Iraqi Journal of Statistical Sciences

http://stats.uomosul.edu.iq

Estimation of Delay Time in Linear Dynamic Systems Using Wavelets

Reem Talal Taha [1], Heyam A.A.Hayawi[2] Mohamed Ahmed Elkhouli[3]

[1,2]Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq
[3] Sadat Academy for Management Science (SAMS), Department of Mathematics and Statistics, Cairo, Egypt.

Article information

Article history:
Received September 1, 2024
Revised: March 5,2025
Accepted March 25, 2025
Available online June 1, 2025

Keywords: Wavelet, Har Wavelet, Delay Time, The Black Box

Correspondence:
Heyam A.A.Hayawi
he.hayawi@uomosul.edu.iq

Abstract

This study delves into the application of wavelet transforms for the estimation of delay time in stochastic linear dynamic systems. The significance of accurately determining delay time lies in its critical role in system diagnostics, where it establishes the temporal gap between the system's inputs and its outputs. In this context, a comprehensive set of simulation experiments was conducted, focusing on the use of the Haar wavelet for data processing. The Haar wavelet was specifically selected due to its simplicity and effectiveness in capturing essential features within time-series dataVarious methodologies for estimating delay time were implemented and subsequently compared across different scenarios to assess their performance. Among the approaches examined, the Haar wavelet-based estimation demonstrated superior outcomes, particularly when applied to an autoregressive model incorporating additional inputs. In comparison to the results obtained from unprocessed data, the Haar wavelet-based method provided a more accurate estimation of delay time, suggesting its potential advantages in practical applications.The primary objective of this research is to explore the efficacy of the Haar wavelet in the estimation of delay time within stochastic linear dynamic models. To achieve this, several estimation techniques were employed and the resulting performances were evaluated and compared based on the outcomes of the simulation experiments. This study contributes valuable insights into the potential of wavelet transforms, particularly the Haar wavelet, in enhancing the accuracy of delay time estimation, which is a key aspect of system modeling and diagnostic analysis. The findings underscore the importance of incorporating advanced signal processing techniques, such as wavelets, for improving the precision and reliability of time delay estimates in complex dynamic systems.

DOI: 10.33899/iqjoss.2025.187788 , ©Authors, 2025, College of Computer Science and Mathematical, University of Mosul.

1. Introduction

The analysis of time series is a crucial area of statistical research, as it provides valuable insights into the underlying behavior of phenomena, offering an accurate and detailed description of the subject under study. This analysis plays a vital role in understanding trends, patterns, and anomalies over time, which is essential for making informed decisions in various scientific and practical domains. A key goal of time series analysis is forecasting, which has become an indispensable tool in modern decision-making processes. Forecasting future values of a phenomenon enables researchers to anticipate potential developments and devise strategies to address challenges that affect human life globally.

A well-designed forecasting system is essential for effectively adapting available plans to suit evolving circumstances and align with managerial expertise. Such systems must be capable of producing reliable predictions to ensure they integrate seamlessly into the existing decision-making framework. When a forecasting system fails to deliver accurate predictions, it compromises the entire decision-making process, rendering it ineffective and incompatible with the goals of the organization or institution. Therefore, forecasting is not just a tool but an integral part of decision-making at various organizational levels, guiding institutions in setting objectives and selecting the most appropriate scientific approaches to achieve them.

In this context, the process of diagnosing black-box models becomes crucial. Identifying the most suitable model to represent a given process or system is essential for accurate forecasting and system optimization. In recent years, researchers have turned to advanced techniques such as wavelet transformation, which has proven to be highly effective in enhancing system diagnostics. Wavelet transformations offer a powerful method for analyzing time series data, capturing both high and low-frequency components of the signal. As a result, they provide more accurate and reliable estimates, significantly improving the effectiveness of system diagnostics and forecasting. This integration of wavelet transformations into the analysis process has led to more precise predictions, making them increasingly popular in modern research and engineering applications.

2. Delay Time Estimation Methods

There are many statistical style and methods that serve to estimate the delay time in dynamic systems or the so-called black box models. We mention:-

a)The Cross-Correlation Function

Also called the cross-correlation function, it is calculate as follows (Box.et.al.,2016):

$$\rho_{uy}(k) = \frac{\gamma_{uy}(k)}{\sigma_u \sigma_y} \qquad , k = 0 \pm 1, \pm 2, 3, \ldots \tag{1}$$

$$\gamma_{uy}(k) = E(u_t - a_u)(y_t - a_y) \tag{2}$$

ρ_{uy}: expresses the multiplicative correlation function between the inputs and the outputs.

$\gamma_{uy}(k)$:The covariance between input and output.

$\sigma_u \sigma_y$: Standard deviation of inputs and outputs straight.

