic Relationships*. Oxford University Press, Oxford.
- HARVEY, D., S. LEYBOURNE, AND P. NEWBOLD (1997a): "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13, 281–291.

- HARVEY, D., S. LEYBOURNE, AND P. NEWBOLD (1997b): “Testing the Equality of Prediction Mean Squared Errors,” *International Journal of Forecasting*, 13, 281–291.
- HORNIK, K., M. STINCHCOMBE, AND H. WHITE (1989): “Multi-Layer Feedforward Networks are Universal Approximators,” *Neural Networks*, 2, 359–66.
- HUSH, J. (1999): “Training a Sigmoidal Node is Hard,” *Neural Computation*, 11(5), 1249–1260.
- HWANG, J. T. G., AND A. A. DING (1997): “Prediction Intervals for Artificial Neural Networks,” *Journal of the American Statistical Association*, 92, 748–757.
- KAASHOEK, J. F., AND H. K. VAN DIJK (1998): “A Simple Strategy to Prune Neural Networks with an Application to Economic Time Series,” Econometric Institute Report 9854/A, Econometric Institute – Erasmus University.
- MACKAY, D. J. C. (1992): “Bayesian Interpolation,” *Neural Computation*, 4, 415–447.
- NGUYEN, D., AND B. WIDROW (1990): “Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights,” in *Proceedings of the International Joint Conference on Neural Networks*, vol. 3, pp. 21–26.
- RECH, G. (2001): “Modelling and Forecasting Economic Time Series with Single Hidden-Layer Feedforward Autoregressive Artificial Neural Networks,” Ph.D. thesis, Stockholm School of Economics.
- RECH, G., T. TERASVIRTA, AND R. TSCHERNIG (2001): “A Simple Variable Selection Method for Nonlinear Models,” *Communications in Statistics, Theory and Methods*, 30, 1227–1241.
- REED, R. (1993): “Pruning Algorithms – A Survey,” *IEEE Transactions on Neural Networks*, 4, 740–747.
- RISSANEN, J. (1978): “Modeling by Shortest Data Description,” *Automatica*, 14, 465–471.
- RIVALS, I., AND L. PERSONNAZ (1999): “On Cross-Validation for Model Selection,” *Neural Computation*, 11, 871–901.
- SCHWARZ, G. (1978): “Estimating the Dimension of a Model,” *Annals of Statistics*, 6, 461–64.
- SIESTMA, J., AND R. J. F. DOW (1988): “Neural Net Pruning – Why and How,” in *IEEE International Conference on Neural Networks*, vol. 1, pp. 325–333.
- (1991): “Creating Artificial Neural Networks that Generalize,” *Neural Networks*, 4(1), 67–69.
- STONE, M. (1977): “Asymptotics For and Against Cross-Validation,” *Biometrika*, 64, 29–35.
- SWANSON, N. R., AND H. WHITE (1997): “Forecasting Economic Time Series Using Flexible versus Fixed Specification and Linear versus Nonlinear Econometric Models,” *International Journal of Forecasting*, 13, 439–461.
- TONG, H. (1990): *Non-linear Time Series: A Dynamical Systems Approach*, vol. 6 of *Oxford Statistical Science Series*. Oxford University Press, Oxford.

- VAN DIJK, D. (1999): “Smooth Transition Models: Extensions and Outlier Robust Inference,” Ph.D. thesis, Tinbergen Institute, Rotterdam, The Netherlands, www.few.eur.nl/few/people/djvandijk/thesis/thesis.htm.
- VAN DIJK, D., T. TERÄSVIRTA, AND P. H. FRANSES (2001): “Smooth Transition Autoregressive Models - A Survey of Recent Developments,” *Econometric Reviews*, forthcoming.
- WEIGEND, A. S., D. E. RUMELHART, AND B. A. HUBERMAN (1991): “Generalization by Weight-Elimination with Application to Forecasting,” in *Advances in Neural Information Processing Systems 9*, ed. by R. P. Lippmann, J. Moody, and D. S. Touretzky. Morgan Kaufmann.
- ZHU, H., AND R. ROHWER (1996): “No Free-Lunch for Cross-Validation,” *Neural Computation*, 8, 1421–1426.

Series	Description	Transformation	Period	T	t_0
1 - M1	US M1 money stock, seasonally adjusted	first diff. of log	1959.1–1999.7	486	365
2 - M2	US M2 money stock, seasonally adjusted	first diff. of log	1959.1–1999.7	486	365
3 - M3	US M3 money stock, seasonally adjusted	first diff. of log	1959.1–1999.7	486	365
4 - CUR	US civilian unemployment rate, seasonally adj.	first diff. of log	1959.1–1999.7	486	365
5 - IP	US industrial production, seasonally adjusted	first diff. of log	1959.1–1999.7	486	365
6 - CPI	US consumer price index, seasonally adjusted	first diff. of log	1959.1–1999.7	486	365
7 - PPI	US producer price index, seasonally adjusted	first diff. of log	1959.1–1999.7	486	365
8 - EPU	US electric power use, seasonally adjusted	first diff. of log	1972.1–1999.6	329	240
9 - EXCAUS	Exchange rate: Canadian \$ to one US \$	first diff. of log	1871.1–1999.8	331	240
10 - EXGEUS	Exchange rate: German marks to one US \$	first diff. of log	1871.1–1999.8	331	240
11 - EXJPUS	Exchange rate: Japanese yens to one US \$	first diff. of log	1871.1–1999.8	331	240
12 - SP500	S&P 500 monthly close	first diff. of log	1950.1–1999.9	596	480
13 - DJIA	Dow Jones Industrial Average monthly close	first diff. of log	1950.1–1999.9	596	480
14 - DVI	Volcanic dust veil index, northern hemisphere	$\log_a(\text{obs.} + 1)$	1500–1969	470	300
15 - ASKEW1	M.ly riverflow: Sacramento R. at Keswick, CA	12th diff. of the log	1939.10–1960.7	252	170
16 - ASKEW2	M.ly r.flow: Madison R. near Yellowstone, MT	12th diff. of the log	1923.1–1960.1	444	300
17 - FISHER	Mean daily flow, cms, Fisher River, Dallas	logarithm	01.01.1988 – 31.12.1991	1461	1000
18 - OLDMAN	Mean daily flow, cms, Oldman River, Brocket	logarithm	01.01.1988 – 31.12.1991	1461	1000
19 - SUNSP	Annual sunspot	$2 \left[\sqrt{(1 + \text{obse.})} - 1 \right]$	1700–1998	299	200
20 - DEATHBF	Deaths in adult population of sheep blow-flies	logarithm	–	318	200
21 - TOTALBF	Total population of sheep blow-flies	logarithm	–	319	200
22 - AZUSA	Ozone concentration, AZUSA	logarithm	1956.1 – 1970.12	180	100
23 - OZONELA	Ozone concentraion, downtown L.A.	logarithm	1955.1 – 1972.12	216	100
24 - TPMON	Monthly temperatures in England	12th difference	1723.1–1970.12	2964	2000
25 - SUMMER	Mean summer temperature in Munich	logarithm	1781 – 1988	208	100
26 - GLOBTP	Changes in global temperature (annual)	first difference	1880 – 1985	105	70
27 - LYNX	Annual n. of Lynx trapped, MacKenzie River	logarithm base 10	1821 – 1934	114	70
28 - CHICKNY	Rept. n. of cases of Chickenpox, m.thly, NYC	12th diff. of the log	1931.1 – 1972.12	486	300
29 - MEASL	Rept. n. of cases of Measles, monthly, NYC	12th diff. of the log	1928.1 – 1972.12	522	300
30 - MUMPS	Rept. n. of cases of Mumps, monthly, NYC	12th diff. of the log	1928.1 – 1972.12	522	300

Table 1: Data sets.

