Volume 26 • Number 2 • April 2026

Cilt 26 • Sayı 2 • Nisan 2026

Contents

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 26 • Sayı 2 • Nisan 2026
SS. 225/236
Doi: 10.21121/eab.20260205
Başvuru Tarihi: 29.01.2024 • Kabul Tarihi: 11.12.2025

Article Type: *Research Article*

Comparison of Model Selection Criteria for Models Including Trend and Seasonal Components in Econometric Time Series

Pınar GÖKTAŞ[1]

ABSTRACT

The paper aims to compare commonly used model selection criteria in time series modeling, such as Adjusted R2, log-likelihood, Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC), Hannan-Quinn (HQ) Information Criterion, and Mean Squared Error (MSE). In this context, for an additive time series, data was produced in different sample sizes from n=60 to n=500 accross 17 different stationary stochastic processes, including constant, trend, seasonal and irregular components. Each production was repeated 10000 times and the criteria were calculated.

For very large sample sizes, the HQ information criterion provides the best results for all types of time series models. It was observed that log-likelihood performed poorly in almost all models. It has been found that "Adjusted R2" is the best option for models with sample sizes less than 120, and "AIC" criterion is the best option for choosing the right model as the sample size increased. These findings offer practical guidance for researchers in econometrics, helping them select the most appropriate model selection criterion to reduce erroneous decisions and improve the reliability of time series analyses based on sample size and component characteristics.

Keywords: Stationary Stochastic Process; Model Selection Criteria; Data Generating Process, Monte Carlo Simulation.

JEL Classification Codes: C15, C32, C52

Referencing Style: APA 7

INTRODUCTION

One of the techniques with a broad range of applications in econometric analyses is time series analysis. The compilation of the majority of economic data in the form of time series plays a crucial role in the emergence of this situation within the economic framework.

Time series is a sequence of observations arranged in equal time intervals over time. It consists of various deterministic and stochastic (probabilistic) components such as trend (long-term trend), seasonality (seasonal fluctuations), periodic (cyclical fluctuations), and irregular movements (error term). The deterministic feature in time series is generally involve examining whether the time series has constant, trend, and seasonal components. Stochastic features, on the other hand, mainly involve investigating the stationarity of variables. Accurately identifying deterministic and stochastic properties in time series and making appropriate modeling is of great importance for reliable predictions.

If the objective is to analyze a time series then the analysis would be considered as the collective term for statistical investigations, applications, and approaches that including determination, forecasting, and testing accuracy with the aim of understanding the underlying structure of the observed data. The goal is to gain insights into the fundamental effects that give rise to the data, comprehend the structure underlying the observed data, and make reliable forecasts for the future.

In addition, it can be easily estimated that an initial analysis of each series is a prerequisite for successful empirical modeling (Franses, Hylleberg and Lee, 1995:249). Fundamental characteristics that need to be addressed in the data include extreme or missing values, sudden changes in the level of series, breakpoints, heteroskedastic structure, and so on. Furthermore, trends in the series over time, characterized by long-term increases and decreases, can be considered as linear and nonlinear or deterministic and stochastic trends. Similarly, regularly recurring peaks and troughs in relation to period of times such as daily, weekly, quarterly, and monthly patterns (seasonality) manifest in two forms: deterministic and stochastic. In this context, the preliminary analysis involves addressing significant features of the time series and accurately defining the structure of the variables under investigation.

[1] Dr., Türkiye İstatistik Kurumu, Ankara/TÜRKİYE, pinar.goktas@tuik.gov.tr

In econometrics of time series, methods and approaches vary significantly depending on the presence of deterministic and stochastic components in the series under investigation, leading to different results. For instance, while adding dummy variables to a regression model might be appropriate for a deterministic trend or seasonality, differentiation of a time series may be necessary for a stochastic trend or seasonal fluctuation. Therefore, before diving into the analysis, it is crucial to correctly identify the deterministic and probabilistic components inherent in the series. This well-known fact among researchers raises one of the most important questions: the cost of treating a deterministic factor as stochastic or vice versa in reality. Just as excluding a deterministic effect in the model causes model determination error, treating the specified effect as stochastic or vice versa can lead to many standard problems. At the same time, some packages may generate all type of results depending on the information provided and the direction of the research, potentially leading to confusion.

As is known, model selection represents the problem of using data to choose a model from among candidate models set (Wasserman, 2000:92; Kadane and Lazar, 2004:79). Essentially, it involves the use of a model selection criterion to find the most appropriate model that fits to the data. In particular, information criterion techniques emphasize minimizing the amount of information required to represent the data and model (Acquah,2019:2). In this framework, the correct model selection is of critical importance in reducing negative consequences (making incorrect decisions, noisy conclutions etc.) that may arise from the use of an inappropriate model (Ding,Tarokh and Yang, 2018:1).

Additionally, since different information criteria may yield different model selection in practice, it may eventually be necessary to choose one criterion to reach a final model. On the other hand, this situation causes confusion among researchers about whether the information criterion used is the best, the reliability of the model, the accuracy of the obtained findings, and so on. In fact, determining when and how to appropriately use information criteria is a complex issue, and the uncertainty in model selection should be assessed properly. Over the past few decades, researchers have found that the successful use of each method may depend on many factors, such as the underlying data generation process, the proposed models (especially how close they are to the data generation process), and the sample size, among others (Zhang, Yang and Ding, 2023:2). There are many studies about several objectives and using different methods in the literature to evaluate of performance of commonly used criteria and goodness-of-fit assessment tool (Clayton, Geisser and Jennings, 1986:437-438; Mills and Prasad,1992:219-221; Mavreski et al,2018:4; Acquah,2019:5-6, Zhang, Ding and Yang,2022:24, Hacker and Hatemi-J,2022:1055-1056).