The highest absolute value of the cross-correlation has taken as follows: $d = Max|pu|(k)$

b)Impulse Response Function

The impulse response function is one of the methods used to estimate the delay time in dynamic systems, and it is an indicator of the relationship between inputs and outputs, as (y_t) the outputs that represent the effect that occurred as a result of changing the inputs (u_t) by one unit(1983,Makridakis et.al):

$$Y_t = V(B)u_t + \varepsilon_i \tag{3}$$

As: y_t: inputs. u_t : outputs. ε_i: expresses a disturbances

$V(B) = v_0 + v_1 B + v_2 B^2 + \cdots$: They represent the weights of the impulse response, that (B) is the back displacement factor and that the first significant value outside the confidence intervals is the one that represents the estimated delay time.

c)Use of autoregressive models with additional input:

Autoregressive with eXogenons Variables(ARX)

This method suggested by the researcher (1995.Ljmg) to estimate the delay time, as the data are divide as being studied from the inputs and outputs together into two parts: the first part is calculated for the estimation and the second part we work on choosing the legitimacy of the estimation (Ljung,1999):-

3. Wavelet

It is a type of function that works to decompose the functions into components of different frequency, and then each component is studied (Gunther 2001). The wavelet is defined mathematically as that function with a real value defined on a complete real axis and is oscillating up and down in an orderly manner around zero (Hamoudat, 2020). The wavelet consists of two main parts (Hayawi & Alsharabi,2022):

1) The first part is the Scaling Function also called the Father Function (Guther 2001), which can be obtained through the following formula(Hayawi,2022):

$$\varphi(x) = \sum_{k=0}^{N} C(k)\varphi(2x - k) \tag{4}$$

$C(k)$): Represent the coefficients of the low pass filter

The above equation is called the detailed equation

2)The second part is Wave Function also called the Mother Function, which can be obtained through the following formula:

$$w(x) = \sum_{k=0}^{N} d(k)\varphi(2x-k) \qquad (5)$$

The above equation is called (the wavelet equation)(Mao,2004)

$d(k)$: Represent the coefficients of the high pass filter

4-1 wavelet properties

The wavelet is a small wave that is generate and fades out at a specific time instead of being a wave that continues forever, such as the sine waves that are determined within the (-∞,∞) period,

There are many types of wavelets, including:

Frist: Haar wavelet:

It was known by the mathematician (Afred Haar) when he proposed this wavelet in 1909 and it is considered one of the most important and simple types of wavelets and is the basis for generating some other types of wavelets. It is also known as the mother wavelet (Green, 1994). Haar Mather Wavelet is calculated through:

$$\psi(x) = \begin{cases} 1 & if \quad 0 < x \le \frac{1}{2} \\ -1 & if \quad \frac{1}{2} < x \le 1 \\ 0 & O.W. \end{cases} \qquad (6)$$

The measurement function is known as the Father wave, and it can be defined as follows:

$$\psi(x) = \begin{cases} 1 & if \quad 0 \le x \le 1 \\ 0 & O.W. \end{cases} \qquad (7)$$

Second: Daubechies wavelet

Third: Coiflets wavelet

Forth: Symmlets wavelet

4. Criteria for Choosing theBest Model

1. Square Error Rate Criterion

It is defined by the following formula

$$MSE = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n} \qquad (8)$$

2. Akaike Information Criterion

This criterion was also known by the scientist Akaike in the year (1973-1974) and is symbolized by the AIC, as he provided information for choosing the appropriate rank for the ARIMA model among several models. The appropriate rank corresponds to the lowest value of the AIC criterion and represents the most appropriate rank, and is expressed mathematically as follows (Wei,2006):

$$AIC = 2k - 2lnL \qquad (9)$$

As:- AIC: expresses the Akaike criteria of information, L: expresses the maximum possible function

Akaike criteria is approximately "for the linear Tyler series criterion for Colpak-Liber information, and the Akaike criterion the corrected for information AIC_C is an approximation of the non-linear Tyler series, so the corrected Akaike criterion is more accurate than the Akaike criterion for the information and through it the model that corresponds to the lowest value of this criterion is chosen as the best model and is calculated in the following form (Burnham,Anderson,2002):

$$AIC_c = nln\sigma_e^2 + 3k + \frac{2k(k+1)}{(n-k-1)} \qquad (10)$$

3. Akaike's Final Prediction Error Criterion

It is one of the important criteria in determining the appropriate rank of the model, as it was known by the scientist Akaike in 1969 and is symbolized by the FPE. It represents the measure of the final prediction error. It is defined as the variance of prediction error for a future period, and it is calculated as follows (Ljung,2004):

$$FPE = \frac{1 + m/N}{1 - m/N} V \quad (11)$$

FPE: expresses the final prediction error, m: expresses the number of parameters in the model

