International Journal of Industrial Engineering & Production Research December 2022 Vol. 33, No. 4: 1-16
DOI: 10.22068/ijiepr.33.4.4

RESEARCH PAPER

Nonlinear System Parameterization and Control using Reduced Adaptive Kernel Algorithm

Rakesh Kumar Pattanaik[1], Muhammad Sarfraz[2] & Mihir Narayan Mohanty[3*]

Received 16 August 2022; Revised 13 September 2022; Accepted 3 October 2022;

ABSTRACT

To develop a system for specific purpose, it needs to estimate its parameters (parameterization). It can be used in different fields like engineering, medicine, industry etc. In this work, authors used adaptive algorithm to model a system that is applicable in industry for control. The model is non-linear and works on kernel-based estimation with Least-mean square (LMS) algorithm. The kernels are verified with Polynomial and Gaussian. As the system is nonlinear, polynomial kernel-based algorithm fails to prove its efficacy, though it is of low complexity approach. Gaussian kernel-based application for nonlinear system control performs better as compared to polynomial kernel. Further, the complexity is reduced and used with Gaussian kernel in LMS algorithm for better performance. The result proves its performance in form of MSE, MAE, RMSE for identification and control that is very useful in industrial application. With the use of reduced Gaussian kernel application, the MSE, MAE and RMSE are found to be -54.622 dB, 0.0362,0.235 respectively, in 0.01062 sec that shows the time consumption is very less compared to other approaches.

KEYWORDS: *Kernel adaptive filtering; Nonlinear system Identification; Least-mean square; Kernel least-mean square; Single input Single-output (SISO) System; Polynomial kernel; Gaussian kernel.*

1. Introduction

In the field of engineering, one of the major challenges is the identification of unknown complex systems [1]. To solve the system identification problem, various statistical and start-of-art methods have been introduced. The adaptive filter has become a popular method to solve various applications of statical system identification problems [2, 3, 4, 5]. In recent years, one of the most well-liked study areas in adaptive signal processing has been kernel adaptive filtering, which was developed at the interface of machine learning and statistical signal processing. To improve the performance of the model, continuous and automatic adjustments are made to the coefficients related to the adaptive filter. An adaptive filter reduces errors by putting the initial data into a format with more dimensions, or even an infinite number of dimensions. A lot of development has been done on kernel methods to reduce the computational complexity of the model. Reproducing kernel Hilbert spaces (RKHS), different types of kernel adaptive filters (KAF), are the improved versions of the basic kernel methods used in solving nonlinear applications, including system identification and control. The design of adaptive filters consists of three main specifications: first, the algorithm; second, the filter structure; and third, application. Given the benefits of kernel adaptive filters in signal processing, adaptive filter algorithms are used successfully in a variety of fields including communications, radar, sonar, seismology, and biomedical engineering. Estimation error is obtained from the process by differentiating between a desired response and an input vector. A popular method named the least mean square (LMS) algorithm was introduced by Widrow and Hoff to identify nonlinear systems, which generates a mean square error [6, 7]. The aim is to design a kernel least mean square model which learns as an adaptive filter to identify and

* **Corresponding author:** ***Mihir Narayan Mohanty***

mihirmohanty@soa.ac.in

1. *Department of Electronics and Communication Engineering, ITER, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India.*
2. *Department of Information Science, Kuwait University, Kuwait.*
3. *Department of Electronics and Communication Engineering, ITER, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India.*

control nonlinear systems. Estimation errors are controlled by a set of filter coefficients that can be changed based on an input vector and the response that is desirable.

Identification and control problems have been one of the major topics of research for several years. Zhou et al. introduce the M-estimation function by modifying the existing LMS algorithm, which is named as the least mean M-estimate (LMM) algorithm and hyperbolic secant LMS (HSLMS) algorithm, further modified and used in different fields [8, 9, 10, 11]. The convergence speed can be increased by implementing the Set of Membership Filtering (SMF) Strategy [12]. The fractional least mean square (LMS) method is suggested in [13] to estimate the parameters of the Hammerstein system with external noise (HN-ARMAX). In order to generate linear-in-parameter models, HN-ARMAX systems are parameterized. The fractional LMS algorithm is then used to adjust unknown parameter vectors. By adapting the unknown parameter vector using a novel fractional least mean square (FLMS) algorithm for a BJ system, the parameters are estimated [14]. For realistic modelling of the Volterra system, Kohli et al. used a RLS technique with a numeric variable forgetting factor [15]. A decomposition method is proposed by implementing an Autoregressive Moving Average (ARMA) nonlinear model for noise reduction and identification of Hammerstein systems with a FIR system to solve the nonlinear system identification problem [16]. In [17], a novel data filtering technique based on maximum likelihood is proposed for parameter estimating of a bilinear system. Volterra [18], Walsh functions [19], Laguerre transform [20], Hammerstein [21], polynomial autoregressive (PAR) [22], Recursive least squares (RLS), WLMS algorithm [23] and Wilcoxon algorithm [24,25] are the most commonly used models for system identification. The input to the linear adaptive filter is a nonlinear mapping to a reproducing kernel-equipped Hilbert space. Even when the dimensionality of the converted input in Hilbert space (H) is unlimited, as it is in the case of a Gaussian kernel, the order of the linear adaptive filter can be constrained provided an appropriate input sparsification technique is employed, according to the framework described previously. By using the above concept as reference, new kernel algorithms like the kernel least-mean-square (KLMS) algorithm [26], the affine projection (KAPA) algorithm [27,28], and mixed kernels [29] have been developed. Various application-oriented work which includes Parameter estimation and Identification in different areas are presented in [30, 31, 32, 33, 34].

For the majority of kernel-based techniques, the Gaussian kernel function is frequently used. The performance of KAF algorithms is ultimately influenced by the sharpness of the Gaussian function fitting the objective function, which is directly determined by the kernel bandwidth.

One component of the Gaussian kernel, Taylor extended to polynomial form has an intentionally mapped kernel function. To approximate Gaussian kernel function, the finite-order Taylor expansion series components are intercepted and multiplied by another factor. For each term of the reduced Gaussian kernel polynomial, there is a clear feature mapping function that lets the right iterative weight operations be done [35]. Table.1 shows the comparisons of different system identification techniques.

