Heart Signal Acquisition Based System Autoregressive Identification Models

Ismail M. Albatrookh
Electrical and Electronic Engineering Department, Misurata University
Misurata, Libya
ismailelbat@eng.misuratau.edu.ly

Ahmed A. Baaiu
Electrical and Electronic Engineering Department, Misurata University
Misurata, Libya
ahmed.baaiu@eng.misuratau.edu.ly

Mustafa A. Elsherif
Electrical and Electronic Engineering Department, Misurata University
Misurata, Libya
m.a.elsherif@misuratau.edu.ly

Almabrok Essa
Electrical and Computer Engineering Department, University of Dayton
Dayton, Ohio, USA
essaa1@udayton.edu

Abstract—**A system identification (SI) model can be constructed without any prior knowledge of the nature or physics of the relationship that has been involved. It is therefore appropriate to examine the question of (heart rate). In this paper, a simple and efficient hardware design is implemented to acquire the heart signal with the help of several linear models of SI. The relationships between different system identification models are discussed with detailed justification of the aid of these typical types. And then characterize the methods that fit the system structure to measure data input and output, as well as the most basic characteristics of the resulting models. For evaluation, compare between SI models to validate the results.**

Index Terms: **Autoregressive; System Identification Models; Box Jenkins; moving average; Output Error; Heart Rate**

I. INTRODUCTION

System identification (SI) plays the crucial role in decision making, forecasting, fault detection, pattern recognition, and many others. The field of SI uses statistical methods to build mathematical models for dynamical systems from measured data. The purpose of the system identification is to learn the general law of things, and solve practical problems by using the identification results [1]. The main objective of this process is, accessing to the model of the system and extraction or understanding the laws and relations between attributes without any prior knowledge about the system under consideration.

There are several ways to identify the mathematical systems which depend on the neural network, of which depend on the approach as the statistical correlation analysis of the modalities, etc. The models obtained from this process usually called external models as they do not have to consider the internal dynamics of the phenomenon at the exits of others [2].

Received 31 October 2017; revised 19 November 2017; accepted 22 November 2017.

Available online 23 November 2017.

Currently, the algorithms of the system identification mainly include least squares algorithm, neural network algorithm, genetic algorithm and swarm intelligence algorithm. Custom systems used in almost all fields of engineering such as aeronautical engineering, electrical engineering, and information technology are relay on the system identification process. Usually system identification is divided to parametric and nonparametric identification models. The parametric identification is a way when information about the structure of the system is available. The nonparametric identification is a way when there is nothing about the structure of the system. In other words, it is customary to color code in shades of grey, the model structure according to what type of prior knowledge has been used [3]:

- White-box models: This is the case when a model is perfectly known, which has been possible to construct it entirely from prior knowledge and physical insight.
- Grey-box models: This is the case when some physical insight is available, but several parameters remain to be determined from data observation. It is useful to consider two sub-cases: 1) Physical modeling which a model structure can be built on physical ground that has a certain number of parameters to be estimated from data. 2) Semi-physical modeling which physical insight is used to suggest certain nonlinear combinations of measured data signal.
- Black-box models: No physical insight is available or used, but the chosen model structure belongs to families that are known to have good flexibility and have been successful in the past.

From the observation that the interactions of organism with the environment and its internal functions are connected with electromagnetic phenomena. Therefore, the study of electromagnetic phenomena, the biological, physical and technical sciences could bring an important contribution to achieve some improved technical devices. Meanwhile, to understand the biological phenomena it need to had the ability of some theoretical interpretations, as well as the achievement of some investigation methods on which the technique can hold or can produce its [4].

The rest of the paper is organized as follows. Section 2 presents the linear black box models in system identification. The hardware implementation and the measurement of heart rates is explained in Section 3. The discussion on the estimation models and results are resented in Section 4.

II. LINEAR BLACK-BOX MODELS

Linear black box models are parametric models that describe a system in terms of differential equations and transfer functions. This provides insight into the system physics and a compact model structure. Generally, they can describe a system using the following equation [5], which is known as the general-linear polynomial model or the general-linear model.

$$y(n) = q^{-k}\ G(q^{-k}, \theta)u(n) + H(q^{-k}, \theta)\ e(n) \qquad (1)$$

Where *u(n)* and *y(n)* are the input and output of the system respectively. *e(n)* is zero-mean white noise, or the disturbance of the system. $G(q^{-1}, \theta)$ is the transfer function of the deterministic part of the system. $H(q^{-1}, \theta)$ is the transfer function of the stochastic part of the system.