4. Bayesian criteria for information

In the years (1979, 1978) the researcher Akaike developed the AIC criteria into a new one that differs slightly from the old criteria, which was called the BIC Criteria for Information. It can be explained it's formula as follows (Akaike,1981):

$$BIC(k) = nlog\sigma_a^2 - (n-k)\ln\left(1 - k/n\right) + kln(n) + kln\left[\left(\frac{\sigma_y^2}{\sigma_a^2} - 1\right)|k\right] \quad (12)$$

σ_y^2: represent the amount of variance of the series, as the model that corresponds to the lowest value for this criterion is chosen

5. Simulation Experiments:

Simulation is defined as a process of representation of the real reality using certain models, as many processes are complex in understanding and analysis, so it is the best solution to describe this process in a comparative manner through an analogy to the real reality, that the degree of similarity must be between a simulated experience with the real reality to obtain an identical or similar in the simulation model to the real system. (Mohammed, 2020).

5-1 Description of the simulation experience

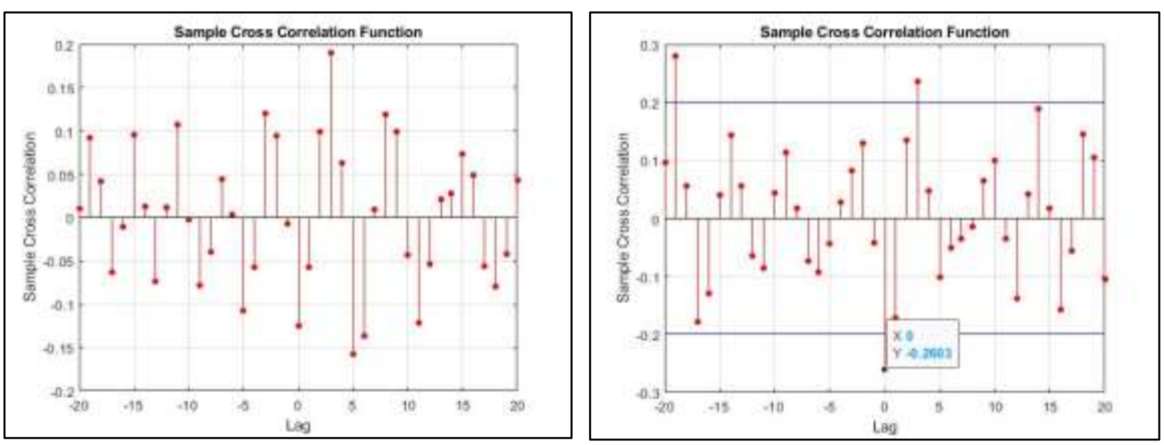
The simulation was designed using the MATLAB program and using the Monte Carlo method, as values were assigned to samples of different sizes (n = 100,250,500) and random data was generated with a specific delay time for the purpose of estimating the delay time and employing the wavelet method in that.

The inputs were generated according to random signals from the standard normal distribution (rgs) using the instruction randn symbolized by the symbol u_t, while the interference was generated using randn also to generate random signals that follow a normal distribution $\varepsilon_t \sim N(0{,}5)$ and other noise signals with a standard normal distribution also $\varepsilon_t \sim N(0{,}1.9)$ through the stochastic linear kinetic model box-Jenkinz with a delay time of (2), as shown in the following equation:

$$Yt = 0.3u_t - 2 + 0.45u_{t-3} - 0.0u_{t-4} + 0.42u_{t-5} + 0.8y_{t-1} \_1.55y_{t-2} + 0.56y_{t-3}$$
$$-0.49y_{t-4} + e_t + 0.1e_{t-1} + 0.94e_{t-2} + 0.05e_{t-3} + 0.21e_{t-4} \quad (13)$$

a) The model when using sample size n=100

1. Some results were obtained, when using the impulse response function and with multiple repetitions as shown in the following figures:

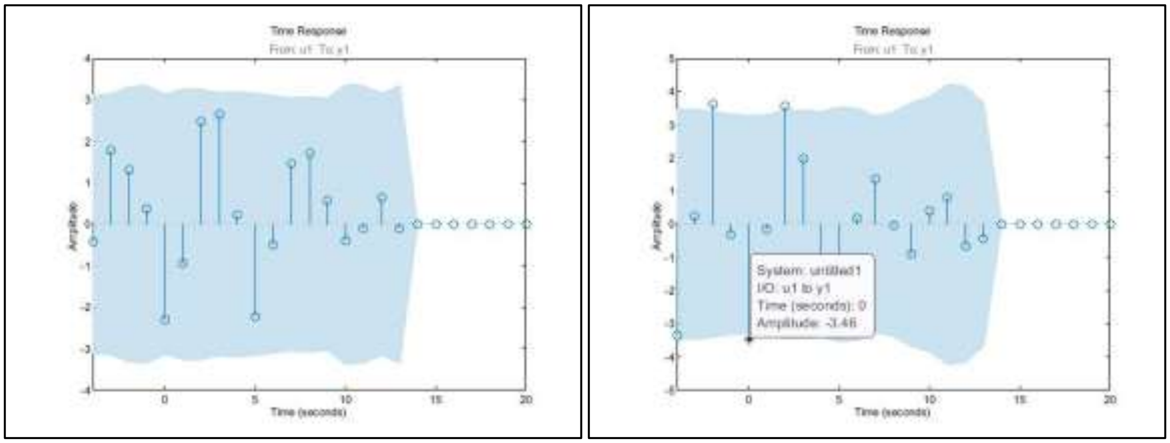


First repetition Second repetition

Figure (1): shows the impulse response function when sample size is n=100

We notice from the above figure that the impulse response function fails to estimate the delay time.