Tab. 1. Comparison of different System Identification techniques

Year	Title, Authors	Work done
Ref [12], 2014	Identification of Hammerstein nonlinear ARMAX systems using nonlinear adaptive algorithms. Naveed Ishtiaq Chaudhary · Muhammad Asif Zahoor Raja	The design schemes consist of parameterization of HN-ARMAX systems to obtain linear-in-parameter models and to use fractional LMS algorithm for adapting unknown parameter vectors.
Ref [16], 2018	Filtering-Based Maximum Likelihood Gradient Iterative Estimation Algorithm for Bilinear Systems with Autoregressive Moving Average Noise. Meihang Li1 · Ximei Liu1 · Feng Ding.	The maximum likelihood principle with the data filtering technique for parameter estimation of bilinear systems with autoregressive moving average noise is proposed. The proposed method can be extended to other linear and nonlinear systems with different structure, which achieve good accuracy.
Ref [19], 2008	Nonlinear system identification using Wiener type Laguerre—Wavelet network model. P. Aadaleesan, Nitin Miglan, Rajesh	The use of Laguerre basis filters coupled with a wavelet network in Wiener type model structure. Laguerre filter models have the ability to approximate linear systems (even with time delay)

		Sharma, Prabirkumar Saha.	with a model order lower than the traditional ARX (e.g., FIR, AR) modelling.
Ref 2018	[21],	Adaptive Model Predictive Control for Wiener Nonlinear Systems Ibrahim Aliskan	A model predictive control system with online identification support has been developed with the combination of MPC algorithm with online identification which constitutes an adaptive model predictive control algorithm. The prominent feature of online system identification may be referred to as accommodating easily to severe changes in system parameters.
Ref 2012	[22],	Analysis of outliers in system identification using WLMS algorithm. Sidhartha Dash, Mihir Narayan Mohanty,	This paper aim towards the Wilcoxon approach, Least Mean Square Algorithm been applied for System Identification problem with Gaussian noise.
Ref 2012	[23],	Variable Sign-Sign Wilcoxon Algorithm: A Novel Approach for System Identification. Sidhartha Dash, Mihir Narayan Mohanty	The application of Wilcoxon norm in linear system identification is analysed.
Ref 2022	[31],	Nonlinear Dynamic System Identification Using Chebyshev Functional Link Artificial Neural Networks. Jagdish C. Patra, Alex C. Kot.	A single-layer functional link ANN (FLANN) in which the need of hidden layer is eliminated by expanding the input pattern by Chebyshev polynomials.
Ref 2022	[32],	Kernel learning for robust dynamic mode decomposition: linear and nonlinear disambiguation optimization Peter J. Baddoo, Benjamin Herrmann, Beverley J. McKeon and Steven L. Brunton,	The dynamic mode decomposition (DMD) has emerged as a cornerstone for modelling high dimensional systems from data. The quality of the linear DMD model is known to be fragile with respect to strong nonlinearity, which contaminates the model estimate. Kernel method efficiently handles high-dimensional data and is flexible enough to incorporate partial knowledge of system physics.
Ref 2022	[33],	Fuzzy Elman Wavelet Network: Applications to function approximation, system identification, and power system control. Zahra Sheikhlar a , Mahdi Hedayati a, Abdolreza Dehghani Tafti a , Hassan Feshki Farahani	The model was the combination of Elman Neural Networks (ENNs), wavelet functions, and fuzzy membership functions (MFs). The integration suggests the use of interval type-2 fuzzy MFs and wavelet functions with self-recurrent and ENN's cross-coupled feedback loops to handle system uncertainties while accurately representing the intrinsic cross-coupled interferences of real dynamic nonlinear systems.
Ref 2012	[40],	Optimization Algorithm-Based Artificial Neural Network Control of Nonlinear Systems. Vishal Srivastava and Smriti Srivastava	A teacher learning-based algorithm (TLBO) has been used with ANN which optimizes the controller parameters and adapts the nonlinearities present in the plants.

Shortcomings

Many works have been performed for this system identification and control as shown in table- 1. However, the simplicity is poor in their approach. Also, the various kernel application is not found for identification and control task. Even they have proposed either for identification and estimation or control with the single approach. To develop a robust algorithm to solve both the tasks of estimation and control, authors have adopted the kernel based popular as well as simple adaptive algorithm and explained in this work as the novelty. In this case no additional optimization algorithm is used.

Significant of research:

For academia and industry system identification and control is an important area of research. To design a model for particular purpose the parameters are to be estimated that leads the system identification problem. Simultaneously the stability of the system needs to be maintained, while the model is useful for any control application. The major objective is to develop an algorithm that can control the system and identify. The system may be of linear or nonlinear or nonlinear dynamic. For nonlinear dynamic system a single algorithm is not yet developed, which is the approach of this work. The kernel application to the adaptive algorithm

works excellent for stabilization and parameter identification of the nonlinear system.

Major Contribution of this work are:

1. The parameters of two nonlinear plants are estimated by applying two Mercer's kernels. After parameterization, the model can be used in a control application.
2. Initially, a Gaussian kernel and a polynomial kernel are used to identify the nonlinear models.
3. Furthermore, a reduced method is applied to the Gaussian kernel, to make the model robust.

These remaining sections are organised as follows: In Section 2, following a brief review, the problem formulation section is discussed. In section 3, the methodology of the work is presented. In section 4, the results and discussion parts are discussed. Finally, the work is concluded in section 5.

2. Problem Formulation

A set of mathematical models can represent any nonlinear model. A general linear time-invariant system (LTI) is an example of nonlinear model representation and is formulated as

$$O(g)\, G(q, \emptyset) M(g) + v(k) \tag{1}$$

Where the transfer function is represented by *G*, input to the LTI system is *M(g)* and output to the system is *O(g)* with an additive disturbance or noise, represented by $v(k)$. The main aim of identification is to match the model to the data. The data set is a collection of nonlinear systems, and is represented as:

$$P_t = \{M(1), O(1), \ldots\ldots, U(N), O(N)\} \tag{2}$$

Where $P(t)$, represents the training set, with *N* being its size. The identification approach is to identify how well the model is able to predict thereby minimizing the error

$$\epsilon(g, \theta) = o(k) - \hat{o}(g, \theta) \tag{3}$$

The size of error represents in scalar norm as

$$l\big(\epsilon(g, \theta)\big) \tag{4}$$

The overall performance is given as

$$O_N(\theta) \sum_{t=1}^{N} l\big(\epsilon(g, \theta)\big) \tag{5}$$

A natural parameter estimation minimizes the prediction fit. It is defined by

$$\widehat{\emptyset}_N = \arg min_{\theta \varepsilon D}\ O_N(\theta) \tag{6}$$

The process is called as prediction error method (PEM) and it is widely used for nonlinear system error predictions.