The deterministic transfer function specifies the relationship between the output and the input signal, while the stochastic transfer function specifies how the output is affected by the disturbance. Some literatures refer to the deterministic and stochastic parts as system dynamics and stochastic dynamics, respectively. The general linear model structure, shown in Figure. 1, provides flexibility for both the system dynamics and stochastic dynamics. However, a nonlinear optimization method computes the estimation of the general-linear model. This method requires intensive computation with no guarantee of global convergence. Simpler models that are a subset of the general linear model structure are possible. By setting one or more of *A(q), B(q), C(q),* or *D(q)* polynomials equal to 1 you can create these simpler models structure as can be seen in Figure. 2. Each of these methods has their own advantages and disadvantages and is commonly used in real world applications. Therefore, for any problem the choice of the model structure depends on the dynamics and the noise characteristics of the system. Using a model with more freedom or parameters is not always better as it can result in the modeling of nonexistent dynamics and noise characteristics. This is where physical insight into a system is helpful.

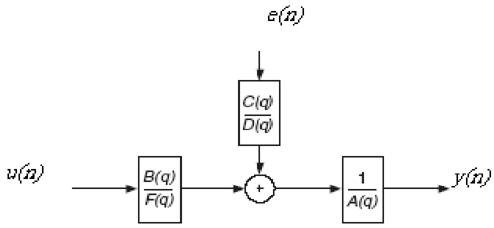

Figure 1. General-Linear Model Structure.

A. Empirical transfer functions Estimated

The empirical transfer function estimation (ETFE) is computed as the ratio of the output Fourier transform to the input Fourier transform, using fast Fourier transform (FFT). The periodogram is computed as the normalized absolute square of the Fourier transform of the time series. ETFE estimates the transfer function (ge) of the general linear model. The smoothed version (ges) of this model is obtained by applying a Hanning window to the output fast Fourier transform (FFT) times the conjugate of the input FFT, and to the absolute square of the input FFT, respectively. Then subsequently forming the ratio of the results.

B. Autoregressive Model

The autoregressive (AR) model specifies that the output variable depends linearly on its own previous values and on a stochastic term (an imperfectly predictable term). Therefore, the model is in the form of a stochastic difference equation.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \qquad (2)$$

$$A(q)y(q) = e(n) \qquad (3)$$

C. Autoregressive With Exogenous Input Model

The model is in the form of a stochastic difference equation.

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{b} \eta_i\ d_{t-i} \qquad (4)$$

Where $\eta_1, \dots \dots \dots, \eta_b$ are the parameters of the exogenous input dt .

When C(q);D(q); and F(q) are equal 1, the general linear polynomial model reduces to an autoregressive with exogenous (ARX) model. This model contains the AR and a linear combination of the last b terms of a known and external time series function [5]. The ARX model structure can be seen in Figure. 2.

$$A(q)y(q) = q^{-k}\ B(q)u(n) + e(n) \qquad (5)$$

D. Autoregressive Moving Average With Exogenous Input

The autoregressive moving average with exogenous input (ARMAX) model with p autoregressive terms, q moving average terms, and b exogenous inputs terms. This model contains the AR, moving average (MA) models, and a linear combination of the last b terms of a known and external time series. The ARMAX model can be created by setting D(q); and F(q) to 1 as seen in Figure. 2

$$X_t = \varepsilon_t + \sum_{i=1}^{p}\varphi_i X_{t-i} + \sum_{i=1}^{q}\theta_i \varepsilon_{t-i} + \sum_{i=1}^{b}\eta_i d_{t-i} \qquad (6)$$

