Adaptation of model selecting criteria for nonlinear time series forecasting

Research practise 2: Progress 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

April 12, 2016

Modeling methodology of time series for forecasting

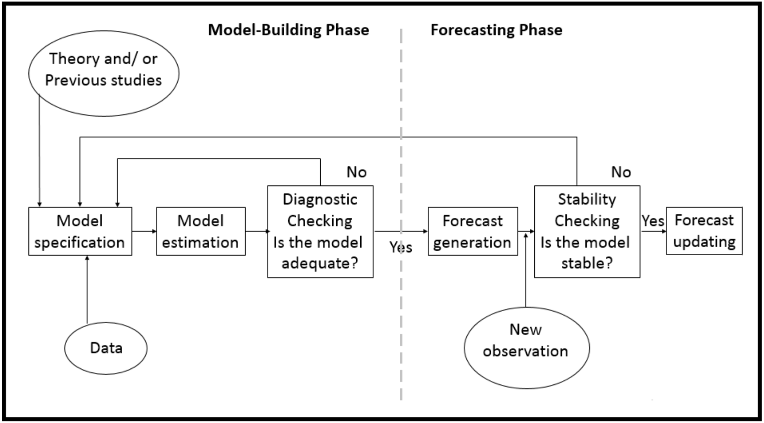


Figure: Conceptual framework of a forecasting system. Taken from [Abraham and Ledolter, 2009]

Types models of time series

A model of *p*th-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	$\frac{1}{T}\sum_{i=1}^{T}\left\|\frac{(y_i-\hat{y}_i)}{y_i}\right\|$	It has consistency problems [Aladag et al., 2010].
DA	$\frac{1}{T}\sum_{i=1}^{T}a_i,$ where $a_1 = \begin{cases} 1 & if\,(y_{i+1}-y_i)(\hat{y}_{i+1}-y_i) > 0 \\ 0 & otherwise \end{cases}$	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]:

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

- 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.

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.

Other model selection criteria

Model selection criterion	Definition
MAE	$\frac{1}{T}\sum_{i=1}^{T}\lvert(y_i - \hat{y}_i)\rvert$
Sign	$\frac{1}{T}\sum_{i=1}^{T} z_i,$ where $z_1 = \begin{cases} 1 & if\,(y_{i+1})(\hat{y}_{i+1}) > 0 \\ 0 & otherwise \end{cases}$
ME	$\frac{1}{T}\sum_{i=1}^{T}(y_i - \hat{y}_i)$

Table: Other model selection criteria, where T is the number of observations

Selected models I

Nonlinear Autoregressive Model (NAR) [Aras and Kocakoç, 2016]:

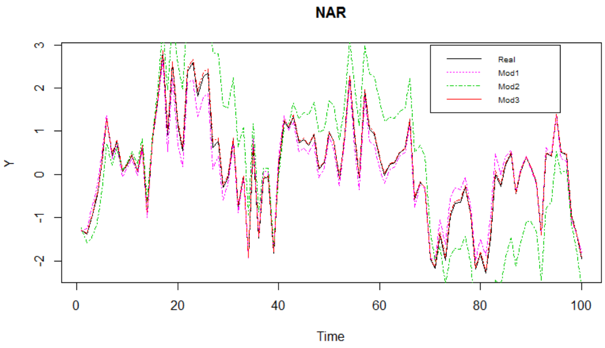


Figure: Variations of the model $y_t = 0.7y_{t-1} - 0.017y_{t-1}^2 + \varepsilon_t$

Selected models II

Generalized Autoregressive Conditional Heteroskedastic (GARCH) [Ennio and Pablo, 2011]:

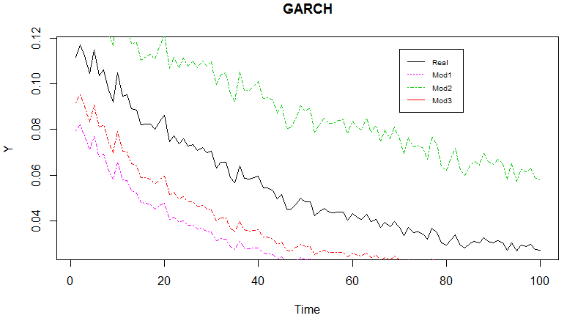


Figure: Variations of the model $y(t) = \sqrt{h_t \varepsilon_t}$ with $h_t^2 = 0.00002281 + 0.0593y_{t-1}^2 + 0.901h_{t-1}^2$

Selected models III

Autoregressive Model (AR):

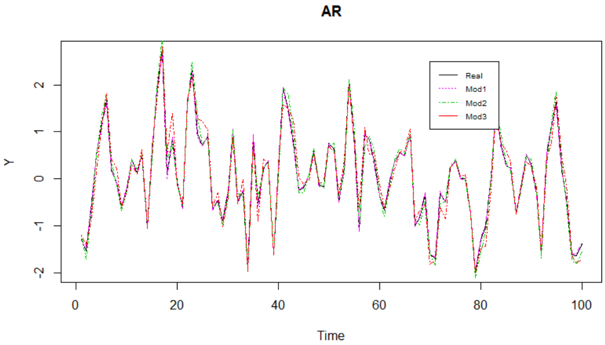
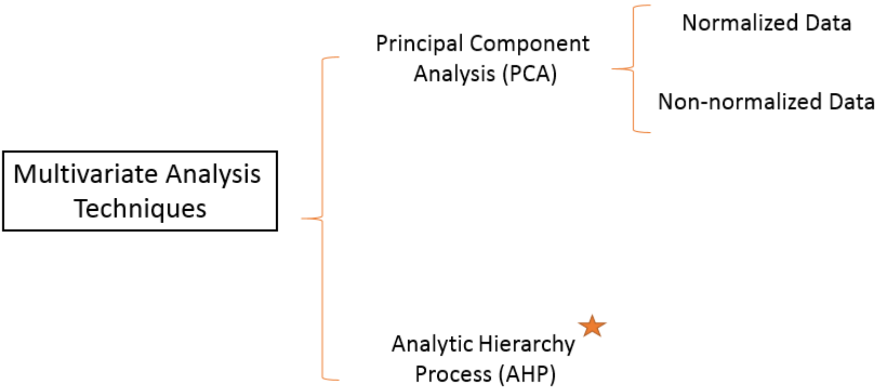


Figure: Variations of the model $y_t = 0.67 y_{t-1} - 0.41 * y_{t-2} \varepsilon_t$

Multivariate analysis techniques



Partial results of PCA I

After implementing the different criteria in the time series models and their variations, we obtained the following correlation matrix:

	SSE	RMSE	MAPE	MAE	ME	DA	MDA	sign	AIC	BIC	AICC
SSE	1.00	0.942	0.46	0.939	-0.10	-0.104	0.69	-0.85	0.574	0.574	0.574
RMSE	0.94	1.000	0.59	0.999	-0.23	-0.076	0.68	-0.89	0.766	0.766	0.766
MAPE	0.46	0.585	1.00	0.576	-0.82	-0.432	0.62	-0.19	0.498	0.498	0.498
MAE	0.94	0.999	0.58	1.000	-0.23	-0.087	0.67	-0.89	0.766	0.766	0.766
ME	-0.10	-0.233	-0.82	-0.234	1.00	0.354	-0.26	-0.13	-0.220	-0.220	-0.220
DA	-0.10	-0.076	-0.43	-0.087	0.35	1.000	-0.49	-0.15	0.067	0.067	0.067
MDA	0.69	0.684	0.62	0.672	-0.26	-0.491	1.00	-0.49	0.506	0.506	0.506
sign	-0.85	-0.889	-0.19	-0.891	-0.13	-0.145	-0.49	1.00	-0.685	-0.685	-0.685
AIC	0.57	0.766	0.50	0.766	-0.22	0.067	0.51	-0.69	1.000	1.000	1.000
BIC	0.57	0.766	0.50	0.766	-0.22	0.067	0.51	-0.69	1.000	1.000	1.000
AICC	0.57	0.766	0.50	0.766	-0.22	0.067	0.51	-0.69	1.000	1.000	1.000