Recent advances refine model selection strategies for complex time series. Dete et al. (2025) demonstrate that Kullback's symmetric divergence, a metric closely related to AIC's theoretical foundation, outperforms traditional criteria in regression models with non-linear trends, echoing our findings on HQ's robustness for stochastic trends. Meanwhile, Roy & Lesaffre (2025) propose Bayesian-focused criteria, highlighting trade-offs between parsimony and fit that parallel our comparisons of AIC (fit-focused) versus SIC (parsimony-focused). Finally, Li et al. (2025) validate deviance-based selection for models with latent components, supporting our analysis of stochastic seasonality (Table 2). Together, these studies underscore the importance of tailoring criteria to component types (deterministic vs. stochastic), a core contribution of our work.

As the study focuses on economic time series, the most frequently encountered deterministic components, trend and seasonality, were emphasized. In this context, series containing deterministic and/or stochastic trends and/or seasonality for deterministic and stochastic processes were generated through a simulation study. Various series with different sample sizes and effects were modeled to investigate the performance of model selection based on certain information criteria and goodness-of-fit tests. It is aimed to determine the success of correct classification rates of models including the basic components of economic time series. In this context, a Monte Carlo study has been conducted to compare these methods.

The structure of the paper is summarized as follows. The data generating proccess for time series under study is introduced within the section of theoretical background. The research methodology of the paper is presented within the methodology section following with an application study presenting results in details. In the last section, concluding remarks are given.

THEORETICAL BACKGROUND

Conventional time series analysis involves decomposing the fluctuations of the unobservable components of the series (trend, seasonal, periodic, and irregular) through purification procedures until the irregular component becomes a white noise process. In traditional methods, time series analysis is based on decomposing the series into trend, seasonal, cyclical, and irregular components. The general mathematical representation of the decomposition approach is expressed (Buteikis, 2020:4-5):

$$Y_t = f(T_t, S_t, C_t, E_t) \quad (1)$$

where, Y_t represents the time series value at period t (real data) and the trend component T_t is the long-term trend in the series. The seasonality component S_t represents fluctuations in the series depending on time intervals such as hours, days, months, quarters, etc., within a calendar year. The periodic component C_t represents cyclical changes that occur at intervals that last longer than a year. The irregular (residual) component E_t represents the non-systematic movements in the series.

Macroeconomic time series are generally determined by trend, seasonal, and periodic components as well. However, while the first two components are extensively studied in the literature, the periodic structure of the series is less emphasized. This situation can be attributed to the fact that the majority of economic time series contain at least one stochastic trend and exhibit strong seasonality that explains a significant portion of the variations (Franses, Hylleberg, and Lee, 1995.249-250; Caporale and Gil-Alana, 2007:155; Hylleberg, 2006.1-2). In this context, the comprehensive representation of the equation where a time series is assumed to combine trend, seasonal, and irregular components is organized as follows:

$$Y_t = T_t + S_t + E_t \quad (t = 1, \ldots, T) \qquad (t=1,...,T) \quad (2)$$

where Y_t represents an observable time series at time t, and T_t, S_t and E_t represent unobservable trend, seasonal, and irregular components, respectively. Additionally, one of the most important models for economic time series is the basic structural model consisting of these three components (Harvey and Peters, 1990:90). The multiplicative model under the assumption of the multiplication of unobservable components is in the form of:

$$Y_t = T_t \cdot S_t \cdot E_t \quad (3)$$

and it can be transformed into an additive form by taking the logarithms of the components.

$$\log(Y_t) = \log(T_t) + \log(S_t) + \log(E_t) \quad (4)$$

As stated before; economic time series often exhibit long-term increasing or decreasing trends, referred to as "trend." However, when a series contains any trend, the analysis begins after removing the effect of that trendto predict the behavior of the series no matter how long the duration is (either short or long). Fitting a direct autoregressive (AR) model to the non-stationary original series is not an appropriate approach (Kitawaga, 2010:159-160). Let's consider the simplest case where a time series is expressed as follows:

$$Y_t = T_t + \omega_t \quad (5)$$

where $T_t = a_0 + a_1 t + \cdots + a_m t^m$ represents the polynomial trend component, and $\omega_t \sim N(0, \sigma^2)$ follows a Gaussian distribution. Deterministic trend models are generally obtained using the following polynomial function (Kitagawa and Gerch, 1996).

$$T_t = \sum_{i=0}^{m} a_i t^i \quad (6)$$

$$\Delta^k T_t = v_t \quad (7)$$

When the actual trend exhibits a polynomial or a close to a polynomial, such a parametric model can provide a good estimate of the trend. However, in other cases, the model may not adequately capture the characteristics of the trend or become overly sensitive to random noise. Here, when a_i is a time-varying coefficient, it can also turn into a stochastic structure such as a random walk. Stochastic trend is based on the trend component model (Kitagawa and Gersch (1984, 1987,1996).

$$T_t = T_{t-1} + v_t \quad (8)$$

where v_t is a white noise series with a mean of 0 and an unknown variance, denoted as $N(0, \sigma_v^2)$. When $k=1$ in the difference equation, the trend component model transforms into a random walk and is referred to as a constant trend. Additionally, when $k=2$, linear trend component with constant slope is obtained, when $k=3$, a quadratic trend is obtained. A generalization of this model is represented as an exponential growth trend model.

$$T_t = \phi T_{t-1} + v_t \quad (9)$$

where $\phi > 1$ or $\phi < 1$ determines an increasing or decreasing trend. The average growth rate of T_t is calculated with $r = \phi - 1$ and it can be observed that the average growth rate is "0" in the random walk process when $\phi = 1$.