2- When using the cross-correlation function, we also note that it fails to estimate the delay time and as shown in the following figure:



First iteration Second iteration

Figure (2): shown the cross-correlation function when sample size is n=100

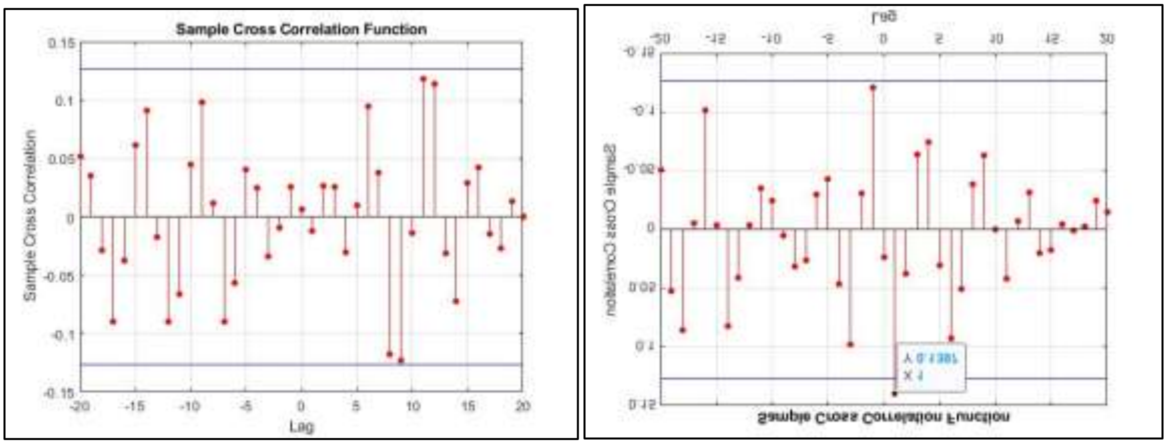
3-When using Linug method, and by reconciling many ARX models with the stability of the parameters $na = nb = 2$ with changing the value of the delay time$nk = 1,2,3, \dots ,10$ and choosing the model that corresponds to the lowest value of the statistical criteria, the success of this method appeared in estimate the delay time as shown in the following table (1):

Table (1): represents the fit of the ARX model with the statistical criteria for choosing the best delay time for one of the iteration when sample size is 100 = n

na	nb	nk	AIC	AIC_c	BIC	FPE	MSE
2	2	1	522.5172	523.6679	536.8093	40.2001	34.5908
2	2	2	522.1513	523.3020	536.4434	40.0166	34.4329
2	2	3	522.9432	524.0938	537.2353	40.4147	34.7754
2	2	4	522.51723	524.5436	537.6850	40.6426	34.9715
2	2	5	523.2845	524.4352	537.5767	40.5875	34.9242
2	2	6	523.6223	524.7730	537.9145	40.7593	35.0719
2	2	7	523.6524	524.8031	537.9446	40.7746	35.0851
2	2	8	523.9796	525.1303	538.2717	40.9417	35.2289
2	2	9	523.9235	525.0742	538.2157	40.9130	35.2042
2	2	10	523.0480	525.0742	537.3401	40.4677	34.8210

b) The model when using sample size is n=250

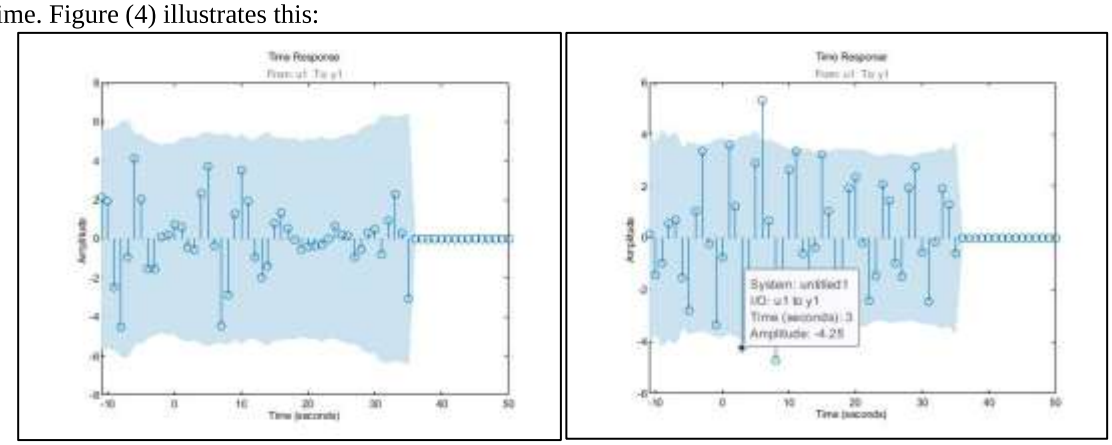
When using the data impulse response function and repeating the experiment (25) times, as the results showed the emergence of failure for this function in estimating the delay time, as shown in the following figures:



First iteration Second iteration

Figure 3: magnitude of the impulse response function of the model data of size n=250

2-When using the cross-correlation function, we also note that there are some failures are obtained of this function in estimating the delay time. Figure (4) illustrates this:



First iteration Second iteration

Figure(4) Amount of cross-correlation function of data model size n=250

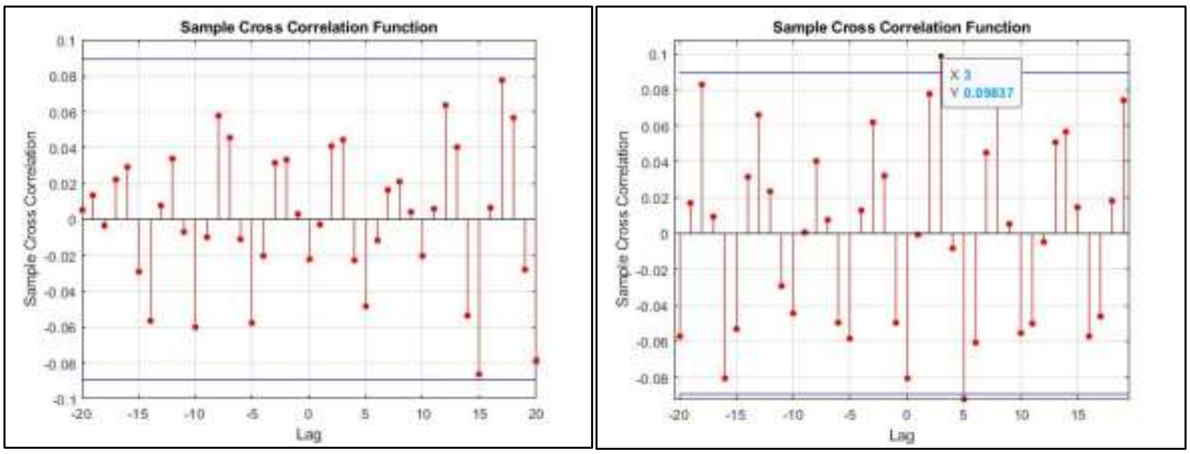
3- When using the Linug method and by reconciling many ARX models with the stability of parameters $na = nb = 2$ with changing the value of the delay time $nk = 1,2,3, \ldots, 10$ and choosing the model that corresponds to the lowest value of the statistical criteria, the success of this method appeared in estimating the time delay, as shown in the table.

Table (2): represents the fit of the ARX model with the statistical criteria for choosing the best delay time for one of the iterations when the sample size is n=250

na	nb	nk	AIC	AIC_c	BIC	FPE	MSE
2	2	1	1291.7753	1292.2106	1311.5652	37.3755	35.1983
2	2	2	1290.1293	1290.5646	1309.9192	37.0692	34.9098
2	2	3	1291.2147	1291.6499	1311.0046	37.2709	35.0998
2	2	4	1295.1577	1295.5929	1314.9476	38.0130	35.7987
2	2	5	1293.8047	1294.2399	1313.5946	37.7567	35.5573
2	2	6	1293.8222	1294.2575	1313.6121	37.7600	35.5604
2	2	7	1295.1685	1295.6037	1314.9584	38.0151	35.8006

2	2	8	1294.9700	1295.4052	1314.7599	37.9773	35.7651
2	2	9	1294.4930	1294.9282	1314.2829	37.8869	35.6799
2	2	10	1294.4522	1294.8875	1314.2421	37.8792	35.6726