Series	ICP		CVP		BRP		SA		AR
	Lags	q	Lags	q	Lags	q	Lags	q	AR Lags
1	1,3,9,12	1	1-3,6-12	1	1-12	4	1,6	1	1,3,5
2	1-3,9,10	1	1-3,5,7-10	1	1-12	4	1-3,9	0	1-3
3	1,3	1	1,2,4,5,7,9-12	1	1-12	4	1,6	1	1-4
4	1-6,8,10,12	1	1,2,4-6,8,9,11,12	1	1-12	4	1,4,12	1	2-4,12
5	1,5-7,9,11,12	1	2-12	1	1-12	4	1,2,5,12	2	1
6	1,8,9	1	1-12	1	1-12	4	1,10	0	1,5,9
7	2-4,7,11,12	1	1-9,11,12	1	1-12	4	2,4,12	1	1-3,5,6
8	7,9-12	1	2-7,9-11	1	1-12	4	1	0	1,6,12,24
9	1,3,8-10	1	1-8,10-12	1	1-12	4	1	0	1-3,8,10,13
10	1,4,6-11	1	1-9,11,12	1	1-12	4	1	0	1
11	1-4	1	1-4,6,8-11	1	1-12	4	1	0	1-3
12	1-3,5-12	1	1-12	1	1-12	4	1,9,12	1	1,5,14
13	1,3,5-10	1	2,4-10,12	1	1-12	4	1,8	1	1,5,14
14	1,2,5,6	1	1-8,10-12	1	1-12	4	1,4,5	0	1,4,5
15	1,12	1	1-12	1	1-12	4	1,12	0	1-3,11-13
16	1,12	1	1-12	1	1-12	4	1,5,12	0	1,12,13
17	1,2,3	1	1-8,10-12	1	1-12	4	1,2,4	2	1-3
18	1,2,3	1	1-3,5-12	2	1-12	4	1,2	1	1-3
19	1,2,9	1	1-5,7-9,11	1	1-12	4	1,2,11	1	1,2,9
20	1,4,11	1	1,5,6,8,10,11	1	1-12	4	1,9,12	1	1,2,8,11
21	1,2,8	1	1,3-5,7,10-12	1	1-12	4	1,3,4,12	1	1,2,8
22	1,4,9,10,12	1	1-5,7-12	1	1-12	4	1,4,6,12	0	1,4,12
23	1,4,5,10,11	1	1,4-9,12	1	1-12	4	1,10,12	0	1,10,12
24	1,2,4	1	1-5,7-12	1	1-12	4	1,2,12	0	1,2,4,6,8,12,13,15,24,25
25	1-7,10-12	1	1-4,6,8-12	1	1-12	4	1,6	0	1
26	1-6,8-11	2	1-3,7-9,11,12	1	1-12	4	2	0	1-4
27	1-12	4	1,4,5,10,12	1	1-12	4	1,3	1	1,2,4,7
28	1,9	1	1-9,11,12	1	1-12	4	1,7,12	0	1,4,9,12,13,20,23-25
29	1,2,11	1	1-7,9,11	1	1-12	4	1,2,9,12	0	1,2,4,11-13,21,24,25
30	1,11	1	1,2,4-12	1	1-12	4	1,12	0	1,11-13,24,25

Table 2: Specification results.

Series	ICP	CVP	BRP	SA	AR
1	0.859	0.849	0.855	0.883	0.871
2	0.682	0.679	0.681	0.693	0.698
3	0.695	0.702	0.609	0.693	0.692
4	0.908	0.918	0.911	0.917	0.923
5	0.874	0.936	0.904	0.912	0.919
6	0.589	0.582	0.415	0.714	0.587
7	0.795	0.801	0.809	0.854	0.835
8	0.951	0.946	0.747	0.990	0.940
9	0.924	0.912	0.931	0.985	0.931
10	0.879	0.922	0.944	0.954	0.954
11	0.906	0.896	0.918	0.938	0.924
12	0.975	0.970	0.988	0.986	0.988
13	0.982	0.979	0.997	0.995	0.986
14	0.663	0.651	0.591	0.656	0.607
15	0.723	0.701	0.700	0.721	0.642
16	0.684	0.676	0.678	0.686	0.663
17	0.124	0.124	0.118	0.141	0.124
18	0.106	0.105	0.098	0.100	0.106
19	0.367	0.363	0.321	0.424	0.377
20	0.419	0.399	0.301	0.400	0.423
21	0.362	0.337	0.154	0.374	0.389
22	0.532	0.524	0.489	0.539	0.538
23	0.575	0.570	0.469	0.593	0.593
24	0.962	0.840	0.821	0.843	0.791
25	0.874	0.882	0.998	0.959	0.972
26	0.745	0.832	0.995	0.964	0.873
27	0.050	0.326	0.180	0.332	0.351
28	0.678	0.590	0.573	0.600	0.537
29	0.250	0.248	0.215	0.245	0.207
30	0.333	0.329	0.322	0.332	0.286

Table 3: $nRMSE$ of estimated models.

Horizon	ICP	CVP	BRP	SA	AR
1	2	4	7	6	11
	4	6	5	4	11
6	4	5	4	7	10
	5	4	6	5	10
12	5	3	5	8	9
	4	3	8	4	11

Table 4: Forecasting results: number of series where each model is the best model according to the $nRMSE$ (first line) and MAE (second line).

Horizon	ICP	CVP	BRP	SA	AR
1	5	8	5	7	5
	9	7	9	5	5
6	6	7	6	3	8
	4	11	10	7	7
12	6	4	6	9	5
	6	6	10	11	10

Table 5: Forecasting results: number of series where each model is the best model according to the MAD (first line) and CR (second line).

	ICP	CVP	BRP	SA	AR	sum of losses	sum of wins - sum of losses
ICP	–	6	11	8	11	36	-22
CVP	2	–	8	7	8	25	-4
BRP	4	6	–	5	8	23	1
SA	4	5	3	–	7	19	2
AR	2	4	2	2	–	10	24
sum of wins	12	21	24	22	34	–	–

Table 6: Forecasting results: number of series where model A (column) is better than model B (line) according to the modified Diebold-Mariano statistic at a 0.10 level for 1-step-ahead (nRMSE test).

	ICP	CVP	BRP	SA	AR	sum of losses	sum of wins - sum of losses
ICP	–	7	10	8	9	34	-20
CVP	4	–	8	6	7	25	-1
BRP	2	6	–	3	5	16	16
SA	7	7	8	–	8	30	-10
AR	1	4	6	3	–	14	15
sum of wins	14	24	32	20	29	–	–

Table 7: Forecasting results: number of series where model A (column) is better than model B (line) according to the modified Diebold-Mariano statistic at a 0.10 level for 6-steps-ahead (nRMSE test).

	ICP	CVP	BRP	SA	AR	sum of losses	sum of wins - sum of losses
ICP	–	6	8	6	9	29	-2
CVP	8	–	10	8	9	35	-14
BRP	7	5	–	6	10	28	-1
SA	9	7	6	–	10	32	-5
AR	3	3	3	4	–	13	15
sum of wins	27	21	27	24	38	–	–

Table 8: Forecasting results: number of series where model A (column) is better than model B (line) according to the modified Diebold-Mariano statistic at a 0.10 level for 12-steps-ahead (nRMSE test).

	ICP	CVP	BRP	SA	AR	sum of losses	sum of wins - sum of losses
ICP	–	6	9	7	10	32	-18
CVP	2	–	7	5	8	20	1
BRP	4	7	–	7	10	28	-7
SA	4	3	2	–	7	16	6
AR	4	5	3	3	–	15	20
sum of wins	14	21	21	22	35	–	–

Table 9: Forecasting results: number of series where model A (column) is better than model B (line) according to the modified Diebold-Mariano statistic at a 0.10 level for 1-step-ahead (MAE test).

	ICP	CVP	BRP	SA	AR	sum of losses	sum of wins - sum of losses
ICP	–	6	9	9	10	34	-16
CVP	3	–	8	7	7	25	-4
BRP	6	7	–	4	7	24	5
SA	7	5	8	–	8	28	-6
AR	2	3	5	2	–	12	20
sum of wins	18	21	29	22	32	–	–

Table 10: Forecasting results: number of series where model A (column) is better than model B (line) according to the modified Diebold-Mariano statistic at a 0.10 level for 6-steps-ahead (MAE test).