As known, seasonal components in time series, like trend, can be deterministic (predictable) or stochastic. Seasonal component S_t can be considered as a "signal" that is replicated at certain intervals, representing a structure (cycle) occurring repeatedly. A periodic signal means a stationary signal, and the aim here is to decompose a univariate periodic time series into simpler periodic functions. In this context, let "s" be a seasonal period signal, where $w_k = 2\pi k/s$ and $k = 0, 1, 2, \ldots, s/2$ are fourier frequencies for each. The first frequency, obtained by multiplying the frequency $w_1 = 2\pi/s$ is called the "fundemental frequency", and other frequencies obtained by multiplying this frequency are called "harmonics" (Giordano, Niglio and Storti, 2000:343-344).

$$S_t = \sum_{k=1}^{s/2} [A_{k,t}\cos(w_k t) + B_{k,t}\sin(w_k t)] \quad (10)$$

In this context, methods known as Fourier analyses turn into a suitable tool for modeling seasonal effects in time series using trigonometric functions. Fourier analysis or harmonic analysis is the decomposition of a time series into the sum of its sinusoidal components (in the form of sinusoidal curves). Sinusoidal curves can be used to approximate seasonal changes, and the effect of these components is removed to make the long-term trend more visible. Moreover, according to the Fourier Theorem, any periodic function can be expressed as a combination of periodic components known as the Fourier series.

Additionally, the trigonometric function in Eq. 10 can be rewritten in Eq. 11 ;

$$S_t = \sum_{k=1}^{s/2} [R_k \cos(w_k t + \phi_k)] \quad (11)$$

where equations for the kth harmonic can also be represented in terms of amplitude R_k and phase ϕ_k (Bloomfield, 2000:7-8, Rodriguez et al,2015:605):

$$\begin{aligned} A_k &= R_k\cos(\phi_k), \quad B_k = -R_k\sin(\phi_k) \\ R_k &= \sqrt{A_k^2 + B_k^2}, \quad \phi_k = \arctan(\tfrac{-B_k}{A_k}) \end{aligned} \quad (12)$$

METHODOLOGY

In this study, Monte Carlo simulation was conducted to model processes, and the results were obtained by following three fundamental stages: "Data Generation Process," "Prediction of Models," and "Evaluation of Criteria.

Data Generating Process and Simulation

In the study; It is aimed to obtain 17 different time series, from the simplest model containing only constants, to a more complex relationship structure in which all states of components exhibiting deterministic and/or stochastic properties are included. In this context; 10000 series representing each model were generated for 9 different sample sizes ranging from 60 and 500 observations. Additionally, the components for each model were constructed separately and then included in the additive equation. The following considerations were taken into account in the simulation design for the generated data:

- It was assumed that the unobservable components (T_t, S_t, E_t) are independent (orthogonal) from each other.
- There are no outliers/extreme/missing values in the process.
- The irregular component $E_t \sim N(0, \sigma_E^2)$. is a white noise process with an independent and identical distribution. For this study, σ_E^2 was set to 1.0.
- No adjustments, transformations, filtering, etc., were applied to the series.

The trend component T_t was simulated through three different cases: constant, deterministic, and stochastic random walk. For the constant and deterministic trend, it was generated with the model $a_i = 0$ for $i = 1, \cdots, m$ and $i = 3, \cdots, m$ for obtaining the stochastic trend, the relationship in Equation (9) was used with $v_t \sim N(0, \sigma_v^2)$. Specifically, ν_t was generated as a white noise process from a normal distribution with mean 0 and variance $\sigma_v^2 = 0.5$. The initial value for the random walk trend, T_0 , was set to 0.

Similarly, for the monthly seasonal component, two cases were considered: deterministic and stochastic seasonality. Equation 10 was used to obtain these components. To avoid complexity in the analysis, only the fundamental frequency was taken, and the phase value was constant while the amplitude was generated as a function of time for the stochastic seasonal component. For stochastic seasonality, the amplitude (A_k in Equation (10) or R_k in Equation (11)) was modeled as a simple random walk process, $A_K = A_{k-1} + \epsilon_t$, where ε_t is a white noise error term drawn from a normal distribution with mean 0 and variance $\sigma_\epsilon^2 = 0.2$. The initial amplitude A_0 was set to 1.

Prediction of Models

In ampirical applications, one widely accepted criterion for choosing among numerous prediction models is the good fit of the model to the data, meaning the model's forecasting accuracy is high. For example, when models satisfy constraints and assumptions, the best model is preferred, and the prediction performances of models are compared to choose the model that provides better predictions.

In this context, information criteria and goodness-of-fit values have been utilized for the purpose of comparing and evaluating the forecast accuracies of models.