2.1. Input (state) output representation of systems

Nonlinear systems are represented in the form of differential equations:

$$\frac{dw(g)}{dt} \triangleq w(g) = \varphi\big(w(g), M(g)\big)\ \ t \in R^{+} O(g) = \mu[w(g)] \tag{7}$$

Where, $w(g) \triangleq [w_1(g), w_2(g), \ldots\ldots w_N(g)]^T$, $M(g) \triangleq [M_1(g), M_2(g), \ldots\ldots M_P(g)]^T$ and $O(g) \triangleq [O_1(g), O_2(g), \ldots\ldots O_m(g)]^T$ represents a p input m output order of the given system is n with $M_i(g)$ as inputs , $w_i(g)$ as state variables and $O_i(g)$as the output. Representation of static maps (nonlinear) is given as $\varphi: \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ and $\mu: \mathbb{R}^n \to \mathbb{R}^m$. The vector $w(g)$ at time *t* is state of system and state at time $t_0 < t$ and the input *M* over time interval is $[t_0, t]$. The output *O(g)* is defined as system's state at time *t*. The system's input-state-output is expressed by Eq. (7). In comparison with Eq. (7), representation of a discrete-time system as a differential equation is

$$M(g+1) = \varphi[w(g), M(g)]$$
$$O(g) = \mu[w(g)] \tag{8}$$

Where *M (.), w (.),* and *O(.)* are discrete time sequences. Assuming that the system in Eq. (8) is linear and time-invariant, the system's behavior can be expressed as

$$w(g+1) = Aw(g) + BU(g)$$
$$O(g) = Cw(g) \tag{9}$$

Where *A, B and C are* $(n \times n), (n \times P), (m \times n)$ matrices respectively. According to the theory of LTI systems, where the terms Controllability, stability, and observability of such a system are defined as *A, B, and C.*

2.2. Identification and control

The problem of unknown system identification is when the functions A, *B*, and *C* in Eq. (9) are unknown from Eq. (8). From the Fig.1. (a), is given as the block diagram representation of system identification. Discrete time plant is presented as *M(g)* and *O(g)* respectively, where the input *M(g)* is a function of time that is

uniformly constrained. It is assumed that the nonlinear system is stable, with a known parameterization. If a system's model can be expressed by an operator *P*, then the model's input and output spaces will be *U* and *O*, respectively. The objective is to categorize the class *p(u(k))* to which *P* belongs. *U(k)* is provided to the nonlinear model *P*, the obtained output from the model is *P(u(k))*.

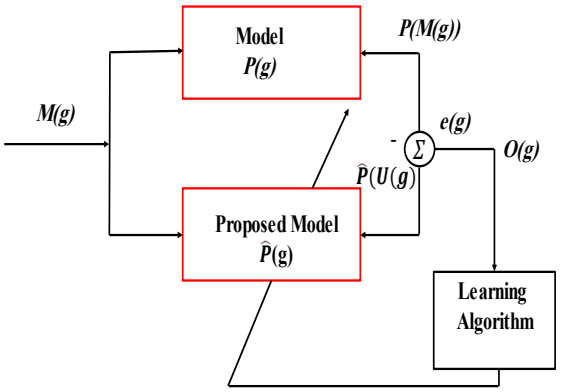

Fig. 1. (a) Block diagram representation of system Identification process.

Now the same input is provided to the proposed model $\hat{P}(K)$,the obtained output is $\hat{P}\big(U(K)\big)$.The identification problem is to determine

$$\|O - \hat{O}\| = \left\|P\big(M(g)\big) - P\big(\widehat{M(g)}\big)\right\| < \epsilon, \quad (10)$$

Where ϵ is a desirable small number greater than 0 and is the norm on the output space based on the Eq. (10), $\hat{P}M(g) = \hat{O}$ and $PM(g) = O$ represent the output of the define model and the plant respectively. The difference between the output of the plant and the estimated model created by $\hat{P}M(g)$ is the error, $e(g) = O - \hat{O}$.

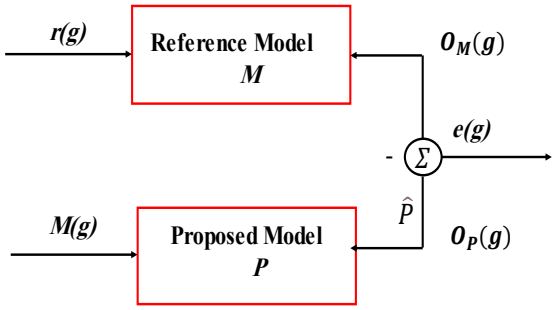

Fig. 1. (b) Block diagram representation of model reference adaptive control.

Control theory is the study and resolution of complex, nonlinear systems in which one or more variables are maintained within predetermined parameters. From Eq. (8), if the functions φ *and* μ are known, the control challenge is to develop a controller that provides the appropriate control input *M(g)* based on all the available information at that moment, *g*. There are a lot of reconciling the controllers of linear systems are reported. Even if the functions *A*, *B* and *C* are known and the functions φ *and* μ are specified from Eq. (9), in case of nonlinear system. That motivates the paper to establish a controller for unknown nonlinear system. Certain model reference adaptive control (MRAC) system. From Fig.1. (b), Block diagram depiction of model reference adaptive control may be observed, that presents a nonlinear plant *P* with the input-output $\{M(g), O_P(g)\}$. The input-output pair of a stable reference model *Z* defines it. $\{r(g), O_Z(g)\}$ are bounded functions indicated by $r: N \to \mathbb{R}$. The plant's output is indicated by output $O_Z(k)$. The control input *M(g)* is defined for every k, $k \geq k_0$ so that,

$$\lim_{k\to\infty} |O_Z(g) - O_P(g)| \leq \varepsilon \quad (11)$$
$$\varepsilon \geq 0.$$

The main goal of the control is to figure out the structure of the controller and change in its parameters to make the difference between actual and estimated outputs, as small as possible. In

this paper, nonlinear difference equations are used to represent two discrete-time models.

Model-1:

$$O_g(g+1) = \sum_{i=0}^{n-1} \alpha_i\, O_g(g-i) + [M(g), M(g-1), \dots, M(g-m+1)] \qquad (12)$$

Model-2:

$$O_g(g+1) = f\left[O_g(g), O_g(g-1), \dots, O_g(g-n+1)\right] + \sum_{i=0}^{m-1} \beta_i\, M(g-i) \qquad (13)$$

were, $[M(g), Og\,(g)]$ denotes the input out pair of the SISO plant at time g and m ≤ n.