	ICP	CVP	BRP	SA	AR	sum of losses	sum of wins - sum of losses
ICP	–	6	8	6	10	30	-7
CVP	8	–	8	7	7	30	-14
BRP	7	4	–	6	8	25	1
SA	6	4	6	–	7	23	1
AR	2	2	4	5	–	13	19
sum of wins	23	16	26	24	32	–	–

Table 11: Forecasting results: number of series where model A (column) is better than model B (line) according to the modified Diebold-Mariano statistic at a 0.10 level for 12-steps-ahead (MAE test).

Table 12: Forecasting results (nRMSE and MAE): 1 (first line), 6 (second line) and 12 (third line) steps ahead.

Series	ICP		CVP		BRP		SA		AR	
	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE
1	0.919	0.003	0.893	0.003	0.848	0.003	0.878	0.003	0.834	0.003
	1.161	0.004	1.119	0.004	1.027	0.004	1.124	0.004	1.057	0.004
	1.374	0.005	1.397	0.005	1.277	0.005	1.290	0.005	1.226	0.005
2	0.686	0.002	0.686	0.002	0.676	0.002	0.640	0.002	0.666	0.002
	1.254	0.004	1.274	0.004	1.137	0.003	0.973	0.003	1.136	0.003
	1.531	0.004	1.589	0.005	1.397	0.004	1.124	0.003	1.281	0.004
3	0.726	0.002	0.750	0.002	1.337	0.003	0.722	0.002	0.711	0.002
	1.212	0.003	1.265	0.003	1.516	0.004	1.108	0.003	1.061	0.003
	1.448	0.004	1.554	0.004	1.824	0.005	1.311	0.003	1.330	0.003
4	0.792	0.020	0.797	0.019	0.777	0.019	0.789	0.019	0.775	0.019
	0.799	0.020	0.779	0.019	0.788	0.019	0.774	0.019	0.776	0.019
	0.813	0.020	0.797	0.020	0.799	0.020	0.791	0.020	0.795	0.020
5	0.583	0.004	0.602	0.004	0.559	0.004	0.513	0.004	0.513	0.004
	0.552	0.004	0.567	0.004	0.563	0.004	0.515	0.004	0.505	0.004
	0.551	0.004	0.563	0.004	0.549	0.004	0.522	0.004	0.494	0.004
6	0.780	0.153	0.755	0.144	1.056	0.187	0.863	0.171	0.781	0.153
	0.755	0.147	0.747	0.141	0.841	0.164	1.012	0.197	0.754	0.144
	0.849	0.166	0.854	0.166	1.102	0.199	1.154	0.228	0.851	0.166
7	0.788	0.003	0.710	0.003	0.709	0.003	1.065	0.005	0.748	0.003
	0.745	0.003	0.715	0.003	0.723	0.003	0.758	0.003	0.749	0.003
	0.798	0.003	0.804	0.003	0.761	0.003	0.735	0.003	0.707	0.003
8	0.583	0.006	0.674	0.007	0.627	0.007	0.623	0.006	0.585	0.006
	0.601	0.006	0.616	0.007	0.618	0.006	0.587	0.006	0.561	0.006
	0.602	0.006	0.613	0.007	0.693	0.007	0.590	0.006	0.551	0.006
9	1.109	0.009	1.079	0.009	1.006	0.008	1.045	0.008	1.008	0.008
	1.037	0.009	1.016	0.008	1.031	0.008	1.101	0.009	0.990	0.008
	1.158	0.009	1.164	0.009	1.151	0.009	1.182	0.009	1.108	0.009
10	1.013	0.023	0.916	0.021	0.904	0.008	0.924	0.021	0.922	0.021
	0.882	0.021	0.866	0.020	0.856	0.008	0.895	0.020	0.893	0.020
	0.849	0.019	0.857	0.020	0.840	0.009	0.868	0.019	0.866	0.019

continued on next page

continued from previous page

Series	ICP		CVP		BRP		SA		AR	
	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE
11	1.040	0.022	1.040	0.023	0.929	0.020	0.957	0.020	0.938	0.021
	1.021	0.022	0.998	0.022	0.958	0.021	0.990	0.021	0.987	0.021
	1.030	0.023	1.034	0.023	1.013	0.022	1.021	0.022	1.022	0.022
12	0.953	0.030	0.969	0.031	0.939	0.030	0.927	0.029	0.939	0.030
	0.934	0.030	0.945	0.031	0.936	0.030	0.928	0.029	0.932	0.030
	0.887	0.028	0.917	0.030	0.897	0.029	0.893	0.029	0.905	0.030
13	0.987	0.031	0.981	0.032	0.964	0.030	0.974	0.030	0.978	0.031
	0.979	0.031	0.968	0.031	0.952	0.030	0.966	0.030	0.965	0.031
	0.982	0.032	0.935	0.030	0.936	0.030	0.946	0.030	0.962	0.031
14	0.655	0.974	0.647	1.000	0.668	0.975	0.671	0.987	0.659	0.971
	1.020	2.094	0.980	2.066	0.981	1.960	0.989	2.010	0.990	2.098
	1.031	2.170	0.988	2.100	0.996	1.922	1.017	2.192	1.019	2.193
15	0.969	0.373	1.042	0.395	1.005	0.390	0.971	0.371	0.943	0.356
	1.102	0.444	1.162	0.476	1.121	0.455	1.056	0.435	1.099	0.427
	1.113	0.446	1.182	0.460	1.133	0.447	1.071	0.446	1.111	0.436
16	0.553	0.081	0.551	0.081	0.544	0.081	0.549	0.082	0.507	0.073
	0.690	0.116	0.688	0.117	0.699	0.119	0.683	0.116	0.651	0.107
	0.704	0.121	0.719	0.125	0.742	0.129	0.709	0.124	0.661	0.111
17	0.135	0.082	0.135	0.082	0.132	0.084	0.138	0.084	0.137	0.083
	0.490	0.352	0.493	0.345	0.539	0.376	0.576	0.526	0.507	0.356
	0.606	0.487	0.610	0.474	1.060	1.078	0.744	0.752	0.619	0.478
18	0.099	0.050	0.099	0.050	0.097	0.053	0.101	0.050	0.099	0.052
	0.328	0.227	0.324	0.228	0.338	0.270	0.334	0.242	0.328	0.238
	0.450	0.361	0.462	0.386	0.449	0.359	0.448	0.421	0.444	0.361
19	0.423	1.856	0.417	1.766	0.492	2.028	0.479	2.009	0.407	1.755
	0.775	3.215	0.739	3.029	0.980	3.746	1.162	4.209	0.690	2.803
	0.890	3.817	0.846	3.627	1.148	4.780	1.147	4.893	0.834	3.560
20	0.525	0.639	0.537	0.669	0.458	0.536	0.411	0.452	0.599	0.761