Evaluation of Results

In the study, the OLS method was used in model predictions. Data generation processes, i.e, each series whose underlying relationship structure is known, were estimated on the basis of all models discussed in the study. The correct model identification rate (correct classification rate) for each criterion (Akaike,1974,1981; Schwarz et al.,1978; Hannan and Quinn,1979; Ebbler, 1975; Hurvich and Tsai, 1989; McQuarrie and Tsai,1998; Ucal,2006; Doğan and Doğan,2020; Riansut, 2023) was obtained and tabulated. To facilitate comparison, aggregated results were presented in Table1-2, while more detailed outcomes were provided in Appendix. MCCR stands for Mean Correct Classification Rate.

Regardless of the model type, meaning for any model, the average performances of criteria were calculated solely based on sample sizes and displayed in Table 1.

The most notable result in the table is the detection success of the adjusted R^2, which is low for small samples. It can be observed that the HQ criterion is effective in models with very large samples, while the AIC criterion is partially effective in models with large samples.

When evaluating the average true model identification rates for each model collectively corresponding to each criterion, the rates are provided in Table 1.

One of the most notable results is that the Log-Likelihood captures the appropriate model less accurately compared to other information criteria. Another noteworthy finding is that the SIC information criterion's effectiveness decreases as the number of parameters in the model increases, consistent with the literature.

CONCLUSION

In this study, a comprehensive Monte Carlo simulation was conducted to compare the performance of six widely used model selection criteria—Adjusted R^2, Log-likelihood, AIC, SIC, HQ, and MSE—across 17 different time series models and 9 varying sample sizes. The primary objective was to determine which criterion best identifies the true underlying data generating process. The correct classification rates for each criterion were meticulously calculated and presented in detailed tables within the manuscript and the Appendix.

Our findings reveal several key insights. Firstly, the ability of all information criteria to select the correct model generally decreases with smaller sample sizes. For very large sample sizes, the Hannan-Quinn (HQ) information criterion consistently yielded the best results across all types of time series models. Conversely, for models with sample sizes less than 120, "Adjusted R^2 proved to be the most effective choice for correct model selection. As sample sizes increased beyond this threshold, the Akaike Information Criterion (AIC) emerged as the superior option. Overall, AIC and Adjusted R^2 were found to be the most reliable criteria for appropriate model selection in various scenarios.

A notable observation was the strong performance of almost all information criteria when the time series included a stochastic seasonal component. In contrast, if the time series contained a deterministic seasonal component, the Schwarz Information Criterion (SIC) demonstrated reduced applicability, suggesting its performance diminishes as the number of parameters in the model increases, consistent with existing literature. Furthermore, contrary to some interpretations, the Log-Likelihood criterion consistently performed poorly in

Table 1: Correct Model Identification Rate (Correct Classification Rate) Independent of Model Type

Criteria	N=500	N=250	N=200	N=150	N=120	N=108	N=96	N=72	N=60
AIC	76.59	73.81	70.74	65.49	61.20	59.04	57.10	52.22	49.29
HQ	86.05	68.46	62.79	57.73	55.18	54.04	53.33	51.21	49.44
MSE	60.46	60.08	59.49	58.36	57.02	55.94	54.68	51.11	48.22
SIC	73.43	59.53	58.63	57.78	57.03	56.55	56.34	55.16	53.91
Log-likelihood	34.98	25.67	22.37	17.03	14.93	14.19	13.69	12.76	12.39
Adjusted R^2	69.04	66.25	65.66	64.37	63.17	62.45	61.37	56.43	53.45

Table 2. Correct Model Identification Rate (Correct Classification Rate) Independent of Sample Size

Models	AIC	HQ	MSE	SIC	Log-likelihood	Adjusted R^2
Y_{1t}=c	40.99	58.61	13.16	74.70	0.00	15.12
Y_{2t}=Deterministic Trend	52.32	66.23	22.66	78.66	0.00	45.95
Y_{3t}=Stochastic Trend	49.87	64.20	21.34	77.82	0.00	15.64
Y_{4t}=c+ Deterministic Trend	60.74	71.35	29.24	81.25	0.00	58.01
Y_{5t}=c+ Stochastic Trend	59.16	70.06	29.22	80.19	0.00	55.27
Y_{6t}= Deterministic Seasonal	37.20	19.99	38.43	4.38	0.00	40.79
Y_{7t}=c+ Deterministic Seasonal	37.36	20.12	38.55	4.33	0.00	40.79
Y_{8t}= Stochastic Seasonal	53.33	76.54	25.10	91.40	0.00	27.76
Y_{9t}=c+ Stochastic Seasonal	71.10	84.91	47.58	93.82	11.11	38.50
Y_{10t}= Deterministic Trend + Deterministic Seasonal	59.45	24.89	88.46	4.69	24.59	94.83
Y_{11t}=c+ Deterministic Trend + Deterministic Seasonal	60.11	24.98	88.92	4.74	20.93	94.76
Y_{12t}= Deterministic Trend + Stochastic Seasonal	83.59	91.68	68.05	96.66	11.66	80.84
Y_{13t} =c+ Deterministic Trend + Stochastic Seasonal	99.99	99.99	99.95	99.95	88.89	99.26
Y_{14t}= Stochastic Trend + Deterministic Seasonal	60.74	26.42	88.86	5.43	32.78	92.54
Y_{15t}=c+ Stochastic Trend + Deterministic Seasonal	59.84	26.64	88.14	5.58	27.26	94.65
Y_{16t}= Stochastic Trend + Stochastic Seasonal	83.43	91.47	67.69	96.45	11.64	67.69
Y_{17t}=c+ Stochastic Trend + Stochastic Seasonal	98.92	98.61	99.21	97.95	88.48	99.52

almost all models, indicating it may not be useful for selecting the appropriate model in these contexts.