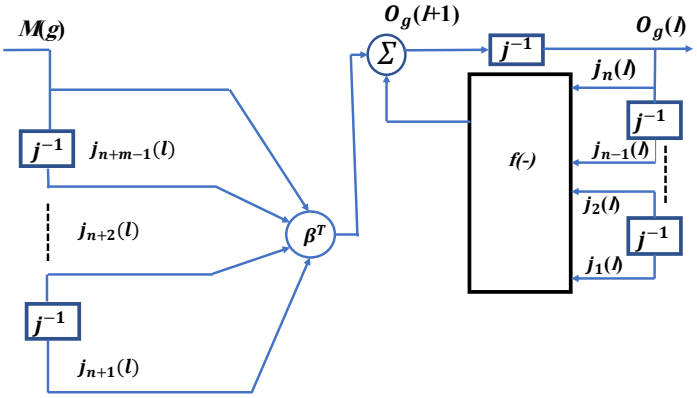

Fig. 2. (a) Block diagram representation of SISO plant model-1.

From Fig. 2. (a) and (b), the block diagram representations of SISO plant model-1 and Model -2 are depicted respectively. Here, the plant's output at time *g+1*, which on past values $O_g(g-i)(i=0,1,\dots.,n-1)$ as well, past m values of the input *M(g-j) (j=0, 1… m-1). In* model-1, the dependency on past values $O_g(g-i)$ acts as linear in the model. Similarly, the dependency on the past values of the input *M(g-j)* acts as linear in the model. The model -2 is perfectly suited for control.

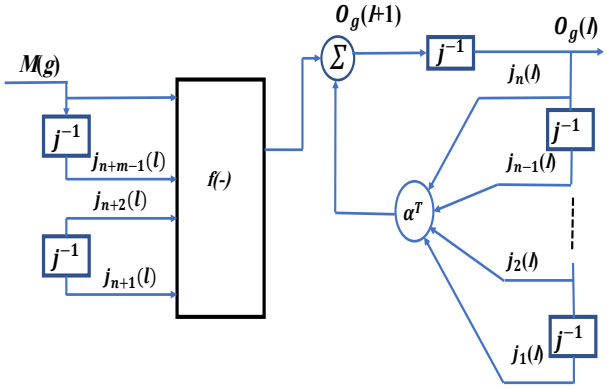

Fig. 2. (b) Block diagram representation of SISO plant model-2.

where $k: M \times M \longrightarrow R$. Example of a vastly used Mercer's kernel is Gaussian kernel.

3. Methodology

The primary objective of identification is to decrease the difference between the real and predicted models by updating the weights of the proposed network. Instead of simply expressing the identification problem in terms of weight (w), the authors are showing their interest in using kernel-based approaches. In this section, two Mercer's kernel's functions are presented with formulation. Using a kernel function, the input may be mapped into a space of high dimensions. The adaptive filtering technique KLMS is an online non-linear filtering approach, that may be considered of as a feature space application of (LMS) algorithm. The input data $M(g) \in M$ is considered as input at time instant k, and $d_n \in R$ is represented as the desired response. The principle of the algorithm starts with mapping the input data $M(g)$ into a high dimensional feature

space by using a Mercer's theorem [17,37]. Further, least mean square is used to formulated_n . A continuous, symmetric, and positive-definite function is what people call Mercer's kernel,

3.1. Polynomial kernel

Scholkopf and Smola in (2001) argue that the input-output map f (.) is defined by a kernel function, is a reproducing kernel Hilbert space(RKHS) $k = M_k, M_j$. As the goal is stated as $\hat{f}$, the estimate of f can be obtained *by* calculating equation

$$\arg\min_{f\in P} \sum_{t=1}^{T} \left(e_g - M_g(M_k)\right)^2 + \alpha^2 \|f\|_{P,}^2 \quad (14)$$

Where e_g is the error of the network and hyperparameter is represented as α . The regularization term is given by the squared RKHS norm *f*. The balance between these two contributions can be changed by changing the hyperparameter, α .$\hat{f}$ can be represented as

$$\hat{f}(M_k) = \sum_{t=1}^{T} \beta_t \; k(M_k, M_j), \quad (15)$$

where $\beta = [\beta_1, \beta_1 \ldots . \beta T]^2$ is equal to $(\mathrm{k} + \alpha^2 I_T)^{-1} O$, $O = [O_1, O_2 \ldots \ldots O_T]^2$ denotes the output vector (output measurements), and Kernel matrix is known as K, as *(k, j)* is (M_k, M_j) .

A kernel function permits an extension in terms of basic functions Υ_q given as

$$k(M_k, M_j) = \sum_{q} \tau_q, \Upsilon_q\,(M_k), \Upsilon_q(M_j) \quad (16)$$

where τ_q is represented as positive scalar and a suitable kernel C_q representation is

$$f(M_k) = \sum_{q} C_{q\Upsilon_q}\,(M_k), \quad (17)$$

Additionally, $if\ \Upsilon_q$ is linearly independent, then it is given as

$$\|f\|_P^2 = \sum_{q} \frac{{C_q}^2}{\tau_q} \quad (18)$$

From the following Eq. (16), it is evident that τ_q coefficients are connected to each Υ_q in calculating the current regularization term presented in Eq. (24). Small values of τ_q , in particular, result in a serious disadvantage of Υ_q.l Hence, the polynomial kernel is presented as

$$k^g(M_k, M_j) = \left(1 + M_k^T M_j\right)^g \quad (19)$$

Where g is an adjustable hyperparameter. The motivation of taking Polynomial kernel is taken from [36].

3.2. Procedure of gaussian kernel least mean square (k-LMS) algorithm

The Gaussian function as kernel parameter is represented as,

$$g(u, u') = e^{\left(-\frac{\|u,u'\|^2}{2\sigma^2}\right)} \quad (20)$$

where $\sigma > 0$ is represented as kernel width. *If g(u,u')* shows that an input space *M* may be transformed into a F feature as mapping φ

$$g(u, u')\varphi(M)^T\varphi(M') \quad (21)$$

Therefore, φ the kernel-induced mapping translates the input *M(g)* in to *F* as $\varphi M(g)$, with the filter weights Υ being updatable through gradient decent over cost function. This cost function represents the mean squared difference between the desired *d* and the expected response $\hat{O}(g)$. As a result, a new pair of samples is treated as { $\varphi M(g), d_n$}, Υ_n.The LMS technique may be used to estimate the weight vector in. *F.*

$$\Upsilon_0 = 0.$$
$$e_n = d_n - \Upsilon_{n-1}^T \varphi M_n$$
$$\Upsilon_n = \Upsilon_{n-1} + \eta e_n \varphi M_n$$

where, η is presented as learning rate for gradient update. Then,

$$\Upsilon_n = \sum_{i=1}^{n} \eta e_i \varphi(M_i) \quad (22)$$

And

$$\hat{O}(g) + 1 = \Upsilon_n^T \varphi(M_{n+1}) = \sum_{i=1}^{n} \eta e_i k(M_i, M_{n+1}) \quad (23)$$

KLMS has been established as a self -regularized algorithm.