continued on next page

continued from previous page

Series	ICP		CVP		BRP		SA		AR	
	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE
	1.787	2.289	2.641	2.499	1.439	1.880	0.799	0.875	1.536	2.003
	0.532	0.573	2.970	3.651	0.944	1.103	0.803	0.839	0.913	1.122
	0.270	0.236	0.268	0.232	0.420	0.375	0.373	0.347	0.357	0.335
21	1.101	1.106	1.166	1.158	1.437	1.350	1.663	1.712	1.298	1.323
	1.043	1.053	1.107	1.113	0.612	0.543	1.157	1.165	0.944	0.938
	0.624	0.193	0.629	0.198	0.573	0.180	0.556	0.183	0.610	0.188
22	0.655	0.190	0.647	0.187	0.620	0.181	0.584	0.188	0.650	0.196
	0.553	0.171	0.545	0.168	0.511	0.155	0.486	0.160	0.753	0.180
	0.657	0.186	0.691	0.202	0.623	0.018	0.606	0.178	0.606	0.178
23	0.656	0.189	0.738	0.211	0.674	0.189	0.649	0.185	0.649	0.185
	0.653	0.190	0.736	0.207	0.637	0.179	0.664	0.192	0.664	0.192
	0.968	1.476	0.815	1.247	0.827	1.260	0.815	1.240	0.778	1.195
24	0.997	1.510	0.828	1.262	0.836	1.274	0.830	1.262	0.795	1.211
	1.006	1.524	0.828	1.263	0.830	1.266	0.830	1.263	0.794	1.210
	0.848	0.039	0.856	0.037	0.928	0.043	0.855	0.038	0.906	0.040
25	0.846	0.038	0.862	0.039	0.887	0.041	0.850	0.038	0.902	0.041
	0.870	0.040	0.874	0.040	0.898	0.041	0.880	0.040	0.913	0.041
	1.059	0.106	1.083	0.119	1.109	0.116	1.080	0.109	0.990	0.099
26	1.022	0.104	1.142	0.117	1.107	0.114	1.106	0.115	1.077	0.113
	1.532	0.158	1.368	0.144	1.339	0.141	1.328	0.141	1.333	0.141
	17.404	4.735	0.507	0.224	0.708	0.327	0.595	0.262	0.532	0.246
27	6.662	1.407	0.649	0.305	1.183	0.543	1.024	0.477	0.820	0.406
	9.621	2.302	0.746	0.332	0.955	0.444	0.886	0.415	0.758	0.342
	0.697	0.221	0.612	0.196	0.611	0.196	0.613	0.195	0.586	0.188
28	0.911	0.302	0.808	0.265	0.770	0.255	0.778	0.253	0.769	0.251
	0.886	0.294	0.768	0.249	0.759	0.246	0.747	0.242	0.759	0.243
	0.239	0.344	0.240	0.341	0.229	0.334	0.235	0.340	0.213	0.307
29	0.603	0.885	0.605	0.897	0.558	0.856	0.590	0.892	0.543	0.824
	0.667	0.982	0.652	0.970	0.582	0.887	0.634	0.967	0.622	0.933

continued on next page

continued from previous page

Series	ICP		CVP		BRP		SA		AR	
	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE	nRMSE	MAE
30	0.331	0.162	0.323	0.160	0.321	0.160	0.296	0.158	0.308	0.155
	0.508	0.285	0.617	0.314	0.588	0.298	0.560	0.303	0.569	0.290
	0.638	0.331	0.716	0.371	0.700	0.364	0.637	0.350	0.673	0.347

Table 13: Forecasting results (MAD and CR): 1 (first line), 6 (second line) and 12 (third line) steps ahead.

Series	ICP		CVP		BRP		SA		AR	
	MAD	CR	MAD	CR	MAD	CR	MAD	CR	MAD	CR
1	0.003	0.382	0.003	0.346	0.003	0.309	0.003	0.382	0.003	0.355
	0.004	0.318	0.004	0.282	0.003	0.336	0.004	0.336	0.003	0.355
	0.005	0.391	0.005	0.409	0.005	0.382	0.005	0.400	0.004	0.391
2	0.001	0.404	0.001	0.464	0.001	0.446	0.001	0.409	0.001	0.436
	0.002	0.436	0.002	0.446	0.002	0.418	0.002	0.391	0.002	0.427
	0.002	0.446	0.002	0.446	0.002	0.418	0.002	0.400	0.002	0.409
3	0.002	0.382	0.001	0.373	0.002	0.446	0.001	0.409	0.001	0.355
	0.002	0.373	0.002	0.391	0.003	0.446	0.002	0.391	0.002	0.346
	0.003	0.391	0.003	0.391	0.003	0.381	0.003	0.400	0.003	0.382
4	0.015	0.300	0.015	0.318	0.014	0.291	0.014	0.300	0.015	0.291
	0.014	0.418	0.014	0.409	0.015	0.418	0.014	0.373	0.015	0.391
	0.016	0.364	0.015	0.364	0.015	0.364	0.015	0.346	0.015	0.373
5	0.003	0.327	0.004	0.363	0.003	0.309	0.003	0.327	0.003	0.355
	0.003	0.200	0.003	0.246	0.003	0.191	0.003	0.282	0.003	0.200
	0.003	0.246	0.003	0.273	0.003	0.218	0.003	0.264	0.003	0.218
6	0.111	0.291	0.108	0.273	0.115	0.355	0.106	0.345	0.115	0.346
	0.112	0.346	0.108	0.346	0.123	0.400	0.111	0.361	0.107	0.355
	0.131	0.300	0.120	0.309	0.128	0.409	0.122	0.319	0.119	0.309

continued on next page

continued from previous page

Series	ICP		CVP		BRP		SA		AR	
	MAD	CR	MAD	CR	MAD	CR	MAD	CR	MAD	CR
7	0.002	0.336	0.002	0.309	0.002	0.327	0.002	0.455	0.002	0.373
	0.002	0.300	0.002	0.263	0.002	0.309	0.002	0.300	0.002	0.318
	0.002	0.364	0.002	0.382	0.002	0.355	0.002	0.364	0.002	0.336
8	0.006	0.192	0.006	0.256	0.006	0.308	0.005	0.218	0.006	0.218
	0.006	0.333	0.006	0.321	0.006	0.269	0.006	0.282	0.006	0.295
	0.006	0.284	0.005	0.333	0.006	0.372	0.006	0.321	0.005	0.282
9	0.008	0.250	0.007	0.225	0.007	0.200	0.006	0.213	0.007	0.263
	0.008	0.263	0.008	0.250	0.007	0.263	0.007	0.275	0.008	0.238
	0.007	0.213	0.008	0.213	0.008	0.225	0.008	0.225	0.008	0.213
10	0.019	0.400	0.018	0.325	0.016	0.375	0.006	0.350	0.017	0.350
	0.016	0.200	0.016	0.188	0.018	0.175	0.007	0.188	0.018	0.188
	0.015	0.200	0.016	0.213	0.016	0.225	0.008	0.250	0.016	0.250
11	0.018	0.363	0.018	0.325	0.017	0.300	0.016	0.350	0.016	0.363
	0.017	0.250	0.017	0.238	0.015	0.275	0.016	0.238	0.016	0.250
	0.016	0.313	0.018	0.300	0.016	0.300	0.017	0.263	0.017	0.288
12	0.022	0.229	0.025	0.267	0.022	0.286	0.023	0.257	0.023	0.276
	0.023	0.210	0.022	0.229	0.021	0.238	0.020	0.248	0.021	0.248
	0.022	0.295	0.022	0.295	0.022	0.295	0.021	0.286	0.023	0.295
13	0.017	0.238	0.022	0.219	0.023	0.210	0.023	0.210	0.023	0.219
	0.023	0.295	0.022	0.276	0.022	0.276	0.022	0.286	0.021	0.295
	0.022	0.267	0.022	0.248	0.022	0.248	0.022	0.248	0.023	0.267
14	0.458	0.503	0.318	0.554	0.448	0.547	0.412	0.506	0.416	0.509
	1.839	0.509	1.852	0.491	1.603	0.503	1.918	0.488	1.844	0.503
	2.391	0.516	2.283	0.503	1.794	0.516	2.077	0.494	2.605	0.503
15	0.205	0.254	0.214	0.324	0.223	0.282	0.219	0.282	0.196	0.268
	0.282	0.225	0.292	0.282	0.257	0.254	0.294	0.197	0.216	0.211
	0.260	0.183	0.219	0.197	0.250	0.169	0.304	0.183	0.229	0.225
16	0.048	0.301	0.044	0.308	0.045	0.293	0.047	0.293	0.042	0.256
	0.090	0.203	0.093	0.196	0.096	0.195	0.090	0.173	0.076	0.158
	0.090	0.203	0.092	0.218	0.094	0.248	0.092	0.196	0.084	0.165