These simulation results provide valuable practical guidance for researchers and practitioners engaged in econometric time series modeling. By considering the sample size and the nature of the seasonal components, researchers can make more informed decisions regarding the choice of model selection criterion, thereby enhancing the reliability of their analyses and avoiding incorrect conclusions.

It is important to acknowledge certain limitations of this study. Our simulation design assumes orthogonality (independence) between the trend, seasonal, and irregular components. While this approach simplifies the analysis and facilitates clear attribution of component effects, real-world economic time series often exhibit more complex, interdependent structures that are not captured under this assumption. Furthermore, the generated data intentionally excluded outliers and missing values, which are common challenges encountered in empirical time series analysis. Future research could explore the performance of these criteria under more complex data generating processes, including correlated components, structural breaks, and the presence of atypical observations.

REFERENCES

Acquah, H. G. (2019). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship, *International Journal of Agricultural Economics and Extension* ISSN 2329-9797 Vol. 7 (1), pp. 001-006, www.internationalscholarsjournals.org.

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19 (6). pp: 716–23.

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics*, 16:3–14.

Bloomfield, P. (2000). Fourier analysis of time series an introduction, *Probability and Statistics Applied Probability and Statistics Section,* John Wiley &Sons, Inc., Canada and USA.

Buteikis, A. (2020). Pratical econometrics and data, *Faculty of Mathematics and Informatics, Institute of Applied Mathematics, Vilnius University, Lithuania, http://web.vu.lt/mif/a.buteikis/*.

Caporale, G. M. and Gil-Alana, L. (2007). Testing for deterministic and stochastic cycles in macroeconomic time series. Empirica 34, 155–169 (2007). https://doi.org/10.1007/s10663-007-9033-4

Clayton, M. K., Geisser, S. and Jennings, D. (1986). A comparison of several model selection procedures Bayesian inference and decision techniques: Essays in Honor of Bruno De Finetti, Studies in Bayesian Econometrics and Statistics, Vol 6, Chapter 27, (Ed: P.K. Goel and A.Zellner), Elseiver Science Publishers B.V.

Dete, C. H., Lokonon, B. E., Gneyou, K. E., Senou, M., Glèlè Kakaï, R., (2025). Relative Performance of Model Selection Criteria for Cox Proportional Hazards Regression Based on Kullback's Symmetric Divergence, Journal of Probability and Statistics, 2025, 3808705, 16 pages,. https://doi.org/10.1155/jpas/3808705.

Ding, J., Tarokh, V. and Yang, Y. (2018). Model selection techniques-an overwiew, *IEEE Signal Processing Magazine*, 35(6), DOI:10.1109/MSP.2018.2867638Corpus ID: 53035396.

Doğan, İ. and Doğan, N. (2020). Model performans kriterlerinin kronolojisine ve metodolojik yönlerine genel bir bakış: bir gözden geçirme, *Türkiye Klinikleri Biyoistatistik Dergisi*, 12(1):114-25.

Ebbler, D. H. (1975), On the probability of correct model selecting using the maximum r-square choıce criterion, *International Economic Review*, 16-2, pp.516-520.

Franses, P. H., Hylleberg, S. and Lee, H. S. (1995). Spurious deterministic seasoanality, *Economics Letters 48* (1995) 249-256.

Kitagawa, G. and Gerch, W. (2012), Smoothness Priors Analysis of Time Series, Lecture Notes in Statistics 116, Springer-Verlag New York.

Gil-Alana, L.A. (2005). Testing of stochastic trends, seasonal and cylical components in macroeconomic time series, *The Korean Communications in Statistics,* 12(1), pp. 101-115.

Giordano, F., Niglio, M. And Storti, G. (2000).A simulation study for the evaluation of the seasonal adjustment and forecasting performances of the TESS System, Statistica Applicata,12(3).

Hacker, R. S. and Hatemi-J, A. (2022), "Model selection in time series analysis: using information criteria as an alternative to hypothesis testing", *Journal of Economic Studies*, Vol. 49 No. 6, pp. 1055-1075. https://doi.org/10.1108/JES-09-2020-0469

Hannan, E. J. and Quinn, B. G. (1979).The Determination of the Order of an Autoregression, *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, 190–95.

Hylleberg, S. (2006). Seasonal adjustment, Working Paper 2006(4), Deparment of Economics, University of AARHUS, Denmark.

Harvey, A. C. and Peters, S. (1990). Estimation Procedures for structural time series. Journal of Forecasting, 9(89), 89-108.

Hurvich, C. M. and Tsai, C. L. (1989). Regression and time series model selection in smal samples. *Biometrica*, 76(2), 297-307.

Ihaka, R. (2005). Time series analysis lecture notes for 475.726, *Statistics Department University of Auckland, New Zeland.*Kadane, J. B. and Lazar, N. A. (2004). Methods and Criteria for model selection, *Journal of the American Statistical Association,* 99(465), pp. 279-290.