Where $\eta_1, \dots \dots \dots, \eta_b$ are the parameters of the exogenous input dt .[5]

$$A(q)y(q) = q^{-k} B(q)u(n) + C(q)e(n) \qquad (7)$$

E. Box-Jenkins Model

Estimating the parameters for Box–Jenkins models involves numerically approximating the solutions of nonlinear equations. For this reason, it is common to use statistical software designed to handle to the approach fortunately, virtually all modern statistical packages feature this capability. When A(q) equals 1, the general-linear polynomial model reduces to the Box-Jenkins model as shown in Figure. 2. The following equation describes a Box-Jenkins model

$$y(n) = \frac{q^{-k}B(q)}{F(q)} u(n) + \frac{C(q)}{D(q)} e(n) \qquad (8)$$

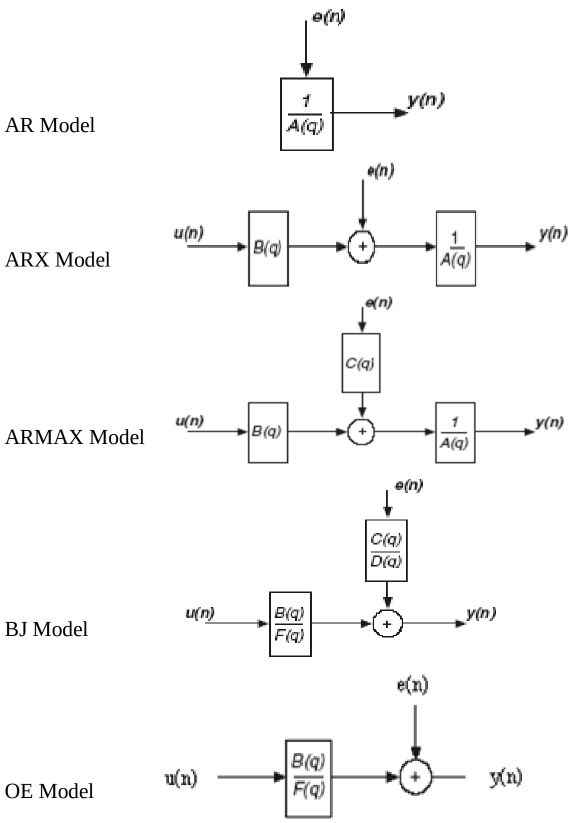

Figure 2. The Different Linear Black-Box Models.

F. Output-Error Model

The output error (OE) method deals with minimizing an objective function, usually a quadratic criterion which is based on the output error OE. This is the error between the measured output y of the system and the output ^y of the transfer function model. The estimation is performed using either time or frequency domain data. The estimated model is delivered as polynomial model with identifiable parameters object. Model contains the estimated values for two polynomials along with their covariance and structure information [5].

$$y(n) = \frac{q^{-k}B(q)}{F(q)} u(n) + e(n) \qquad (9)$$

When A(q);C(q); and D(q) are equal 1, the general-linear polynomial model reduces to the output-error model as shown in Figure. 2.

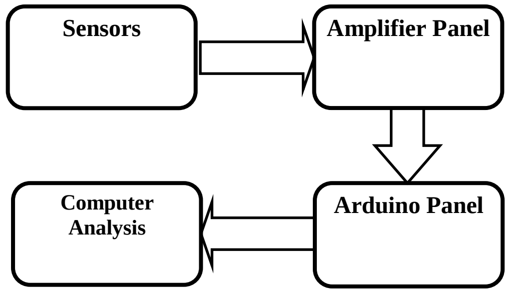

Figure 3. Block Diagram of Assembling the Heart Signal.

III. HARDWARE IMPLEMENTATION & MEASUREMENT OF HEART RATE

Figure. 3, describes the block level presentation of the signal processing system. The heart signal was transferred by connecting three sensors to the human body with the amplifier panel. The amplifier panel signal is brought to the user (computer) via the Arduino panel using the signal transmission cables. The simplified electronic design of the amplifier board shown in Fig. 4. Fig. 5, shows a real heart signal that has been extracted using this system. The main output equation of the board can be defined by

$$V_{out} = (V_2 - V_1)\left[1 + \frac{2R_2}{R_1}\right]\left(\frac{R_4}{R_3}\right) \qquad (10)$$