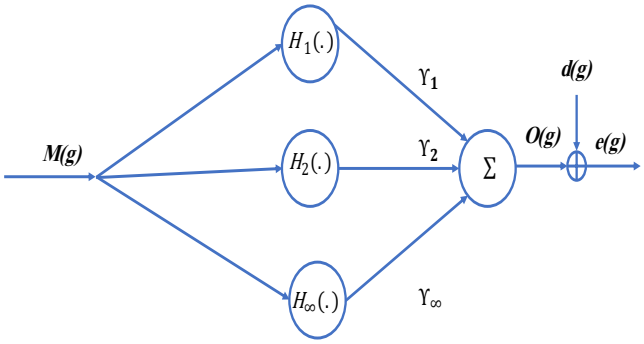

Fig. 3. Block diagram representation of kernel identification process.

The block diagram representation of kernel identification process is depicted in Fig. 3 where *M(g)* is the input vector, $g_1(.),\ldots.g_\infty(.)$ are the kernel vectors, Υ is the proposed kernel, *O(k)* is the output of the model, *e(n)* is the error of the network. In the above dissection, the operation of the kernel identification process is explained. Further, more investigation has been done on gaussian KLMs to make it more robust.

Let us consider input and output pair as $M_g(g)$ and $O_g(g)$,where $g = 1,2,\ldots\ldots\ldots N$ and $M_g(g)$ is considered to be g dimensional of the input space. The goal is to learn a nonlinear input-output mapping as a continuous function.

$$O_g = f(M_g), \quad M_g \epsilon M \subseteq R^N \ , O_g \epsilon R \tag{24}$$

The objective of a kernel adaptive filter is to approximate the mapping *f (.)* based on the input data.

3.3. Use of reduced gaussian kernel

Let $\Upsilon = \{((M_g, O_g)), g = 1,2,\ldots\ldots m\} \in (R^n \times \gamma)^m$ that is taken a data set containing m samples, where $M_g \in R^n, O_g \in \gamma = \{-1,+1\}, g = 1\ldots\ldots m$. Further step is to establish relation Rn to a kernel feature *H*. *H* = (Hilbert space).

$R\hat{}n \dashrightarrow H$

$\emptyset: M_g = (\emptyset(M_g))$

The mapping $\emptyset$ chosen aims to turn nonlinear relations into linear ones. It can be represented as

$$\Upsilon = \{(M_g)), g = 1,2,\ldots\ldots m\} \in (\mathrm{H} \times \gamma)^m \tag{25}$$

Due to the fact that the dimensionality of *H* is far bigger than that of *Rn* and is even finite, direct analysis in *H* using fundamental techniques will be extremely difficult and expensive. Instead of using fundamental techniques, a kernel-trick method is used which calculates the inner product, *H* is computed in R^n with the kernel function as presented in [37, 38. 39]. The formulation is given as

$$k(.): < (\emptyset(M_g)), (\emptyset(M_j)) \geq k^g(M_k, M_j) \tag{26}$$

where $M_k, M_j$are the two vectors in R^n. Without mapping $\emptyset$ the inner can compute sometime directly consequence of the input characteristics. In the case were $M_k, M_j \in R^n$, for $\sigma > 0$, Gaussian kernel method is explained in [37]and presented as

$$k(M_k, M_j) = e^{-\frac{\|M_k, M_j\|^2}{2\sigma^2}} \tag{27}$$

Further to increase the performance and decrease the computational complexity of the kernel functions, more investigation has been done. According to [35], the computational complexity is reduced for Gaussian kernel function and it is used as

$$O_g(g) = \delta(g-1)^T \emptyset(M_g) = \sum_{P=0}^{P} \delta_P^T(g - 1)\emptyset_P(M_g) \tag{28}$$

4. Result and Discussion

The simulation outcomes of nonlinear plant identification using the previously described SISO models. Two benchmark nonlinear system examples are presented in this section to verify the supremacy of the proposed model. Each case has been chosen to highlight a certain issue. To identify the given plant, all samples are used a series-parallel model, and Kernel least mean

square was used to update the weights. In this case, the input is found to appear linearly in the differential equation characterizing the plant. The SISO plant model was adapted from [40]. All the simulation results have been achieved using MATLAB 2021 (a) platform. The parameters evolution the performance is done through MSE and RMSE. The parameters are calculated as :

$$MSE = \frac{1}{N}\sum_{i=1}^{N}|T_i - P_i|^2 \quad (29)$$

MSE calculation in dB

$$10 * log10\big(mean(MSE)\big) \quad (30)$$

$$MSE = \sqrt{\sum_{i=1}^{N}(T_i - p_i)^2} \quad (31)$$

Example-1

The nonlinear plant's output is given by the difference equation [40].

$$O_g(g+1) = 0.3O_g(g) + 0.6O_g(g-1) + f[M(g)] \quad (32)$$

The unknown function is

$$f(M) = 0.6\sin(\pi M) + 0.3\sin(3\pi M) + 0.1\sin(5\pi M) \quad (33)$$

A series parallel model governed in form of difference equation was used

$$O_g(g+1) = 0.3O_g(g) + 0.6O_g(g-1) + Q[M(g)] \quad (34)$$

Q[M(g)] represents the proposed model. The identification operation was continued using an input consisting of a combination of two sinusoidal signals denoted as

$$M(g) = \sin\left(\frac{2\pi k}{250}\right) + \sin\left(\frac{2\pi k}{25}\right), \text{for k} = 500 \quad (35)$$

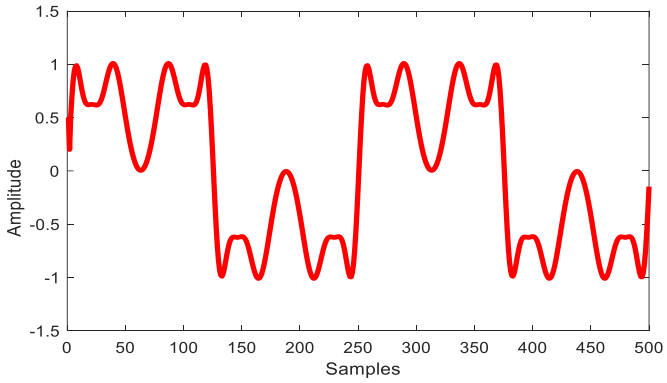

Fig. 4. (a) Output of the actual nonlinear system.