continued on next page

continued from previous page

Series	ICP		CVP		BRP		SA		AR	
	MAD	CR	MAD	CR	MAD	CR	MAD	CR	MAD	CR
17	0.040	0.378	0.039	0.387	0.041	0.389	0.041	0.384	0.041	0.402
	0.178	0.447	0.180	0.420	0.191	0.442	0.185	0.533	0.188	0.427
	0.273	0.453	0.252	0.429	0.266	0.560	0.236	0.538	0.278	0.424
18	0.016	0.360	0.019	0.404	0.020	0.424	0.015	0.362	0.019	0.409
	0.109	0.398	0.119	0.409	0.113	0.447	0.123	0.422	0.122	0.462
	0.225	0.420	0.259	0.418	0.218	0.398	0.277	0.398	0.255	0.416
19	1.635	0.080	1.526	0.090	1.788	0.125	1.939	0.182	1.436	0.136
	1.889	1.890	1.890	0.068	2.628	0.091	3.037	0.125	2.019	0.068
	3.142	2.985	2.985	0.250	4.054	0.352	4.180	0.398	3.075	0.261
20	0.261	0.523	0.229	0.533	0.260	0.551	0.317	0.411	0.250	0.561
	0.005	0.439	1.235	0.402	0.454	0.430	0.694	0.336	0.589	0.439
	0.352	0.280	1.409	0.486	0.539	0.421	0.533	0.327	0.364	0.430
21	0.155	0.411	0.149	0.398	0.207	0.407	0.176	0.463	0.165	0.426
	0.306	0.435	0.339	0.444	0.617	0.398	0.402	0.472	0.385	0.435
	0.293	0.509	0.292	0.537	0.295	0.370	0.308	0.519	0.294	0.482
22	0.192	0.261	0.186	0.290	0.157	0.290	0.163	0.304	0.170	0.290
	0.186	0.044	0.159	0.044	0.169	0.044	0.159	0.058	0.175	0.058
	0.177	0.362	0.166	0.333	0.163	0.333	0.149	0.348	0.189	0.377
23	0.149	0.267	0.160	0.295	0.151	0.238	0.152	0.248	0.152	0.248
	0.147	0.038	0.164	0.057	0.136	0.028	0.161	0.057	0.161	0.057
	0.141	0.429	0.185	0.467	0.143	0.419	0.177	0.438	0.177	0.438
24	1.203	0.267	1.011	0.210	1.027	0.218	0.988	0.209	0.938	0.200
	1.195	0.258	1.037	0.208	1.053	0.212	1.017	0.208	0.979	0.198
	1.200	0.172	1.029	0.155	1.041	0.156	1.017	0.158	0.980	0.155
25	0.034	0.299	0.030	0.247	0.032	0.289	0.031	0.227	0.030	0.299
	0.033	0.330	0.033	0.320	0.033	0.299	0.034	0.351	0.032	0.309
	0.036	0.289	0.035	0.289	0.033	0.258	0.033	0.299	0.033	0.247
26	0.070	0.208	0.098	0.208	0.075	0.208	0.084	0.208	0.075	0.250

continued on next page

continued from previous page

Series	ICP		CVP		BRP		SA		AR	
	MAD	CR	MAD	CR	MAD	CR	MAD	CR	MAD	CR
27	0.099	0.333	0.072	0.375	0.080	0.292	0.083	0.292	0.073	0.292
	0.137	0.417	0.117	0.250	0.165	0.208	0.105	0.208	0.107	0.208
	0.367	0.242	0.231	0.091	0.288	0.182	0.216	0.242	0.251	0.152
28	0.579	0.273	0.249	0.121	0.470	0.242	0.291	0.152	0.346	0.152
	0.618	0.424	0.296	0.091	0.245	0.091	0.304	0.061	0.258	0.121
	0.169	0.354	0.143	0.274	0.146	0.286	0.138	0.286	0.139	0.286
29	0.240	0.229	0.220	0.189	0.213	0.183	0.206	0.189	0.195	0.200
	0.229	0.131	0.196	0.109	0.195	0.103	0.189	0.097	0.190	0.126
	0.302	0.228	0.292	0.228	0.280	0.213	0.308	0.218	0.257	0.232
30	0.725	0.071	0.722	0.062	0.786	0.076	0.798	0.062	0.765	0.114
	0.847	0.119	0.821	0.109	0.807	0.085	0.863	0.114	0.840	0.095
	0.132	0.308	0.135	0.313	0.125	0.303	0.123	0.322	0.136	0.294
30	0.224	0.147	0.225	0.147	0.234	0.161	0.222	0.151	0.199	0.119
	0.273	0.128	0.263	0.147	0.294	0.147	0.260	0.142	0.265	0.137

Table 14: forecasting results (S^* for the $nRMSE$ test): 1 (first line), 6 (second line) and 12 (third line) steps ahead. Note: the apex 1 denotes rejection of the null hypothesis of equal accuracy in forecast performance at the 0.10 significance level, 2 at 0.05, 3 at 0.01. nc = not computed.

Ser.	SA	SA	SA	SA	ICP	ICP	ICP	CVP	CVP	BRP
	vs ICP	vs CVP	vs BRP	vs AR	vs CVP	vs BRP	vs AR	vs BRP	vs AR	vs AR
1	-1.457	-0.384	1.846 ¹	1.629	2.572 ²	3.005 ³	2.253 ²	2.122 ²	1.493	0.261
	-1.316	0.390	2.309 ²	2.070 ²	1.910 ¹	2.369 ²	1.890 ¹	1.627	0.976	-1.790 ¹
	-1.417	-1.307	0.470	1.185	-0.699	1.076	1.333	1.020	1.265	1.596
2	-3.100 ³	-2.972 ³	-2.542 ²	-4.086 ³	-0.151	1.794 ¹	1.442	2.046 ²	1.401	0.382
	-4.268 ³	-4.396 ³	-3.651 ³	-3.909 ³	-6.615 ³	5.079 ³	4.875 ³	5.307 ³	5.081 ³	-3.255 ³
	-3.323 ³	-3.460 ³	-2.929 ³	-2.577 ²	-4.509 ³	4.183 ³	3.893 ³	4.281 ³	4.008 ³	3.522 ³
3	-0.333	-0.727	-5.325 ³	0.654	-0.627	-5.323 ³	2.155 ²	-5.286 ³	1.918 ¹	5.410 ³
	-3.357 ³	-3.315 ³	-3.645 ³	3.362 ³	-2.291 ²	-3.412 ³	3.539 ³	-3.499 ³	3.540 ³	3.773 ³

continued on next page

continued from previous page

Ser.	SA vs ICP	SA vs CVP	SA vs BRP	SA vs AR	ICP vs CVP	ICP vs BRP	ICP vs AR	CVP vs BRP	CVP vs AR	BRP vs AR
	-2.104 ²	-2.179 ²	-2.221 ²	-1.244	-2.315 ²	-2.250 ²	2.064 ²	-2.224 ²	2.169 ²	2.208 ²
4	0.291 -0.720 -0.323	0.108 0.351 0.157	0.943 -0.187 0.105	1.151 0.531 0.596	-0.448 3.845 ³ 2.561 ²	1.555 2.248 ² 1.965 ²	1.160 2.022 ² 0.666	1.703 ¹ -1.874 ¹ -0.479	1.543 0.271 0.107	0.183 0.940 0.189
5	-0.708 0.602 1.211	-1.340 0.006 2.400 ²	0.533 0.132 0.568	0.302 0.647 0.998	-0.491 -0.507 -0.793	0.819 -1.093 0.078	0.704 -0.282 0.683	1.947 ¹ 0.116 0.344	1.443 0.519 0.742	-0.456 0.756 0.597
6	2.017 ² 1.870 ¹ 2.445 ²	2.404 ² 1.875 ¹ 2.487 ²	-1.348 1.474 0.270	1.872 ¹ 1.960 ² 2.547 ²	1.352 0.437 -0.550	-1.872 ¹ nc -1.209	-0.121 0.117 -0.409	-2.036 ² nc -1.221	-2.448 ² -0.639 0.423	1.876 ¹ nc 1.216
7	4.861 ³ 0.744 -1.239	5.363 ³ 3.157 ³ -3.631 ³	5.914 ³ 2.799 ³ -1.140	5.246 ³ 0.358 1.789 ¹	1.275 1.596 -0.194	1.507 1.222 1.084	0.673 -0.361 2.106 ²	0.067 -0.873 6.494 ³	-1.799 ¹ -1.846 ¹ 5.183 ³	-2.905 ³ -2.138 ² 3.340 ³
8	1.213 -0.614 -1.250	-1.882 ² -2.180 ² -7.042 ³	-0.445 -1.020 -3.688 ³	1.025 0.535 1.769 ¹	-2.259 ² -0.528 -1.072	-1.263 -0.416 -3.206 ³	-0.274 0.678 5.359 ³	1.460 -0.026 -3.017 ³	2.046 ² 1.423 4.062 ³	1.058 1.096 5.412 ³
9	-1.539 1.881 ¹ 0.104	-0.890 1.641 -0.011	0.729 1.676 ¹ 1.704 ¹	0.452 1.581 2.019 ²	0.350 1.132 -0.314	1.838 ¹ 0.366 0.046	1.747 ¹ 1.204 0.556	1.319 -1.098 0.152	1.200 0.993 0.631	-0.170 1.223 1.675 ¹
10	-2.653 ³ -1.313 0.059	-0.386 1.710 ¹ -0.572	-0.053 1.690 ¹ 1.149	-0.829 -1.213 -1.436	2.583 ³ 2.024 ² -4.609 ³	2.727 ³ 2.990 ³ 0.540	2.652 ³ 1.305 -0.063	0.430 2.709 ³ 1.253	0.385 -1.711 ¹ 0.567	0.051 -1.692 ¹ -1.151
11	-2.001 ² -1.889 ¹ -0.600	-2.144 ² -1.229 -1.459	0.748 1.470 -0.660	0.349 -0.634 -0.908	-0.400 1.953 ¹ -0.585	2.835 ³ 2.993 ³ 0.450	2.620 ³ 2.150 ² 0.460	3.019 ³ 2.855 ³ 2.746 ³	2.765 ³ 1.335 1.729 ¹	-0.566 -2.318 ² 0.272
12	-1.245 -0.527 0.792	-1.723 ¹ -0.796 -2.802 ³	-1.186 -0.786 -1.977 ²	-1.176 -0.584 -3.232 ³	-0.527 -0.897 -2.490 ²	0.857 -0.320 -1.675 ¹	0.589 -0.014 -2.548 ²	1.379 0.704 3.240 ³	1.097 0.663 1.100	-0.220 0.252 -1.799 ¹
13	-0.980 -8.783 ³	-0.360 -0.276	0.892 2.681 ³	-0.695 -0.399	0.192 0.462	1.464 6.912 ³	0.381 0.880	0.595 0.705	0.053 0.041	-1.222 -1.199