Kitawaga, G. (2010).Introduction to time series modelling, *Momographs on Statistics and Applied Probability 114*, A Chapman & Hall Book, Taylor&Francis Group, Japan.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research, 33(2)*, 188–229. https://doi.org/10.1177/0049124103262065

Li, Y., Sushanta, K. M., Wang, N., Yu, J. and Zeng, T. (2025). Deviance Information Criterion for Bayesian model selection: Theoretical justification and applications, Journal of Econometrics, 2025, 105978, ISSN 0304-4076, https://doi.org/10.1016/j.jeconom.2025.105978.

McQuarrie, A. D. and Tsai, C. L. (1998) Regression and time series model selection,World Sciencetific, 1998

Mavreski, R., Milanov, P.,Traykov, M. and Pencheva, N. (2018). Performance comparison of model selection criteria by generated experimental data, ITM Web of Conferences16, 022006 AMCSE 2017, https://doi.org/10.1051/itmconf/20181602006.

Mills, J. A. and Prasad, K. (1992) A comparison of model selection criteria, Econometric Reviews, 11:2, 201-234, DOI: 10.1080/07474939208800232.

Riansut, W. (2023). A study of the effectiveness of model selection criteria for multiple regression model. *Rajamangala University of Technology Srivijaya Research Journal*, 15(1), 198–212. Retrieved from https://li01.tci-thaijo.org/index.php/rmutsvrj/article/view/249278.

Rodríguez, E. G., Villalobos, H., Muñoz, V. M. G.and Rodríguez, A. R. (2015). Computational method for extracting and modeling periodicities in time series, *Open Journal of Statistics,*5(6), 604-617, doi: 10.4236/ojs.2015.56062.

Roy, B., & Lesaffre, E. (2025). Focused information criteria for model selection – a Bayesian perspective. Journal of Applied Statistics, 1–19. https://doi.org/10.1080/02664763.2025.2514152.

Schwarz, Gideon, and others. 1978. "Estimating the Dimension of a Model." The Annals of Statistics 6 (2). Institute of Mathematical Statistics: 461–64.

Ucal, M. Ş. (2006). Ekonometrik model seçim kriterleri üzerine kısa bir inceleme, *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 7(2), 41-56.

Wasserman, L. (2000).Bayesian model selection and model averaging. Journal of Mathematical Psychology, 44(1), 92–107. https://doi.org/10.1006/jmps.1999.1278

Zhang, J., Ding, J. and Yang, Y. (2022). Is a classification procedure good enough?—A goodness-of-fit assessment tool for classification learning. Journal of the American Statistical Association, 1–11. https://www.tandfonline.com/doi/abs/10.1080/01621459.2021.1979010

Zhang, J., Yang, Y. And Ding, J. (2023). Information criteria for model selection, WIREs Computantional Statistics, Advanced Review, Wiley, DOI: 10.1002/wics.1607

APPENDIX

True Model Identification Rates (Correct Classification Rate) For Each Criteria and Goodness of Fit Assessment Tools

Table 1. True Model Identification Rates (Correct Classification Rate) via AIC for 17 Different Generated Models Against 6 Different Sample Sizes

AIC	N=500	N=250	N=200	N=150	N=120	N=108	N=96	N=72	N=60
Y_{1t}	41.69	41.36	41.73	41.92	41.31	39.92	40.03	40.64	40.31
Y_{2t}	53.26	52.58	53.30	52.38	52.29	51.85	52.48	51.75	50.99
Y_{3t}	51.69	51.70	51.26	50.94	50.04	49.68	48.89	47.90	46.72
Y_{4t}	60.80	61.67	61.12	61.03	60.49	60.41	60.72	60.72	59.71
Y_{5t}	61.82	61.14	59.87	59.33	59.75	58.71	58.78	57.67	55.41
Y_{6t}	65.33	58.34	52.08	41.32	33.10	28.56	25.17	16.87	14.07
Y_{7t}	64.79	59.09	51.69	41.09	33.24	29.26	24.35	18.00	14.74
Y_{8t}	63.26	61.48	61.32	61.11	59.46	57.61	54.30	37.09	24.37
Y_{9t}	70.83	71.41	72.30	71.69	71.65	71.30	71.94	70.65	68.13
Y_{10t}	99.93	91.32	81.25	65.83	53.42	46.67	41.46	29.84	25.31
Y_{11t}	99.91	91.58	83.15	66.47	52.84	47.82	42.07	31.01	26.17
Y_{12t}	84.38	84.64	83.18	83.76	83.25	83.41	83.15	83.20	83.32
Y_{13t}	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.99	99.88
Y_{14t}	99.94	92.21	83.25	68.04	54.16	48.68	42.93	31.32	26.16
Y_{15t}	99.94	91.90	83.44	65.20	52.23	46.46	41.60	30.85	26.92
Y_{16t}	84.48	84.28	83.71	83.22	83.31	83.52	83.21	82.82	82.35
Y_{17t}	100.00	100.00	100.00	100.00	99.93	99.84	99.66	97.43	93.41
MCCR*	**76.59**	**73.81**	**70.74**	**65.49**	**61.20**	**59.04**	**57.10**	**52.22**	**49.29**

* Mean Correct Classification Rate

Table 2. True Model Identification Rates (Correct Classification Rate) via HQ for 17 Different Generated Models Against 6 Different Sample Sizes

