Adaptation of model selecting criteria for nonlinear time series forecasting

Research practise 2: Proposal presentation

Andrea Molina-Alonso[1] Myladis Rocio Cogollo-Flórez[2]

[1]Responsible
amolin13@eafit.edu.co

[2]Tutor
mcogollo@eafit.edu.co

EAFIT University
Mathematical Engineering
Department of Mathematical Sciences

March 4, 2016

Time series forecasting

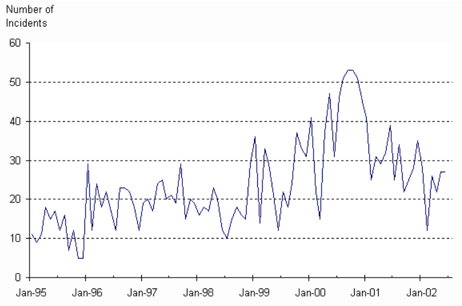

Figure: Time series plot of the monthly number of incidents of international piracy. Taken from *http : //www.rita.dot.gov*

The interest is to discover future values of a signal or a function of time, X_t, based on its past values [Camina and Janacek, 2012].

Modeling methodology of time series for forecasting

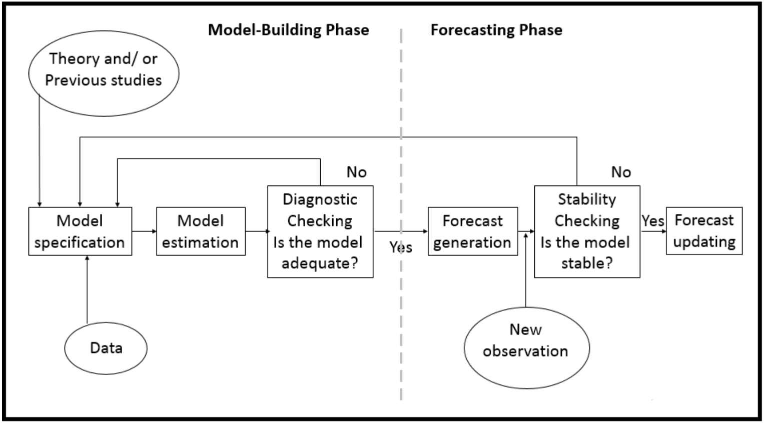

Figure: Conceptual framework of a forecasting system. Taken from [Abraham and Ledolter, 2009]

Types models of time series

A model of pth-order of time series is defined as [Li, 2003, Hwang et al.,]

$$X_t = f\left(F_{t-1};\phi\right) + a_t \quad , \tag{1}$$

where

- f is a known linear or nonlinear function of past X_t's.
- ϕ is a $p \times 1$ vector of parameters.
- The noise process $\{a_t\}$ is assumed to be independent, with mean zero, variance σ_a^2, and finite fourth order moment.

- Compared to the linear case, the nonlinear time series have been little explored and theory is not sufficient to uncover nonlinearities [Anders and Korn, 1999].
- One of the most critical issues is to select the appropriate forecasting nonlineal model [Qi and Zhang, 2001].

Model selection criterion	Definition	Disadvantages
SSE	$\sum_{i=1}^{T}(y_i - \hat{y}_i)^2$	It has overfitting problems. It is not invariant to linear scale transformations [De Gooijer and Hyndman, 2006] .
RMSE	$\sqrt{\frac{1}{T}SSE}$	It has overfitting and consistency problems [Aladag et al., 2010].
$\bar{R}^2$	$1 - \frac{SSE/(T-m)}{\sum(y_i - \hat{y}_i)^2/(T-1)}$	Inappropriate measure in the field of nonlinear fitting [De Gooijer and Kumar, 1992].

Table: Disadvantages of model selection criteria, where m is the number of parameters and T the number of observations

Model selection criterion	Definition	Disadvantages
AIC	$\log\left(\frac{SSE}{T}\right) + \frac{2m}{T}$	It requires normal data [Qi and Zhang, 2001]. Poor performance in nonlinear time series [Spiess and Neumeyer, 2010]. Presents a bad performance in nonlinear time series [De Gooijer and Hyndman, 2006].
AICC	$\log\left(\frac{SSE}{T}\right) + \frac{2m}{T-m-1}$	Presents a bad performance in nonlinear time series [De Gooijer and Hyndman, 2006].
BIC	$\log\left(\frac{SSE}{T}\right) + \frac{m\log(T)}{T}$	The penalty term can become dominant [Qi and Zhang, 2001].