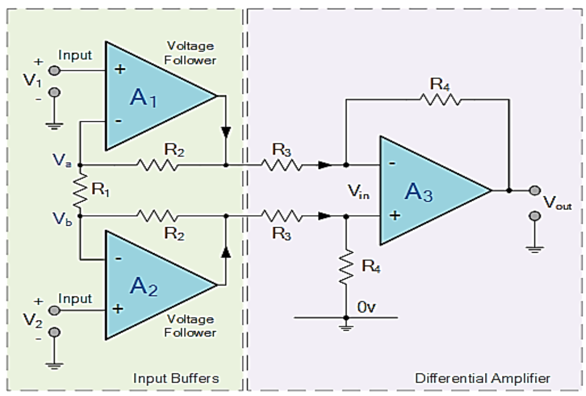

Figure 4. Signal Amplifier Panel

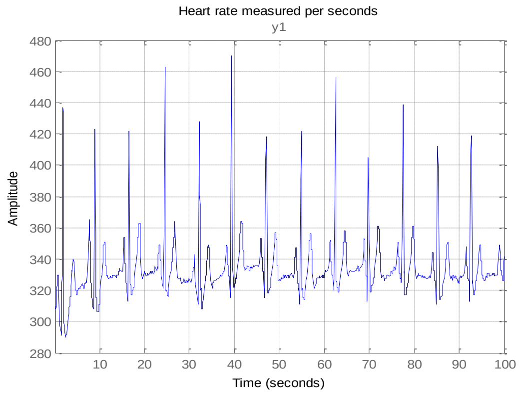

Figure 5. Heart Rate Measured per Seconds n Taken by the Amplifier Panel.

IV. ESTIMATION MODELS & RESULTS

A non-parametric analysis of the signal is performed using an autoregressive function for modeling the signal. Tools for choosing a reasonable model order are then discussed along with the use of autoregressive for signal modeling [6]. Methods for fitting a model to only a chosen range of harmonics are also discussed. Signals can be considered as the impulse response of an autoregressive linear model, and can thus be modeled using autoregressive model. Data for signals can be encapsulated into identification objects. Standard identification Models are used to estimate the characteristics of the output-only data. These models are assessed for their spectral estimation capability, as well as their ability to predict the future values of the signal from a measurement of their past values.

A. ETFE Model Estimation

The ETFE is estimated on time domain data (ge), This periodogram reveals several harmonics as shown in Figure. 6, which is not very smooth. A smoothed version of ge (ges) is obtained by applying a Hanning window to the output which can be seen in Figure. 7. Figure. 8, shows the comparison of the (ges) with the spectral analysis of the original data (SPA). The Hann window is used to compute the spectral amplitudes (as opposed to ETFE which just computes the raw periodogram). Note that a very large window will be required to see all the fine resonances of the signal, which needs a more sophisticated model that can be provided by parametric autoregressive modeling techniques.

B. AR, ARX, ARMAX Models Estimation

Another estimation called parametric modeling of the heart signal which used to compute the spectra by parametric AR methods. In this case, models of 2^{nd} and 4^{th} order are obtained by

$$A(z)\,y(t) = e(t). \quad (11)$$

From linear spectrum analysis shown in Figure. 9 .The parametric spectra are not capable of picking up all the harmonics. The reason is that the AR-models attach to much attention to the higher frequencies, which are difficult to model. Therefore, to go to high order models before the harmonics are picked up. The orders up to 30 are checked to determine the best order for the calculation. Figure. 10 shows the best AR order selection strategy.

As Figure. 10 shows, there is a dramatic drop at the number of parameters equals 10 (n=10). All the harmonics are now picked up, but the model has the level dropped that is caused by the 10^{th} order, which contains very thin and high peaks. With the crude grid of frequency points, simply don't saw the true levels of the peaks. Figure. 11 shows comparison between 2^{nd} order AR (t2), 4^{th} order AR (t4), 10^{th} order AR (tnn) and smoothed ETFE using SPA (ges) models.

When it comes to the ARX and ARMAX models, the estimation parameters of ARX and ARMAX models yields discrete-time ARX and ARMAX models. Fig. 12 shows the comparison of AR model (tnn), ARX model (tx), and the actual output (gs). While Fig. 13 shows the comparison between 2^{nd} order ARMAX (tmx2), 4^{th} order AR (tmx4) and the actual output (gs).

Box-Jenkins and Output-Error structure do not apply to time series data, so at the last it can compare between all models to choose the best method that represent the heart signal from this study.