The Fig. 4. (a) represents the actual output of nonlinear model given in example-1. The number of samples taken is 500. The output of the model is $O_g(g+1)$ and input *M(g)*. Initially, the process was done using Gaussian and polynomial kernel functions. From the Fig. 4. (b), the actual vs estimated plot by polynomial kernel is detected. The tracking output can be seen between the actual and reference models. A feedback controller is used to keep the nonlinear 'plant in check' during this process; however, the controller is not satisfactory. Further, the weights of the controller are updated using suitable learning algorithms like the LMS method to eliminate the harmonics and to provide good performance. The weight initially was taken randomly. The program executes until the parameter of the proposed model is estimated, once the parameter of the model is known, then the model can be controlled.

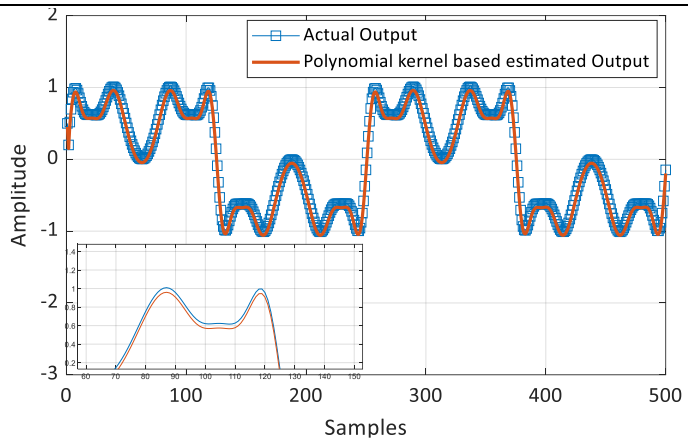

Fig. 4. (b) Actual vs Predicted plot showing output of the plant and Identification model for example-1. The solid line shows the output plot of polynomial Kernel function and the square line shows the actual output of the nonlinear system.

The performance of the Gaussian and polynomial kernel function is analyzed during the process. The complexity of the model increases during the process. Further from the analysis, in order to make the model more robust and with less computational complexity, a reduced method is applied to the Gaussian Kernel functions.

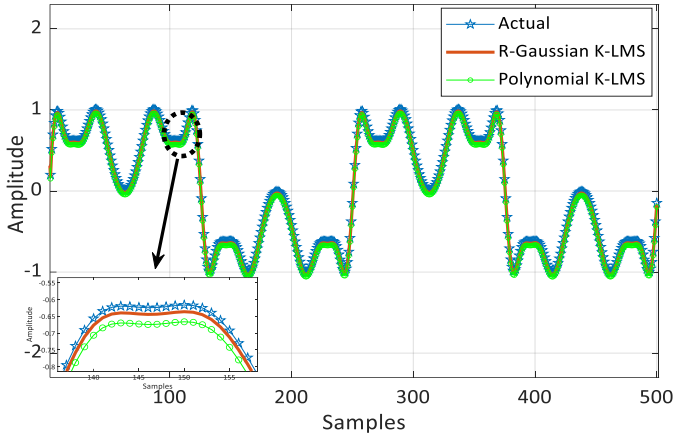

Fig. 4. (c) Actual vs Predicted plot comparing the output of the plant and Identification models example-1. The star line shows the performance of the actual system. The solid line shows the output plot of Reduced Gaussian Kernel function and the circle line shows the performance of polynomial Kernel function.

In Fig. 4. (c), the actual model output versus the predicted model's output can be seen. From the figure, the Reduced Gaussian kernel's output tracks closely as compared to polynomial kernel.

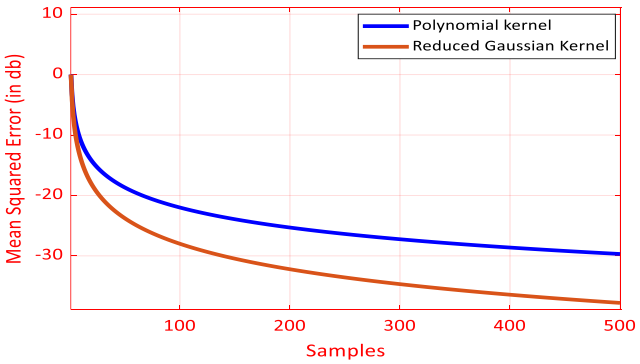

Fig. 4. (d) The learning curve shows the representation of MSE vs Number of samples example-1.

The learning curve between the MSE and the number of samples of the proposed model is presented in Fig. 4. (d). From the curve, it is seen that the performance of the Reduce Gaussian Kernel is better than Polynomial kernel.

Example-2

The second order difference equation describes the plant to be identified [40]

$$O_g(g+1) = f\left[O_g(g), O_g(g-21)\right] + M(g) \quad (36)$$

Were

$$f[O_K(K), O_K(K-1)] = \frac{O_g(g)O_g(g-1)\left[O_g(g)+2.5\right]}{1+{O_g}^2+{O_g}^2(g-1)} \quad (37)$$

This is equivalent to model-2. Using input-output data and an equation, a series parallel identifier is utilized to determine the plant's identity.

$$O_g(g+1) = Q\left[O_g(g), O_g(g-1)\right] + M(g) \quad (38)$$

Where, $Q\left[O_g(g), O_g(g-1)\right]$ is the proposed plant.

In example 2, the presented plant has 200 numbers of samples. Output to the plant is $O_g(g+1)$. In Fig. 5 (a), the output of the plant is depicted. The Fig. 5. (b) shows the Actual vs Predicted plot (polynomial kernel) output and Identification model.

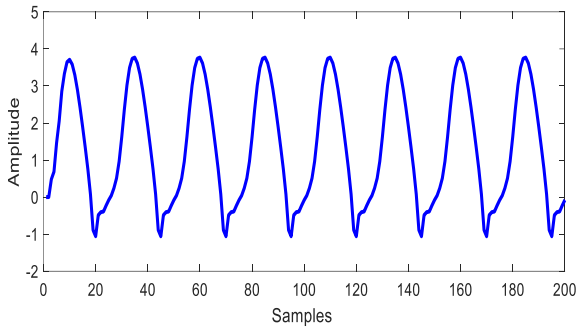

Fig. 5. (a) Output of the actual nonlinear system.

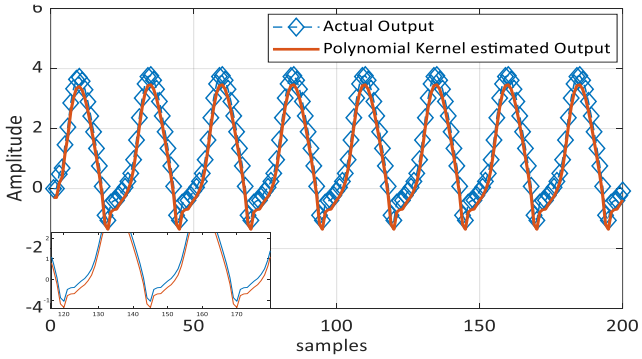

Fig. 5. (b) Actual vs Predicted plot showing output of the plant and Identification model for Example 2. The solid line shows the output plot of polynomial Kernel function and the diamond line shows the actual output of the nonlinear system.