continued on next page

continued from previous page

Ser.	SA vs ICP	SA vs CVP	SA vs BRP	SA vs AR	ICP vs CVP	ICP vs BRP	ICP vs AR	CVP vs BRP	CVP vs AR	BRP vs AR
	-4.620 ³	3.605 ³	2.840 ³	-2.169 ²	5.069 ³	4.259 ³	1.033	-0.462	-4.352 ³	-3.141 ³
14	0.258	1.101	-0.477	-1.106	0.643	-0.529	-0.265	-0.946	-1.111	0.473
	-0.770	0.512	0.238	-0.941	1.071	1.477	0.768	-0.019	-0.537	-0.249
	-0.295	1.083	0.299	-0.528	1.487	0.548	0.292	-0.158	-1.105	-0.303
15	0.448	-0.856	-0.387	0.797	-1.678 ¹	-0.878	0.468	1.181	1.769 ¹	1.272
	-2.721 ³	-0.999	-0.638	-0.713	-0.727	-0.284	-0.029	2.210 ²	0.672	0.252
	-0.764	-0.564	-0.352	-0.370	-0.522	-0.228	-0.026	1.100	0.394	0.144
16	-0.307	-0.234	0.949	2.173 ²	0.292	0.770	1.998 ²	0.915	2.134 ²	1.971 ²
	-0.763	-0.952	-0.871	1.748 ¹	0.305	-0.400	1.597	-0.646	1.695 ¹	1.655 ¹
	0.581	-0.656	-0.936	4.733 ³	-0.983	-1.099	2.627 ³	-1.119	2.632 ³	1.935 ¹
17	1.091	0.784	0.942	-0.431	-1.473	0.597	-1.417	0.721	-1.212	-1.064
	3.569 ³	2.646 ³	0.873	2.525 ²	-0.366	-1.493	-1.130	-1.327	-0.827	1.138
	3.058 ³	2.368 ²	-5.440 ³	2.325 ²	-0.268	-5.693 ³	-0.791	-5.393 ³	-0.561	5.821 ³
18	0.608	0.789	1.695 ¹	0.571	0.919	1.077	-0.231	0.857	-0.849	-1.073
	0.559	1.036	-0.143	0.727	1.048	-0.571	0.032	-0.919	-0.883	0.592
	2.214 ²	1.925 ¹	2.153 ²	1.739 ¹	-1.052	0.135	0.441	0.576	0.955	0.085
19	1.756 ¹	1.953 ¹	-0.307	2.316 ²	1.111	-2.095 ²	1.549	-2.366 ²	0.672	2.511 ²
	2.930 ³	3.031 ³	1.617	3.107 ³	3.102 ³	-1.613	3.927 ³	-1.889 ¹	2.987 ³	2.057 ²
	2.731 ³	2.652 ³	0.150	2.710 ³	1.874 ¹	-2.634 ³	2.222 ²	-2.572 ²	0.342	2.602 ³
20	-3.089 ³	-3.264 ³	-1.177	-5.053 ³	-0.908	3.777 ³	-5.480 ³	3.798 ³	-7.211 ³	-6.925 ³
	-4.475 ³	-3.366 ³	-5.093 ³	-4.869 ³	-1.541	2.416 ²	3.525 ³	2.012 ²	2.201 ²	-1.186
	2.200 ²	-2.943 ³	-5.521 ³	-1.506	-2.939 ³	-3.483 ³	-4.373 ³	2.875 ³	2.813 ³	0.612
21	7.554 ³	8.268 ³	-2.110 ²	1.722 ¹	0.720	-5.669 ³	-8.973 ³	-5.582 ³	-7.986 ³	2.956 ³
	5.718 ³	5.482 ³	3.251 ³	5.818 ³	-6.804 ³	-2.396 ²	-4.974 ³	-2.161 ²	-4.307 ³	1.238
	22.22 ³	nc	4.403 ³	6.764 ³	-10.293 ³	3.832 ³	5.132 ³	4.292 ³	7.260 ³	-3.555 ³
22	-1.084	-1.020	0.795	-1.283	-0.367	2.939 ³	0.605	3.770 ³	0.615	-1.139
	-1.037	-0.743	0.301	-2.051 ²	0.620	1.834 ¹	0.214	1.863 ¹	-0.111	-1.050
	-2.902 ³	-1.665 ¹	0.589	-4.045 ³	0.935	3.132 ³	-2.930 ³	4.724 ³	-2.223 ²	-3.005 ³
23	0.186	-0.920	0.930	nc	-0.921	1.093	1.402	2.050 ²	2.489 ²	0.661
	1.877 ¹	2.036 ²	2.070 ²	nc	-1.633	-0.983	0.092	2.027 ²	1.873 ¹	0.700