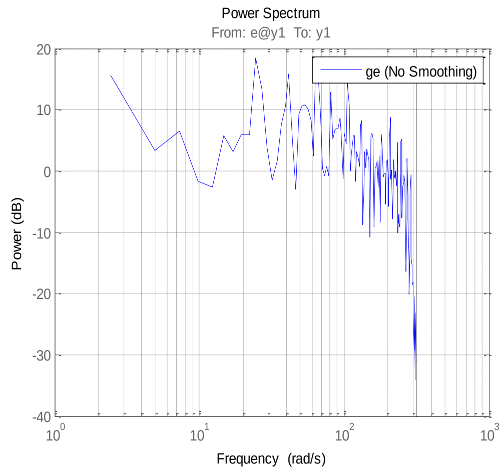

Figure 6. Spectrum Analysis by ETFE.

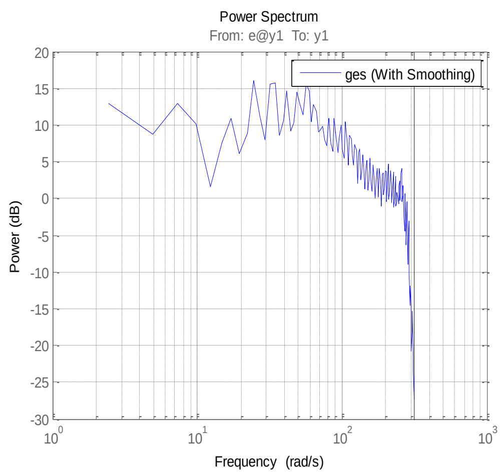

Figure 7. Smoothing Spectrum Analysis by ETFE.

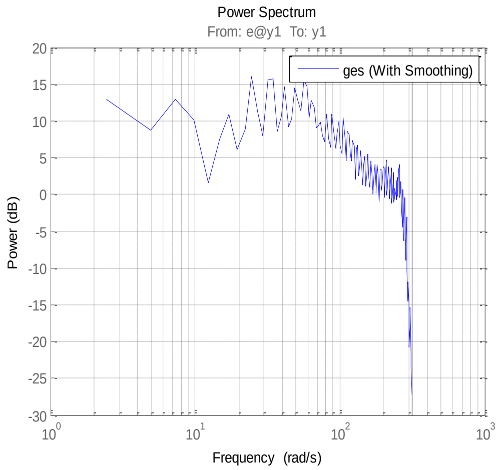

Figure 8. Comparison between ETFE and Data Spectrum.

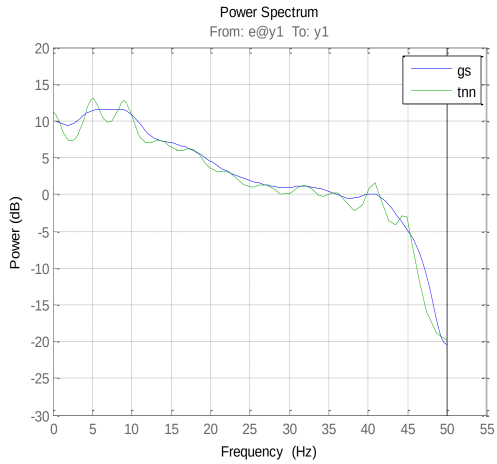

Figure 11. Comparison between 2nd Order AR (t2), 4th Order AR (t4), 10th Order AR (tnn) and the Actual Output (ges).

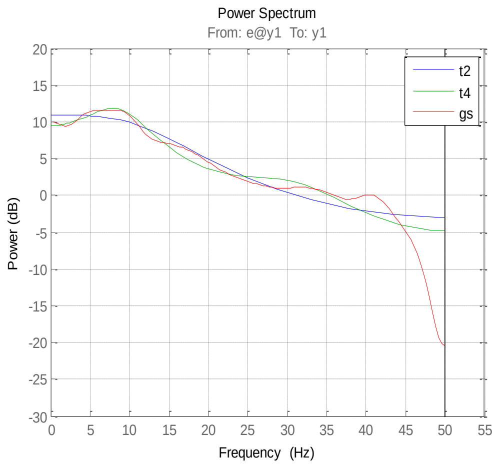

Figure 9. Comparison between 2nd Order AR (t2), 4th Order AR (t4) and the Actual Output (gs).