The Fig. 5. (c) shows the output plot of the actual plant and proposed model. From the plot, the Reduced Gaussian Kernel function tracks closely to the output as compare to the polynomial kernel function.

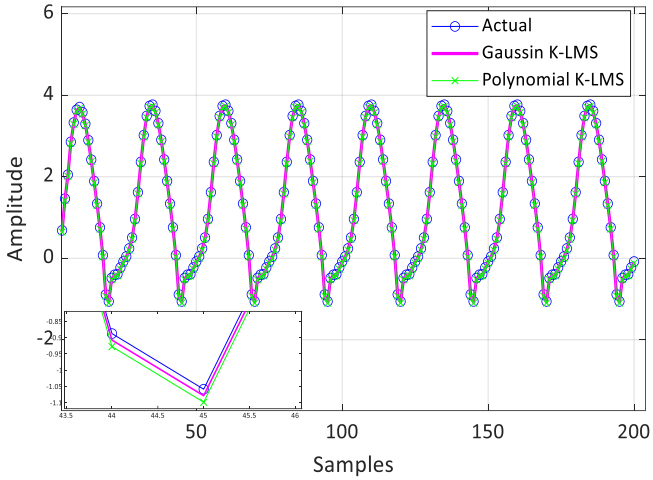

Fig. 5. (c) Actual vs Predicted plot comparing the output of the plant and Identification models example-2. The circle line shows the performance of the actual system. The solid line shows the output plot of Reduced Gaussian Kernel function and the cross line shows the performance of polynomial Kernel function.

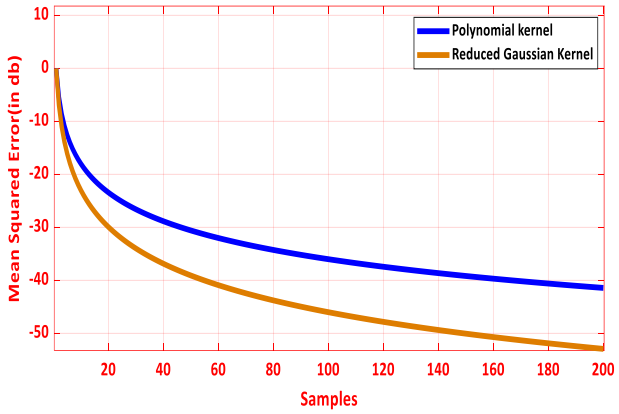

Fig. 5. (d) The learning curve shows the representation of MSE vs Number of Samples example-2.

The Fig. 5. (d) shows the performance curve between MSE and the number of samples of the proposed model.

Tab. 2. Performance of the proposed model.

	Kernel Methods	MSE in (dB)	RMSE	MAE	Timing in (sec)
Exp- 1	Polynomial	-27.769	0.409	0.166	0.030951
	Reduced Gaussian	-38.256	0.251	0.125	0.011030
Exp- 2	Polynomial	-40.9731	0.0564	0.2704	0.02456
	Reduced Gaussian	-54.622	0.0362	0.235	0.01062

From the Table 2, the performance of the models is analysed**.** The polynomial kernel function in example 1 archives -27.769 dB of MSE, 0.409 RMSE and 0.166 MAE. Were the Reduced Gaussian kernel function archives -38.256 of MSE ,0.251 RMSE and 0.125 MAE. Similarly, in example 2, the polynomial kernel function achieves -40.9731 of MSE ,0.0564 RMSE and 0.2704 MAE. Were the Reduced Gaussian kernel function archives -54.622 of MSE ,0.0362 RMSE and 0.235 MAE. Similarly, the timing of the models' cab be compared as, the polynomial kernel for example 1 is 0.030951 sec, for Reduced Gaussian kernel function is 0.011030. In example 2 the polynomial kernel function takes 0.02456 sec to complete the process and Reduced Gaussian kernel function takes 0.01062 sec to execute the process. From the above analysis the performance of the Reduced Gaussian kernel function is better in minimising the error with less computational time as compare to polynomial kernel.

Tab. 3. Performance models sample wise (MSE in dB)

No of examples	Kernel Methods	50 Samples	100 Samples	150 Samples	200 Samples	500 Samples
Example1	Polynomial	-20.3876	-24.00	-25.12	-26.33	-27.769
	Gaussian	-22.8722	-27.622	-30.7545	-33.743	-38.256
Example2	Polynomial	-30.5815	-33.33	-36.699	-40.9731	-
	Gaussian	-39.0763	-46.09	-50.122	-54.622	-

The table 3 presents the performance of the models, minimization of MSE with respective the number of samples. In example 1 polynomial kernel function archives -27.769 dB MSE of 500 samples. The study is continued for every 50 samples to verify the status of the MSE. At 50 samples the MSE is -20.3876 dB, at 100 samples the MSE, at 150 samples MSE is -25.12 dB, at 200 samples the MAE is -26.33. Similar at 500 samples the MSE is -27.769. In example 2, the polynomial kernel function achieves -30.5815 dB at 50 samples, at 100 samples -33.33 dB, at 150 samples -36.699, at 200 samples –40.9731. In example 1 the Reduced Gaussian kernel function archives -22.8722 at 50 samples, at 100 samples -27.622, at 150 samples -30.7545 dB, at 200 samples -40.973 dB.

Similarly, in example 2 at 50 samples -39.0763 dB, at 100 samples -46.09 dB, at 150 samples -50.122 dB at 200 sample -54.622 dB. From the above discussion it is cleared that by applying reduced method to the Gaussian kernel function the overall performance of the model is increased and further it can be used for identification and control of models in industrial application.