Table: Disadvantages of model selection criteria, where m is the number of parameters and T the number of observations

Model selection criterion	Definition	Disadvantages
MAPE	$\frac{1}{T}\sum_{i=1}^{T}\left\lvert\frac{(y_i - \hat{y}_i)}{y_i}\right\rvert$	It has consistency problems [Aladag et al., 2010].
DA	$\frac{1}{T}\sum_{i=1}^{T} a_i$, where $a_1 = \begin{cases} 1 & \textit{if}\,(y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) > 0 \\ 0 & \textit{otherwise} \end{cases}$	Large model penalization [Egrioglu et al., 2008].
MDA	$\frac{\sum_{i=1}^{T-1} D_i}{T-1}$, where $D_i = (A_i - F_i)^2$	Large model penalization [Egrioglu et al., 2008].

Table: Disadvantages of model selection criteria, where m is the number of parameters and T the number of observations

Weighted selection criterion

A weighted selection criteria using optimization was proposed by [Aladag et al., 2010]:

$$AWIC = w_1 RMSE + w_2 MAPE + w_3(1 - DA) + w_4 MDA + 0.1AIC + 0.1BIC \quad (2)$$

- It is not shown a criteria for determining the weights of AIC and BIC.
- There are no guidelines to know which criteria to use, bearing in mind the inherent behavior of the time series.
- Heuristic methods have not been successful in the estimation of weights for the combined methods.
- This method does not consider the time series characteristics.

Importance and originality of this research

- A solution for the selection of models for nonlinear time series including multivariate techniques would be investigated.
- Heuristic methods have not been successful in the estimation of weights for the combined methods.
- The criteria are applied without validating the assumptions required, leading to bad decisions.
- It would be implemented a method that not only considers the time series characteristics, but the advantages of the existing criteria.

Objectives I

General Objective

Formulate a criterion for selecting models of nonlinear time series using multivariate analysis techniques and the inherent characteristics of the series.

Objectives II

Specific Objectives

- Identify the different selection criteria formulated in the literature for non-linear time series.
- Determine the multivariate analysis techniques that allow the creation of synthetic indicators according to the characteristics of the data.
- Establish a methodological framework that considers the characteristics of the data and consider the advantages of the proposed selection criteria in literature to date.
- Validate the feasibility of the proposed methodology by experimental data.

Schedule of Activities

Activity	Start	End
Review of literature	January 30	March 10
Proposal report	January 30	February 12
Oral presentation of the proposal report	February 12	February 19
Identification of the methods	February 19	March 10
Selection of the multivariate technique	March 10	April 1
Oral progress report	April 1	April 8
Method implementation	April 8	April 20
Validation with experimental data	April 20	May 5
Project report	April 10	May 20
Project presentation	May 20	June 7

Table: Schedule of Activities.

References I

Abraham, B. and Ledolter, J. (2009).
Statistical methods for forecasting, volume 234.
John Wiley & Sons.

Aladag, C. H., Egrioglu, E., Gunay, S., and Basaran, M. A. (2010).
Improving weighted information criterion by using optimization.
Journal of computational and applied mathematics, 233(10):2683–2687.

Anders, U. and Korn, O. (1999).
Model selection in neural networks.
Neural Networks, 12(2):309 – 323.

References II

Camina, A. R. and Janacek, J. (2012).
Mathematics for seismic data processing and interpretation.
Springer Science & Business Media.

De Gooijer, J. and Hyndman, R. (2006).
25 years of time series forecasting.
International Journal of Forecasting, 22(3):443–473.

De Gooijer, J. G. and Kumar, K. (1992).
Some recent developments in non-linear time series modelling, testing, and forecasting.
International Journal of Forecasting, 8(2):135–156.

References III

- Egrioglu, E., Aladag, C. H., and Gunay, S. (2008).
  A new model selection strategy in artificial neural networks.
  Applied Mathematics and Computation, 195(2):591–597.
- Hwang, S. Y., Basawa, I., and Reeves, J.
  The asymptotic distributions of residual autocorrelations and related tests of fit for a class of nonlinear time series models.
  Statistica Sinica, 4:107–125.
- Li, W. K. (2003).
  Diagnostic checks in time series.
  CRC Press.

References IV

Qi, M. and Zhang, G. P. (2001).
An investigation of model selection criteria for neural network time series forecasting.
European Journal of Operational Research, 132(3):666–680.

Spiess, A.-N. and Neumeyer, N. (2010).
An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach.
BMC pharmacology, 10(1):6.

Thanks for your attention!!

Adaptation of model selecting criteria for nonlinear time series forecasting.