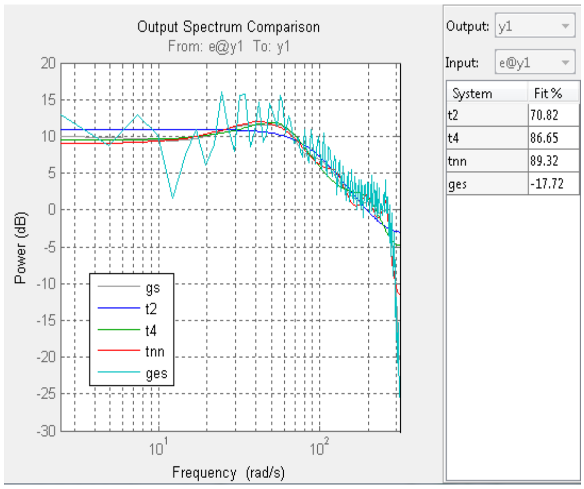

Figure 12. Shows Comparison between 2nd ,4th,10th Order and Smoothed ETFE Models

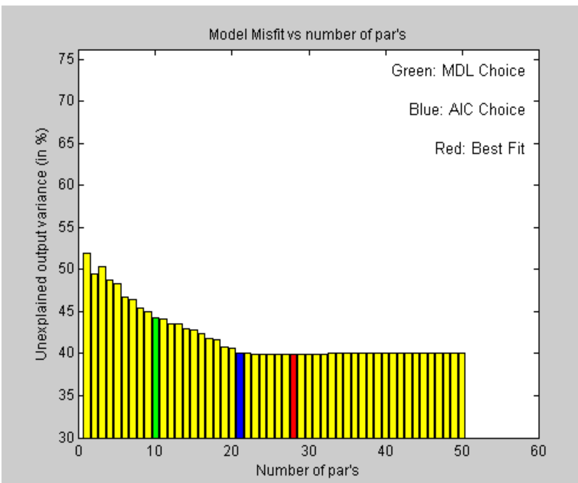

Figure 10. The Best AR Order Selection Strategy.

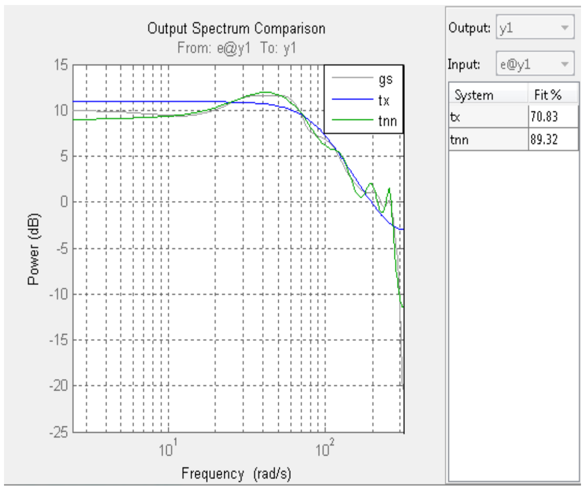

Figure 13. Comparison of AR Model (tnn), ARX Model (tx), and the Actual Output (gs).

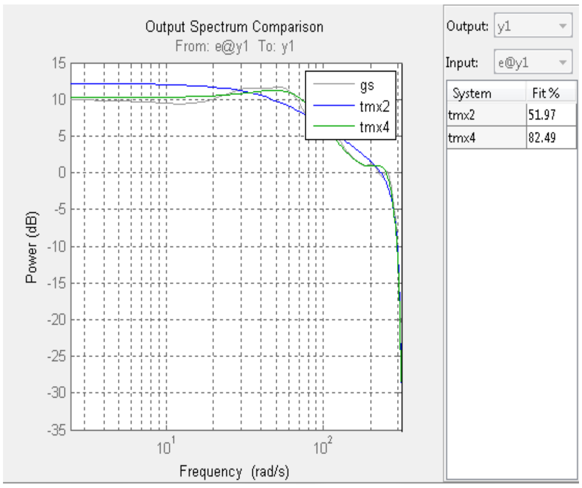

Figure 14. Comparison between 2nd Order ARMAX (tmx2), 4th Order AR (tmx4) and the Actual Output (gs).