HQ	N=500	N=250	N=200	N=150	N=120	N=108	N=96	N=72	N=60
Y_{1t}	64.34	61.19	60.94	60.27	57.95	56.99	56.46	55.52	53.85
Y_{2t}	69.82	68.34	67.98	66.92	66.43	65.14	65.21	63.93	62.34
Y_{3t}	69.58	67.7	67.03	65.5	64.45	62.88	62.04	60.36	58.26
Y_{4t}	73.43	72.57	72.33	72.45	71.13	70.57	70.81	69.89	68.98
Y_{5t}	74.34	73.36	71.83	71.01	70.05	69.16	68.81	67.35	64.66
Y_{6t}	81.32	36.57	23.47	12.73	8.02	6.3	5.32	2.88	3.27
Y_{7t}	81.16	37.55	23.45	12.21	8.67	6.82	5.04	3.34	2.84
Y_{8t}	85.85	83.24	82.11	80.03	78.78	77.94	76.59	67.58	56.7
Y_{9t}	88.91	87.54	87.27	85.62	85.14	83.97	83.97	81.88	79.92
Y_{10t}	96.06	45.15	29.81	16.73	11.17	8.2	7.24	5.3	4.31
Y_{11t}	95.85	46.52	30.35	16.52	10.02	8.91	7.1	4.91	4.67
Y_{12t}	94.55	92.95	92.99	92.28	91.31	90.98	90.91	89.74	89.42
Y_{13t}	100	100	100	100	100	100	100	99.99	99.89
Y_{14t}	96.42	48.7	32.37	18.96	12.13	9.65	8.11	5.75	5.73
Y_{15t}	96.43	49.17	32.9	18.34	11.66	10.11	8.8	6.23	6.09
Y_{16t}	94.82	93.3	92.61	91.87	91.22	91.38	90.75	89.33	87.94
Y_{17t}	100	100	100	99.98	99.92	99.74	99.51	96.64	91.67
MCCR*	**86.05**	**68.46**	**62.79**	**57.73**	**55.18**	**54.04**	**53.33**	**51.21**	**49.44**

Table 3. True Model Identification Rates (Correct Classification Rate) via MSE for 17 Different Generated Models Against 6 Different Sample Sizes

MSE	N=500	N=250	N=200	N=150	N=120	N=108	N=96	N=72	N=60
Y_{1t}	14.14	14	13.37	13.43	13.23	12.4	12.98	13.07	11.8
Y_{2t}	24.25	22.87	23.1	23.37	22.47	22.44	22.65	21.43	21.4
Y_{3t}	23.26	22.61	22.41	21.6	21.34	21.15	21.04	19.84	18.79
Y_{4t}	30	30.56	29.96	29.18	29.89	28.92	28.7	28.59	27.4
Y_{5t}	31.31	30.2	29	29.38	29.57	29.24	28.73	28.5	27.06
Y_{6t}	43.97	42.11	42.69	39.9	39.86	38.18	36.79	32.36	29.97
Y_{7t}	43.29	43.33	41.93	40.88	39.36	38.37	35.59	33.59	30.57
Y_{8t}	34.62	33.46	33.31	32.81	29.85	26.37	22.5	9.39	3.55
Y_{9t}	46.71	48.02	48.06	48.62	48.89	48.56	49.57	48.38	41.39
Y_{10t}	100	99.11	97.94	94.35	89.65	87.27	83.57	74.72	69.51
Y_{11t}	100	99.14	98.01	94.64	90.08	87.53	84.58	75.47	70.85
Y_{12t}	68.46	68.76	67.46	67.74	67.78	68.46	67.69	67.6	68.54
Y_{13t}	100	100	100	100	100	100	100	99.96	99.58
Y_{14t}	100	99.34	98.24	94.72	90.37	87.75	84.31	75.68	69.32
Y_{15t}	100	99.37	98.02	93.94	89.67	86.51	83.07	74.49	68.15
Y_{16t}	67.78	68.47	67.78	67.49	67.35	67.94	68.1	67.63	66.69
Y_{17t}	100	100	100	100	99.93	99.88	99.74	98.12	95.18
MCCR*	**60.46**	**60.08**	**59.49**	**58.36**	**57.02**	**55.94**	**54.68**	**51.11**	**48.22**

* Mean Correct Classification Rate

Table 4. True Model Identification Rates (Correct Classification Rate) via SIC for 17 Different Generated Models Against 6 Different Sample Sizes