continued on next page

continued from previous page

Ser.	SA vs ICP	SA vs CVP	SA vs BRP	SA vs AR	ICP vs CVP	ICP vs BRP	ICP vs AR	CVP vs BRP	CVP vs AR	BRP vs AR
	2.456 ²	2.659 ³	5.573 ³	nc	-1.675 ¹	0.347	-0.304	5.886 ³	1.141	-0.458
24	-8.720 ³	0.425	-2.290 ²	3.748 ³	8.804 ³	8.146 ³	9.662 ³	-2.811 ³	3.790 ³	4.658 ³
	-8.200 ³	0.885	-1.462	4.156 ³	8.157 ³	7.942 ³	9.196 ³	-2.013 ²	4.016 ³	4.521 ³
	-9.287 ³	0.672	-0.158	4.397 ³	8.865 ³	9.221 ³	9.263 ³	-0.683	4.306 ³	4.245 ³
25	0.312	0.139	-1.816 ¹	-1.626	-0.203	-1.638	-1.108	-1.497	-0.956	1.003
	0.491	-0.248	-0.822	-0.885	-1.100	-0.895	-0.950	-0.577	-0.644	-1.809 ¹
	1.703 ¹	1.586	-0.597	-0.680	-0.313	-0.968	-1.016	-1.186	-1.226	-1.570
26	0.451	0.037	-0.417	1.418	-0.375	-0.761	0.570	-0.245	0.827	1.175
	nc	-0.283	1.325	1.524	-1.599	nc	nc	0.362	0.665	1.647
	nc	nc	-0.429	nc	nc	nc	nc	nc	nc	-0.216
27	-2.893 ³	1.697 ¹	-1.420	0.879	2.895 ³	2.892 ³	2.895 ³	-3.231 ³	-1.264	2.173 ²
	-1.092	1.669 ¹	-0.872	1.316	1.107	1.078	1.102	-3.830 ³	-1.854 ¹	2.057 ²
	-1.116	0.785	-0.287	0.617	1.123	1.115	1.122	-1.798 ¹	-0.719	3.598 ³
28	-3.341 ³	0.397	0.403	1.425	3.263 ³	3.442 ³	3.941 ³	0.139	1.245	1.150
	-2.617 ³	-2.177 ²	0.708	0.497	1.999 ²	2.721 ³	2.692 ³	2.279 ²	1.405	0.021
	-2.502 ²	-1.284	-0.601	-0.864	2.201 ²	2.553 ²	2.773 ³	1.802 ¹	0.311	-0.034
29	-1.094	-1.073	0.885	2.781 ³	-0.241	1.206	2.912 ³	1.349	2.965 ³	1.848 ¹
	-0.715	-1.152	0.907	1.438	-0.315	0.986	1.391	1.040	1.519	0.315
	-1.575	-1.129	1.941 ¹	0.427	2.213 ²	1.852 ¹	1.969 ²	1.781 ¹	1.070	-0.630
30	-2.588 ³	-1.132	-0.565	1.253	2.563 ²	2.086 ²	2.303 ²	0.459	1.586	1.479
	2.026 ²	-2.899 ³	0.609	1.121	-2.405 ²	-1.324	-0.074	1.114	1.628	0.596
	1.883 ¹	-2.345 ²	-0.936	0.528	-2.761 ³	-2.270 ²	-0.898	1.187	1.519	1.104

Table 15: forecasting results (S^* for the MAE): 1 (first line), 6 (second line) and 12 (third line) steps ahead. Note: the apex 1 denotes rejection of the null hypothesis of equal accuracy in forecast performance at the 0.10 significance level, 2 at 0.05, 3 at 0.01. nc = not computed.

Ser.	SA vs ICP	SA vs CVP	SA vs BRP	SA vs AR	ICP vs CVP	ICP vs BRP	ICP vs AR	CVP vs BRP	CVP vs AR	BRP vs AR
1	-0.924	-0.437	1.778 ¹	1.416	1.289	2.776 ³	1.958 ¹	2.624 ³	1.562	-0.390
	-0.457	0.762	2.097 ²	1.790 ¹	1.407	1.980 ²	1.332	1.117	0.414	-1.689 ¹
	-1.254	-1.184	0.465	1.121	-0.638	0.989	1.233	0.932	1.169	1.083
2	-4.270 ³	-4.084 ³	-3.223 ³	-3.650 ³	-0.513	1.866 ¹	2.297 ²	2.486 ²	2.321 ²	1.050
	-6.003 ³	-6.121 ³	-4.569 ³	-4.587 ³	-6.854 ³	8.212 ³	8.943 ³	8.349 ³	8.748 ³	-1.556
	-5.635 ³	-6.066 ³	-4.776 ³	-3.856 ³	-12.46 ³	7.548 ³	7.217 ³	8.492 ³	7.821 ³	6.659 ³
3	-0.452	-0.715	-6.519 ³	0.871	-0.318	-6.415 ³	2.823 ³	-6.479 ³	1.922 ¹	6.737 ³
	-3.357 ³	-3.592 ³	-5.053 ³	2.972 ³	-3.466 ³	-4.813 ³	3.693 ³	-4.672 ³	3.947 ³	5.102 ³
	-1.972 ¹	-2.086 ²	-2.533 ²	-0.574	-2.312 ²	-2.721 ³	2.270 ²	-2.713 ³	2.330 ²	2.599 ³
4	-0.190	-0.043	0.534	0.310	0.330	1.580	0.806	1.190	0.557	-0.551
	-1.307	-0.044	-0.664	-0.094	3.668 ³	2.530 ²	2.006 ²	-2.832 ³	-0.051	0.895
	-0.553	-0.077	-0.141	0.009	3.849 ³	3.239 ³	0.735	-0.769	0.111	0.203
5	0.217	-1.712 ¹	-0.353	-0.917	-1.708 ¹	-0.362	-0.380	1.934 ¹	1.606	-0.051
	1.330	0.280	0.931	1.252	-1.225	-0.503	-1.231	0.873	1.114	-0.249
	1.655 ¹	2.215 ²	0.918	1.270	-1.366	0.322	0.704	0.702	1.045	0.247
6	2.116 ²	3.108 ³	-0.832	2.012 ²	2.752 ³	-1.861 ¹	-0.065	-2.418 ²	-3.281 ³	1.874 ¹
	1.972 ²	2.273 ²	1.601	2.335 ²	1.217	-1.976 ²	0.625	-2.576 ³	-0.892	2.631 ³
	2.182 ²	2.358 ²	0.778	2.376 ²	0.144	-1.322	0.302	-1.395	0.331	1.363
7	7.061 ³	7.422 ³	7.526 ³	6.815 ³	0.949	0.690	-0.297	-0.920	-2.216 ²	-2.495 ³
	1.390	3.395 ³	2.230 ²	0.947	0.750	0.118	-0.707	-1.510	-3.003 ³	-4.721 ³
	-1.374	-4.803 ³	-1.483	4.417	-0.784	1.031	2.563 ³	34.01 ³	7.375 ³	3.927 ³
8	0.535	-1.099	-0.235	0.518	-1.278	-0.607	-0.093	0.918	1.158	0.564
	-0.720	-3.208 ³	-0.702	0.191	-0.661	-0.092	0.577	0.417	1.342	0.814
	-0.526	-5.681 ³	-5.168 ³	2.395 ²	-1.150	-2.723 ³	2.079 ²	-2.219 ²	3.417 ³	nc
9	-1.227	-0.760	0.935	0.170	0.275	1.732 ¹	1.194	1.364	0.860	-0.747
	0.652	0.869	1.923 ¹	1.133	1.078	1.467	1.064	0.163	0.425	0.206
	0.224	-0.034	0.890	1.229	-1.502	-0.098	0.219	0.167	0.504	0.738
10	-1.689 ¹	0.235	0.215	-0.711	2.178 ²	1.737 ¹	1.689 ¹	-0.177	-0.236	-0.217
	-0.868	nc	0.564	-1.288	0.689	1.062	0.864	0.637	nc	-0.571
	0.978	-0.403	1.374	-1.535	-2.224 ²	-0.450	-0.982	1.161	0.395	-1.375