V. CONCLUSION & FUTURE WORK

In this paper, the heart signal is estimated with several linear models of SI and a simple/efficient hardware design. Experiments performed using different types of SI models on real-world heart signal data demonstrated the effectiveness of the proposed SI models that is evidenced by yielding promising approximation accuracy of the heart rate. The approximation accuracy was 89.32% using the AR model, 67.88% using the ARX model, and 82.49% using the ARMAX model. Even though, AR model with the 10th order has provided the best approximation of the original data, that does not mean AR is the best model forever, since the accuracy of the model construction depends on the nature of the signal. The work is in progress to use the BJ and OE models as well as nonlinear SI models.

REFERENCES

[1] J. Zhai, C. Zhu, C. He, Z. Yao, and Y. Dai, "*A system identification method to hammerstein model based on modified shuffled frog leaping algorithm*," in *Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2017.

[2] A. Essa and I. Ibrahim, "*Linear black box modeling in system identification*," in *International Conference on Electrical Engineering Design and Technologies (ICEEDT09)*, 2009.

[3] L. Milici, S. G. Penifuc, and M. Milici, "*Mathematical analysis applied for of heart signals*," in *IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS*, 2007.

[4] L. Ljung, "*Mathematical analysis applied for of heart signals*," in *IEEE Instrumentation and Measurement Technology Conference, IMTC*, 2001.

[5] E. Tohme, R. Ouvrard, T. Poinot, J.-C. Trigeassou, and A. Abche, "*Initialization of output-error identification methods comparison between arx and rpm models*," *IFAC Proceedings Volumes*, vol. 42, no. 10, pp. 302–307, 2009.

[6] I. Kollr, R. Pintelon, and J. Schoukens, "*Frequency domain system identification toolbox for matlab*," *IFAC Proceedings Volumes*, vol. 24, no. 3, pp. 1243–1247, 1991.

BIOGRAPHIES

Ismail Albatrookh was born in Misurata, Libya, on 6th of June1977. In 2000 he received his B. Sc from the Department of Electrical Engineering, Faculty of Engineering, OMAR ALMUKHTAR university, ELBEIDA, Libya. His B. Sc project was focused on study simplified model of A direct sequence spread spectrum communication system. In 2008 he received his M.Sc. from Tripoli university, Tripoli, Libya. His master subject was entitled"Using Artificial Neural Network to Control an Autonomous Underwater Vehicle". On September 2009 he appointed as a lecturer Assistant at the department of Electrical and Electronic Engineering at the Faculty of Engineering at Misurata University, Misurata, Libya. From August 2013 he promoted to be an Lecturer at Misurata University, Misurata, Libya.

Ahmed Baaiu was born in Misurata, Libya, on March 1976. In 1998 he received his B. Sc from the Department of Electrical and Computer Engineering, Faculty of Engineering, NASSER university, Alkhoms, Libya. His B. Sc project was focused on the Analysis of control systems with a delay time effects. In 2003 he received his M.Sc. from Lyon 1 university, Lyon, France. His master subject was entitled "Studying a reference model behavior using the Ramadge and Whonam theory for the synthesis of control laws for complex discrete events systems". In 2007 he received his Ph.D. in automatic control (section 61: computer and control) from Lyon 1 university, Lyon, France. His Ph.D. subject was entitled "Hamiltonian Approach for modeling, estimation and control of a separation process". On September 2007 he appointed as a lecturer at the department of Electrical and Electronic Engineering at the Faculty of Engineering at Misurata University, Misurata, Libya. From February 2011 he promoted to be an Assistant Professor at Misurata University, Misurata, Libya.

Almabrok Essa received his BS degree in electrical and electronic engineering from Sirte University, Libya, and his MS degree in electrical and computer engineering from the University of Tripoli, LY. He recently received his PhD in electrical and computer engineering from the University of Dayton, Dayton, OH, USA. He has authored and coauthored more than 40 publications, including refereed journals and conference proceedings in the field of computer vision, image processing, pattern recognition, machine learning, and remote sensing. Dr. Essa also serves as a technical reviewer for several signal processing journals and international conferences. He is a member of IEEE and SPIE.

Mustafa Elsherif was born in Misurata /Libya, on September 5, 1981. He received B.Sc. degree in Electrical and Electronics from University of Sirt, in 2003. He got M.Sc. degree in Engineering from Nottingham Trent University /UK in 2007. Moreover, he got PhD degree in Electrical Power Systems from Durham University/UK in 2013, where he is currently assistant professor in Department of Electrical Engineering at Misurata University / Libya. His research field is applied power systems control and Superconductor Power Systems.