Figure: Correlation matrix of model selection criteria

Partial results of PCA II

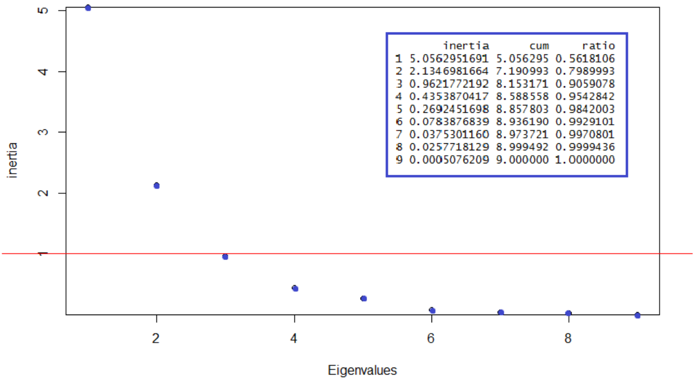


Figure: Inertia plot

Partial results of PCA III

	Comp1	Comp2	Comp3
SSE	-0.9333479	0.1186563	0.14185485
RMSE	-0.9791722	0.1491286	-0.05941615
MAPE	-0.5658536	-0.7180668	-0.28950029
MAE	-0.9796977	0.1513640	-0.05323925
ME	0.2715033	0.8309280	0.39172496
DA	0.1494437	0.6720636	-0.66085177
MDA	-0.7941954	-0.3111628	0.34918324
sign	0.8040604	-0.5325492	-0.15928829
AIC	-0.7570164	0.1931493	-0.33823676

Figure: Coordinates 1

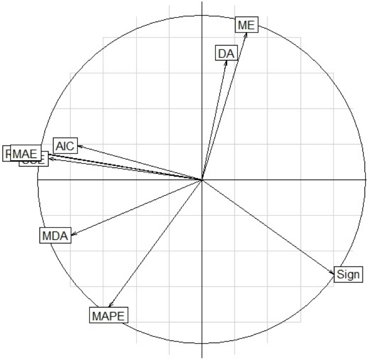
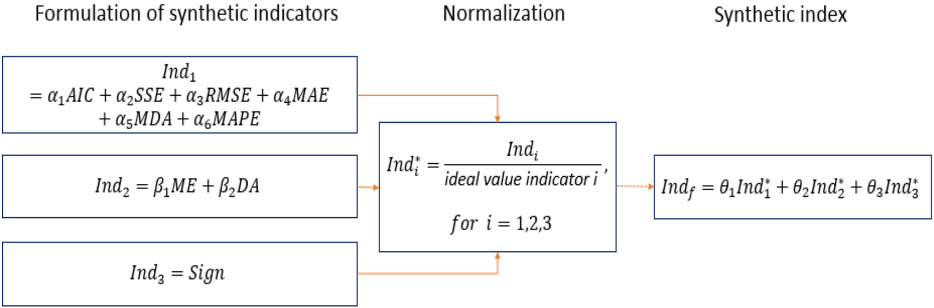


Figure: Coordinates 2

Methodological strategy using PCA



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

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

Aras, S. and Kocakoç, İ. D. (2016).
A new model selection strategy in time series forecasting with artificial neural networks: IHTS.
Neurocomputing, 174:974–987.

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

- 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.
- Ennio, A. P. and Pablo, O. (2011).
  A Switching Threshold Model for Oil Prices.
  Systems Engineering Procedia, 1:490–498.
- 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.

References

Li, W. K. (2003).
Diagnostic checks in time series.
CRC Press.

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.