continued on next page

continued from previous page

Ser.	SA vs ICP	SA vs CVP	SA vs BRP	SA vs AR	ICP vs CVP	ICP vs BRP	ICP vs AR	CVP vs BRP	CVP vs AR	BRP vs AR
11	-2.137 ²	-2.362 ²	-0.377	-0.762	-0.300	2.049 ²	1.786 ¹	2.285 ²	1.938 ¹	-0.987
	-2.080 ²	-1.397	0.385	-0.780	1.844 ¹	3.036 ³	2.304 ²	2.169 ²	1.466	-0.915
	-0.736	-0.947	-0.865	-1.201	-0.028	0.553	0.546	1.525	0.772	0.370
12	-1.267	-1.598	-1.260	-1.094	-0.547	0.760	0.605	1.183	0.983	-0.034
	-1.215	-1.648	-1.235	-1.199	-1.588	0.288	0.251	1.579 ¹	1.461	0.153
	1.207	-3.263 ³	-3.499 ³	-7.454 ³	-2.906 ³	-2.612 ³	-4.999 ³	3.207 ³	1.211	-2.333 ²
13	-1.068	-1.472	0.699	-1.718 ¹	-0.532	1.372	-0.156	1.814 ¹	0.411	-2.162 ²
	-12.46 ³	-2.285 ²	2.704 ³	-1.356	0.268	9.644 ³	0.655	3.815 ³	0.376	-3.567 ³
	-4.688 ³	1.452	3.200 ³	-3.384 ³	3.892 ³	4.402 ³	1.355	0.707	-4.169 ³	-4.958 ³
14	-0.100	-1.237	-0.096	1.705 ¹	-0.990	-0.018	0.109	0.651	1.248	0.102
	0.033	0.708	1.773 ¹	-1.195	0.335	2.522 ²	-0.042	2.186 ²	-0.735	-1.797 ¹
	0.219	1.310	1.453 ¹	-0.391	1.234	1.739 ¹	-0.226	1.406	-1.332	-1.460
15	-0.158	-0.972	-0.789	0.645	-1.138	-0.646	0.745	0.824	1.7592 ¹	1.432
	-1.676 ¹	-0.840	-0.479	0.488	-0.668	-0.276	1.124	2.594 ³	0.988	0.637
	-0.033	-0.155	-0.015	0.261	-0.176	-0.010	0.322	0.624	0.260	0.147
16	0.224	0.292	0.770	2.646 ³	0.056	0.242	2.271 ²	0.318	2.256 ²	2.319 ²
	-0.416	-0.462	-0.591	1.730 ¹	0.196	-0.305	1.703 ¹	-0.547	1.653 ¹	1.575 ¹
	3.258 ³	-0.108	-0.559	3.385 ³	-0.818	-0.910	2.515 ²	-0.942	2.300 ²	1.811 ¹
17	1.917 ¹	1.327	0.193	0.949	-1.257	-1.119	-1.866 ²	-0.791	-0.803	0.493
	7.816 ³	7.388 ³	5.815 ³	6.552 ³	1.112	-1.803 ¹	-0.431	-2.105 ²	-0.951	1.320
	5.985 ³	5.830 ³	-10.02 ³	5.330 ³	1.384	-8.006 ³	0.634	-7.852 ³	-0.400	7.779 ³
18	0.299	-0.256	-1.878 ¹	-1.206	-1.588	-2.365 ²	-5.773 ³	-1.830 ¹	-2.075 ²	0.940
	1.735 ¹	1.625	-1.972 ²	0.425	-0.438	-2.749 ³	-1.738 ¹	-2.670 ³	-1.240	2.055 ²
	2.312 ²	1.648	2.232 ²	1.566	-1.897 ¹	0.102	0.012	0.937	1.013	-0.077
19	1.244	1.833 ¹	-0.112	2.047 ²	2.233 ²	-1.240	1.560	-1.929 ¹	0.143	1.770 ¹
	2.941 ³	3.500 ³	1.124	3.666 ³	5.256 ³	-1.092	6.458 ³	-1.545	3.004 ³	1.793 ¹
	3.210 ³	3.255 ³	0.521	3.285 ³	1.916 ¹	-3.333 ³	2.111 ²	-3.255 ³	0.506	3.104 ³
20	-3.937 ³	-4.745 ³	-1.875 ¹	-6.533 ³	-1.772 ¹	4.810 ³	-6.592 ³	5.222 ³	-7.271 ³	-9.178 ³
	-5.866 ³	-4.815 ³	-5.989 ³	-6.255 ³	-1.274	2.037 ²	3.457 ³	1.775 ¹	2.110 ²	-0.970
	1.589	-4.875 ³	-15.12 ³	-2.651 ³	-4.213 ³	-3.187 ³	-5.170 ³	4.448 ³	3.870 ²	-0.213

continued on next page

continued from previous page

Ser.	SA vs ICP	SA vs CVP	SA vs BRP	SA vs AR	ICP vs CVP	ICP vs BRP	ICP vs AR	CVP vs BRP	CVP vs AR	BRP vs AR
21	8.268 ³	8.527 ³	-1.555	1.344	0.501	-7.057 ³	-8.869 ³	-6.829 ³	-7.560 ³	2.504 ²
	11.25 ³	10.12 ³	4.138 ³	10.78 ³	-7.263 ³	-2.064 ²	-7.493 ³	-1.674 ¹	-6.176 ³	0.287
	nc	nc	5.841 ³	nc	-10.33 ³	4.997 ³	23.61 ³	5.782 ³	nc	-4.215 ³
22	-1.250	-1.468	0.284	-1.547	-0.900	1.998 ²	0.652	3.102 ³	0.963	-0.770
	-0.251	0.141	0.759	-1.022	0.656	1.420	-0.608	0.761	-0.752	-1.493
	-2.919 ³	-1.397	0.893	-7.030 ³	0.916	3.354 ³	-1.588	3.427 ³	-1.456	-2.801 ³
23	0.568	-0.771	1.213	nc	-1.386	0.815	0.764	2.319 ²	2.185 ²	0.036
	1.633	1.332	2.420 ²	nc	-1.393	0.033	0.300	4.014 ³	1.640	0.300
	1.522	1.249	4.887 ³	nc	-1.018	0.797	-0.077	5.977 ³	0.793	-0.559
24	-9.171 ³	-1.942 ¹	-2.148 ²	2.782 ³	8.854 ³	8.154 ³	9.691 ³	-1.537	3.336 ³	3.685 ³
	-8.586 ³	0.018	-1.353	3.288 ³	8.825 ³	8.698 ³	9.249 ³	-1.480	3.228 ³	3.508 ³
	-9.384 ³	-0.137	-0.437	3.591 ³	9.152 ³	9.735 ³	9.157 ³	-0.354	3.628 ³	3.479 ³
25	-0.152	0.664	-2.014 ²	-1.407	0.801	-1.400	-0.603	-2.017 ²	-1.300	1.653 ¹
	-0.001	-0.634	-1.399	-1.473	-1.621	-1.127	-1.220	-0.825	-0.941	-2.099 ²
	0.556	0.000	-0.582	-0.667	0.511	-0.570	-0.643	-0.864	-0.918	-1.347
26	0.157	-0.699	-1.104	1.382	-0.902	-0.787	0.594	0.218	1.517	1.778 ¹
	7.448 ³	-0.136	1.178	1.229	-1.203	-3.267 ¹	nc	0.213	0.315	0.742
	nc	-0.689	0.711	0.145	nc	nc	nc	0.496	0.434	-0.767
27	-3.010 ³	1.536	-1.657 ¹	0.570	3.062 ³	2.960 ³	3.046 ³	-2.969 ³	-1.201	1.917 ¹
	-1.762 ¹	1.236	-0.586	0.771	1.861 ¹	1.492	1.820 ¹	-3.264 ³	-2.051 ²	1.765 ¹
	-1.552	0.746	-0.255	0.614	1.766 ¹	1.536	1.717 ¹	-1.532	-0.353	3.035 ³
28	-3.172 ³	-0.279	-0.148	1.009	2.834 ³	2.747 ³	3.374 ³	0.135	1.079	0.984
	-2.860 ³	-2.604 ³	-0.299	0.219	1.993 ¹	2.771 ³	2.763 ³	1.295	1.293	0.335
	-3.323 ³	-1.412	-0.614	-0.300	2.967 ³	3.389 ³	3.937 ³	0.867	0.598	0.255
29	-0.769	-0.191	0.522	2.324 ²	0.724	0.912	2.593 ³	0.710	2.440 ²	1.853 ¹
	0.313	-0.366	0.802	1.135	-1.563	0.626	0.896	0.832	1.131	0.289
	-0.577	-0.150	2.138 ²	0.505	1.250	1.632	0.961	1.708 ¹	0.646	-0.461
30	-1.565	-0.786	-0.762	0.534	1.224	0.736	1.258	-0.184	0.829	0.917
	1.869 ¹	-3.000 ³	0.366	0.919	-2.401 ²	-1.502	-0.310	1.062	1.726 ¹	0.557
	1.036	-2.482 ²	-1.641	0.234	-2.130 ²	-1.968 ²	-0.670	1.312	1.677 ¹	1.245

continued on next page

continued from previous page

Ser.	SA	SA	SA	SA	ICP	ICP	ICP	CVP	CVP	BRP
	vs	vs	vs	vs	vs	vs	vs	vs	vs	vs
	ICP	CVP	BRP	AR	CVP	BRP	AR	BRP	AR	AR