SIC	N=500	N=250	N=200	N=150	N=120	N=108	N=96	N=72	N=60
Y_{1t}	82.43	78.83	77.61	76.34	73.93	72.45	71.98	70.44	68.29
Y_{2t}	84.07	82	81.03	79.49	78.59	77.16	76.79	75.29	73.54
Y_{3t}	85.27	82.2	80.48	79.28	77.87	76.19	75.17	72.54	71.42
Y_{4t}	85.57	83.06	82.95	82.52	80.47	80.04	79.72	79.15	77.81
Y_{5t}	86.1	84.17	82.87	81.4	79.6	79.18	78.49	76.26	73.68
Y_{6t}	35.95	2.14	0.85	0.24	0.07	0.07	0.03	0.01	0.1
Y_{7t}	35.68	2.01	0.72	0.18	0.17	0.07	0.06	0.03	0.02
Y_{8t}	96.63	94.57	93.91	92.27	90.94	90.66	90.29	87.63	85.72
Y_{9t}	97.26	96.03	95.91	94.5	94.08	92.72	93.13	91.09	89.62
Y_{10t}	38	2.39	1.04	0.26	0.16	0.15	0.11	0.04	0.05
Y_{11t}	38.25	2.43	0.92	0.34	0.21	0.2	0.13	0.09	0.06
Y_{12t}	98.76	97.63	97.8	97.25	96.51	96.06	96.07	95.22	94.68
Y_{13t}	100	100	100	100	100	100	100	99.97	99.6
Y_{14t}	42.52	3.4	1.45	0.56	0.31	0.19	0.24	0.11	0.11
Y_{15t}	43.09	3.38	1.62	0.68	0.39	0.25	0.28	0.22	0.29
Y_{16t}	98.66	97.75	97.61	96.92	96.3	96.47	96.25	94.72	93.4
Y_{17t}	100	100	100	99.98	99.83	99.55	99.11	94.96	88.1
MCCR*	**73.43**	**59.53**	**58.63**	**57.78**	**57.03**	**56.55**	**56.34**	**55.16**	**53.91**

* Mean Correct Classification Rate

Table 5. True Model Identification Rates (Correct Classification Rate) via Log-Likelihood for 17 Different Generated Models Against 6 Different Sample Sizes

Log.L	**N=500**	**N=250**	**N=200**	**N=150**	**N=120**	**N=108**	**N=96**	**N=72**	**N=60**
Y_{1t}	0	0	0.02	0	0	0	0	0	0
Y_{2t}	0	0	0	0	0	0	0	0	0
Y_{3t}	0	0	0	0	0	0	0	0	0
Y_{4t}	0	0	0	0	0	0	0	0	0
Y_{5t}	0	0	0	0	0	0	0	0	0
Y_{6t}	0	0	0	0	0	0	0	0	0
Y_{7t}	0	0	0	0	0	0	0	0	0
Y_{8t}	0	0	0.02	0	0	0	0	0	0
Y_{9t}	0.01	0	100	0	0	0.01	0.01	0	0
Y_{10t}	97.68	50.77	32.31	16.34	9.18	6.06	4.85	2.56	1.53
Y_{11t}	97.5	52.05	0	16.09	8.48	6.51	4.34	2.12	1.3
Y_{12t}	0.75	0.71	100	0.69	0.84	0.81	0.77	0.13	0.2
Y_{13t}	100	100	0	100	100	100	100	100	99.99
Y_{14t}	99.03	66.26	47.95	28.5	17.77	13.53	10.9	6.3	4.82
Y_{15t}	98.88	66.05	0	27.35	16.76	13.49	11.21	6.5	5.08
Y_{16t}	0.8	0.61	100	0.59	0.76	0.8	0.74	0.22	0.28
Y_{17t}	100	100	0	100	99.97	99.96	99.88	99.05	97.48
MCCR*	**34.98**	**25.67**	**22.37**	**17.03**	**14.93**	**14.19**	**13.69**	**12.76**	**12.39**

* Mean Correct Classification Rate

Table 6. True Model Identification Rates (Correct Classification Rate) via Adj. R^2 for 17 Different Generated Models Against 6 Different Sample Sizes

Adj.R^2	**N=500**	**N=250**	**N=200**	**N=150**	**N=120**	**N=108**	**N=96**	**N=72**	**N=60**
Y_{1t}	17.14	16.58	16.22	15.37	14.74	15.15	14.64	13.77	12.49
Y_{2t}	73.08	47.4	44.73	43.21	41.1	41.42	41.81	40.44	40.37
Y_{3t}	14.45	14.54	15.5	15.39	15.91	15.76	16.14	16.11	16.97
Y_{4t}	76.93	59.05	57.76	55.17	55.95	54.55	53.96	54.75	53.93
Y_{5t}	55.43	56.19	55.42	55.27	55.15	55.5	55	54.98	54.47
Y_{6t}	44.3	43.25	43.46	42.01	42.04	41.31	40.5	36.4	33.85
Y_{7t}	43.65	43.37	43.56	42.62	41.83	41.04	38.73	37.38	34.96
Y_{8t}	34.64	33.48	33.32	33.17	31.85	29.65	27.51	16.74	9.48
Y_{9t}	46.91	48.08	48.11	48.66	48.7	47.7	47.08	11.04	0.19
Y_{10t}	100	99.99	99.72	98.73	96.66	95.36	93.07	86.82	83.1
Y_{11t}	100	99.95	99.61	98.73	96.45	95.4	93.15	86.75	82.76
Y_{12t}	99.43	96.31	92.3	83.37	75.75	73.63	71.5	68.42	66.87
Y_{13t}	100	100	100	100	100	99.99	100	99.14	94.2
Y_{14t}	100	99.65	99.08	96.94	93.9	92.27	89.56	83.16	78.28
Y_{15t}	100	99.95	99.7	98.23	96.54	95.01	92.6	86.87	82.94
Y_{16t}	67.78	68.47	67.78	67.49	67.36	67.95	68.1	67.61	66.7
Y_{17t}	100	100	100	100	99.95	99.96	99.86	98.92	97.02
MCCR*	**69.04**	**66.25**	**65.66**	**64.37**	**63.17**	**62.45**	**61.37**	**56.43**	**53.45**

* Mean Correct Classification Rate