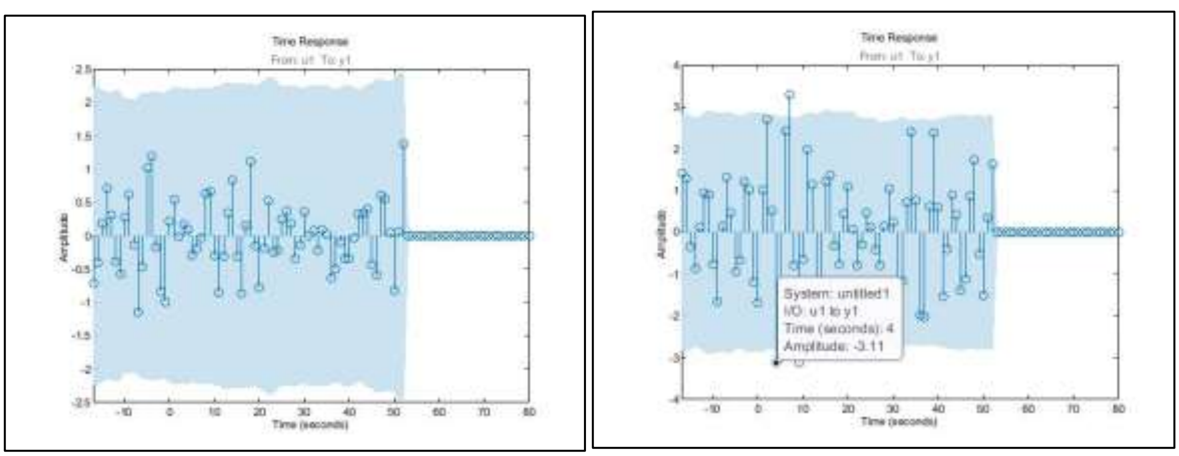
c) The model when using a sample size of $n = 500$
1- When using the impulse response function and by performing multiple repetitions exceeding 25 times, it was found that this function failed to estimate the delay time, as its first significant value outside the confidence interval is (lag=3), and the following figure shows some of these repetitions:



First iteration Second iteration

Figure (5) magnitude of the impulse response function of the model data of size n=500

2- When using the cross-correlation function, we also note that is failures in estimating the delay time, as shown in the following figure:



First repetition Second repetition

Figure (6): shown the cross-correlation function when sample size is n=500

3- When using Linug method, and by reconciling many ARX models with the stability of the parameters $na = nb = 2$ with changing the value of the delay time$nk = 1,2,3, \dots ,10$ and choosing the model that corresponds to the lowest value of the statistical criteria, the success of this method appeared in estimate the delay time as shown in the following table (3):

Table (3): represents the fit of the ARX model with the statistical criteria for choosing the best delay time for one of the iteration when sample size is $n = 100$

na	nb	nk	AIC	AIC_c	BIC	FPE	MSE
2	2	1	2693.1273	2693.3411	2717.0761	49.1532	47.7004
2	2	2	2690.7877	2691.0014	2714.7365	48.8666	47.4222
2	2	3	2693.5318	2693.7456	2717.4806	49.2030	47.7487
2	2	4	2696.1754	2696.3892	2720.1242	49.5292	48.0653
2	2	5	2700.3496	2700.5634	2724.2984	50.0488	48.5695
2	2	6	2697.0110	2697.2248	2720.9598	49.6328	48.1658
2	2	7	2698.1784	2698.3921	2722.1272	49.7779	48.3066
2	2	8	2699.7845	2699.9983	2723.7333	49.9781	48.5009
2	2	9	2698.7609	2698.9747	2722.7097	49.8504	48.3770
2	2	10	2700.5026	2700.7164	2724.4514	50.0679	48.5881

6. Delay time estimation using wavelet

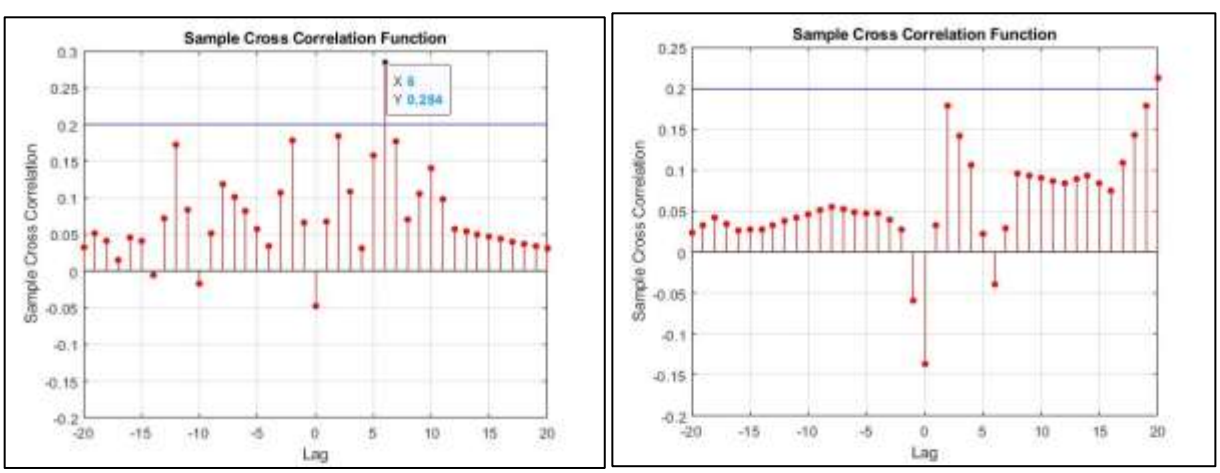
For the purpose of estimating the delay time with the wavelet, the data generated using the simulation will be processed in different sizes. The original (raw) data is entered into the ready-made MATLAB software in the form of a two-column matrix that includes the first column the input and the second column represents the output. The instructions for the wavelet are used, as shown in the following figures:

(Start ⟶ Toolboxes ⟶ More ⟶ Wavelet ⟶ Wavelet Toolbox Main Menu(Wave menu))

By using the sample sizes that were generated, the delay time was estimated by the methods mentioned in the theoretical side and with the wavelet Haar. The results were as follows:

a)The model when using sample size n=100

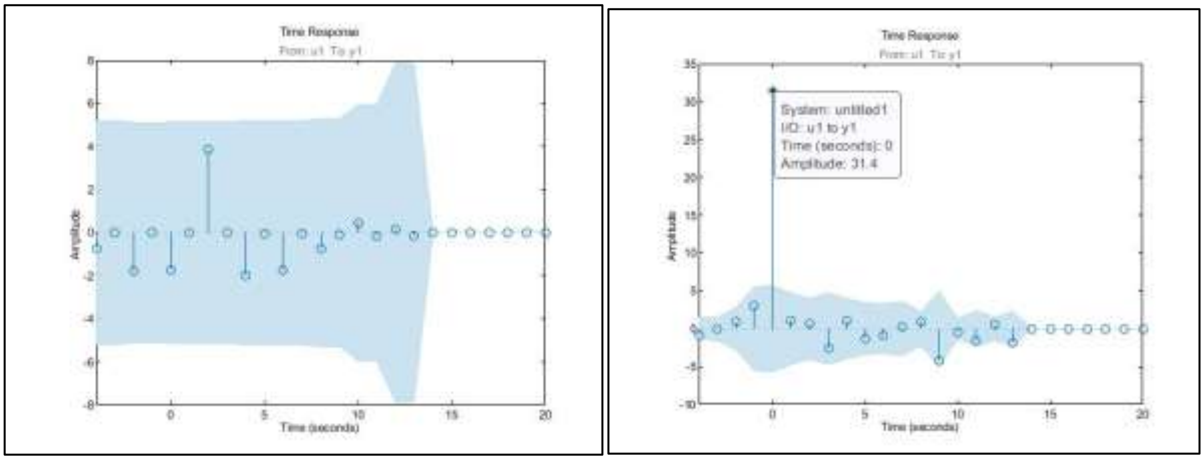
1.When using the impulse response function for wavelet data, the delay time was estimated for multiple repetitions, as it was found that this function failed to estimate the delay time, as shown in the following figures



First iteration Second iteration

Figure (7): The magnitude of the impulse response function of the wavelength Haar when n=100

1. When using the cross-correlation function of the wavelet Haar data, we also notice a great failure in estimating the delay time, as shown in the following figures:



First iteration Second iteration

Figure (8): shows the cross-correlation function by converting to wavelet Haar when $n = 100$

2. When using Linug method, and by reconciling many ARX models with the stability of the parameters $na = nb = 2$ with changing the value of the delay time $nk = 1,2,3, \dots, 10$ and choosing the model that corresponds to the lowest value of the statistical criteria, the success of this method appeared in estimate the delay time as shown in the following table:
3.

Table (4): represents the fit of the ARX model with the statistical criteria for choosing the best delay time for one of the iterations when sample size is $n = 100$

na	nb	nk	AIC	AIC_c	BIC	FPE	MSE
2	2	1	418.4550	419.6057	432.7472	10.9473	9.4197
2	2	2	418.3255	419.4762	432.6177	10.9296	9.4045
2	2	3	421.6571	422.8078	435.9492	11.3943	9.8044
2	2	4	426.4956	427.6463	440.7878	12.1047	10.4157
2	2	5	426.2142	427.3649	440.5064	12.0622	10.3791
2	2	6	425.7264	426.8771	440.0185	11.9889	10.3160
2	2	7	424.4925	425.6432	438.7847	11.8054	10.1581
2	2	8	424.0015	425.1522	438.2937	11.7332	10.0960
2	2	9	423.7994	424.9501	438.0916	11.7036	10.0705
2	2	10	424.1342	425.2849	438.4264	11.7527	10.1127

In the same way, simulation experiments were conducted for the wavelet data for the rest of the sample sizes, and we obtained a clear superiority for the autoregressive model with additional inputs in estimating the delay time and the failure of the rest of the methods.

7. **Conclusions**

The research reached some conclusions, including:

1.Most of the methods failed to estimate the delay time in the dynamic systems, including the impulse response method and the cross-correlation method for the simulation data before and after processing it using the wavelet.
2.When using the autoregressive model fit with additional inputs by fixing the parameters $na = nb = 2$ and taking several values for the delay time $nk = 1,2,3, \ldots, 10$, which is the method developed by the scientist Ljung, we noticed that this model achieved better results than the rest of the methods in estimating the delay time of the simulation data before and after using the wavelet.
3.We noticed some failures in estimating the delay time with different sample sizes in most of the methods used, and that the larger the sample size, the better the results.

References

1. Akaike ,H.(1981),"Likelihood of a model and Information Criteria", Journal of Econometrics , 16 .(30-14).
2. Box ,G.,Jenkins,G., Reinsel , G., and Ljung G.,(2016):"Time Series Analysis Forecasting and control", John wiley &Sons ,Inc ,Hoboken ,New Jersey.
3. Burnham ,K.P., and Anderson , D.R. (2002):"Model Selection and Multimodel Inference: A Practical Information- theoretic approach",2nd – edition Springer-Verlag ,New York.
4. Green,O.,& Silverman,B., (1994):" Nonparametric Regression and Generalized Linear Models : ARoughness Penalty Approach" , Chapman and Hall ,London.
5. Gunther ,U.,(2001), Advanced NMR Processing ,Eurolab Course:"Advance Computing in NMR Spectroscopy", Florence,pp6-9.
6. Hammoudat, Alaa Abdul-Sattar, (2020): "Using some wavelet reduction techniques and immunization methods to improve the efficiency of the capabilities of the generalized aggregate model" PhD thesis (unpublished) College of Computer Science and Mathematics, University of Mosul, Iraq
7. Heyam A. Hayawi, Najlaa Saad Ibrahim Alsharabi,(2022),"Analysis of multivariate time series for some linear models by using multi-dimensional wavelet shrinkage",Periodicals of Engineering and Natural Sciences 10 (4), 120-129. https://xjournals.com/collections/articles/Article?qt=7PzIpVggCvea+kXIRpdERyKclJE59lgkRqB5N4lmpYCnJ33rwBcEfxmhUhdkNsmyomnZWqDC91eL5xMLv7CSBA==
8. Heyam A. Hayawi,(2022),"Using wavelet in identification state space models",Int.J. Nonlinear Annl.Appl. 1 (13), 2573-2578. https://ijnaa.semnan.ac.ir/article_5958_3214a7502075753bc47161edba0d2ae5.pdf
9. Ljung,L.,(1995):"System Identification Toolbox for use with MATLAB", The Mathworks Inc.Natick Massachusetts.
10. Ljung,L.,(1999):"System Identification Theory for user ",2nd ed.Prentice Hall Upper Saddle River N.J. London UK.
11. Ljung,L.,(2004):):"System Identification Toolbox for use with MATLAB",6.0 Methworks Inc.
12. Makridakis,s.,Wheelwright,S.,&Mcgee,E.,(1983):"Forecasting :Methods and Applications" , 2nd ed . John Wiley &Sons New York U.S.A.
13. Mao,Yinian,(2004):"Wavelet Smoothing ",EcE, University of Maryland
14. Muhammad, Shaima Mahoud: "Possibility to Analyze Traffic Congestion in Karbala Governorate, Master's Thesis (published), College of Administration and Economics, University of Karbala, Iraq.
15. Wei, William W.S. ,(2006), :"Time Series Analysis Univariate and Multivariate Method ",2th ,Pearson Education ,Inc.

تقدير زمن التأخير في النظم الحركية الخطية باستعمال المويجات

ريم طلال طه[1] ، هيام عبد المجيد حياوي [2] ، محمد احمد الخولي[3]

[2,1]قسم الإحصاء والمعلوماتية ، كلية علوم الحاسوب والرياضيات ، جامعة الموصل ، الموصل ، العراق.

[3] أكاديمية السادات للعلوم الإدارية (SAMS)، قسم الرياضيات والإحصاء، القاهرة، مصر .

الخلاصة: يتناول هذا البحث استخدام المويجات في تقدير وقت التأخير في الأنظمة الديناميكية الخطية العشوائية ، حيث أن وقت التأخير له أهمية كبيرة في تشخيص النظام لمعرفة الفترة الزمنية التي تسببها المدخلات إلى المخرجات. تم إجراء العديد من تجارب المحاكاة وتم استخدام أحد أنواع المويجات ، وهو مويجات هار ، في معالجة البيانات ، ثم تم تطبيق بعض طرق تقدير وقت التأخير وتمت مقارنة النتائج. كما وجد أن تقدير وقت التأخير باستخدام المويجات هار حققت نتائج أفضل باستخدام نموذج الانحدار الذاتي مع مدخلات إضافية مما كان عليه في البيانات قبل معالجتها.

الكلمات المفتاحية: المويجات ، المويجات هار ، زمن التأخير ، الصندوق الأسود.