5. Conclusion

Identifying complex nonlinear plants is a major issue in today's control systems. From the literature it is found that the simplicity is poor in the existing algorithms. The system needs to estimate the parameters accurately. Simultaneously, control can be performed with stability. Hence it is required to design new methodology. The model is considered as adaptive SISO model. Initially SISO model is chosen because of simplicity and industrial control application. To adjust the weights of the model kernel based LMS algorithm is chosen, so that the job of estimation and control can be performed. After verification of different kernels, Gaussian kernel is set as reduced kernel for less computational complexity. It is explained in section 3. A Reduced method is applied to a Gaussian kernel function in this work. To validate the superiority of the model, two nonlinear plants modelling SISO plants are used. Since the kernel parameter has no effect on complexity, it is best to establish the ideal value for the kernel parameter prior to applying the algorithm, according to the complexity analysis described previously. From the above analysis the model has good prediction and control accuracy. The computational complexity also reduced. Comparison is done between the two kernels methods. The performance of the models is verified in every 50 samples and presented in the table. In example1 the Reduced Gaussian kernel function achieve -38.256 dB, and in example 2, -54.622 dB with time 0.011030sec and 0.01062 sec respectively. From a deep analysis it is found that the performance of Gaussian kernel is increased by implementing Reduced method. In recent time certain works, [41,42,43], are considered the VHDL implementation of earlier methods. The proposed prospective algorithm may be implemented and is kept for future work. The findings of the work are:

1. The adaptive filtering technique (K-LMS) is in this case, a non-linear filtering approach, that may be considered of as a feature space application of (LMS) algorithm. This can perfectly estimate the weight vector (parameters) of the nonlinear model.
2. In a practical environment, the complexity of the proposed method depends mostly on the reduced Gaussian kernel. This makes it possible to obtained good amount of tracking performance.
3. From the result analysis part, the supremacy of the model is verified and this model is useful for industrial control application.

The proposed method is suitable for industrial application, automated system, smart city application, estimation of aerodynamics, stability and control of flight vehicle, biomedical signal identification etc.

In future, complex and dynamic nonlinear system examples may be considered to verify proposed method. Also, the deep machine learning approach may be applied for practical application. Simultaneously, the implementation of these algorithms can be verified for different applications.

References

[1] Ljung, L., "*Perspectives on system identification*". Annual Reviews in Control, Vol. 3, No. 1, (2010), pp.1-12.

[2] Tangirala, A. K. *Principles of system identification: theory and practice*. Crc Press. (2018).

[3] Cluett, W. R., "*Principles of System Identification: Theory and Practice [Bookshelf]*". IEEE Control Systems Magazine, Vol. 37, No-2, (2017), pp.181-184.

[4] Benesty, J., & Huang, Y., "*Adaptive signal processing: applications to real-world problems",* Springer Science & Business Media, (2013).

[5] John, G., & Charles, M. R., "*Algorithms for statistical signal processing,"* Prentice Hall, London, (2002).

[6] Feuer, A., & Weinstein, E., "*Convergence analysis of LMS filters with uncorrelated Gaussian data*," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 33, No. 1, (1985), pp.222-230.

[7] Zou, Y., Chan, S. C., & Ng. T. S., "*Least mean M-estimate algorithms for robust adaptive filtering in impulse noise,"* IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, Vol. 47, No. 12, (2000), pp.1564-1569.

[8] Zhou, Y., Chan, S. C., & Ho, K. L., "*New sequential partial-update least mean M-estimate algorithms for robust adaptive system identification in impulsive noise*," IEEE Transactions on Industrial Electronics, Vol. 58, No. 9, (2010), pp. 4455-4470.

[9] Song, P., & Zhao, H., "*Affine-projection-like M-estimate adaptive filter for robust filtering in impulse noise*," IEEE Transactions on Circuits and Systems II: Express Briefs, Vol. 66, No. 12, (2019), pp. 2087-2091.

[10] Wang G., & Zhao, H., "*Robust adaptive least mean M-estimate algorithm for censored regression*," IEEE Transactions on Systems, Man, and Cybernetics: Systems, (2021).

[11] Lu, L., Chen, L., Zheng, Z., Yu Y., & Yang, X., "*Behaviour of the LMS algorithm with hyperbolic secant cost,*" Journal of the Franklin Institute, Vol. 357, No. 3, (2020), pp-1943-1960.

[12] Chaudhary, N. I., & Z, M. A., "*Raja Identification of Hammerstein nonlinear ARMAX systems using nonlinear adaptive algorithms*," Nonlinear Dynamics, Vol. 79 No. 2, (2015), pp.1385-1397.

[13] Raja, M. A. Z., &, Chaudhary, N. I., "*Adaptive strategies for parameter estimation of Box–Jenkins systems*," IET Signal Processing, Vol. 8, No. 9, (2014), pp. 968-980.

[14] Kohli,K., & Rai, A., "*Numeric variable forgetting factor RLS algorithm for second-order Volterra filtering,*" Circuits, Systems, and Signal Processing, Vol. 32, No. 1, (2013), pp. 223-232.

[15] Ding K., & Liu, X., "*Decomposition based Newton iterative identification method for a Hammerstein nonlinear FIR system with ARMA noise*," Circuits, Systems, and Signal Processing, Vol. 33, No. 9, (2014), pp. 2881-2893.

[16] Li, M., Liu, X., & Ding, F., "*Filtering-based maximum likelihood gradient iterative estimation algorithm for bilinear systems with autoregressive moving average noise,*" Circuits, Systems, and Signal Processing, Vol. 37, No. 11, (2018), pp. 5023-5048.

[17] Janjanam, L., Saha, S. K., Kar R., & Mandal, D., "*Volterra filter modelling of non-linear system using Artificial Electric Field algorithm assisted Kalman filter and its experimental evaluation*," ISA transactions, (2020).

[18] Kubat, M. *"Neural networks: a comprehensive foundation by Simon Haykin, Macmillan"*, ISBN 0-02-352781-7. The Knowledge Engineering Review, Vol. 13, No. 4, (1999), pp. 409-412.

[19] Aadaleesan, P., Miglan, N., Sharma, R., & Saha, P., "*Nonlinear system identification using Wiener type Laguerre–Wavelet network model*," Chemical Engineering Science, Vol. 63, No. 15, (2008), pp. 3932-394.

[20] Janjanam, L., Saha, S. K., Kar, R., &, Mandal, D., "*Volterra filter modelling of non-linear system using Artificial Electric Field algorithm assisted Kalman filter and its experimental evaluation*." ISA transactions, Vol. 125, (2020), pp. 614-630.

[21] Aliskan, I., "Adaptive model predictive control for Wiener nonlinear systems. *Iranian Journal of Science and Technology,"* Transactions of Electrical Engineering, Vol. 43, No. 1, (2019), pp. 361-377.

[22] Dash, S., & Mohanty, M.N., "Analysis *of outliers in system identification using WLMS algorithm,"* International Conference on Computing, Electronics and Electrical Technologies (ICCEET),(2012), a.

[23] Dash, S., & Mohanty, M. N., "*Variable sign-sign Wilcoxon algorithm: a novel approach for system identification," International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 2, No. 4, (2012)(b) , pp. 481-486.

[24] Dash, S., & Mohanty, M. N., "*Identification of System with Non-stationary Signal Using Modified Wilcoxon Approach*," Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) (2013).

[25] Jain, S., & Majhi, S., "*Zero-Attracting Kernel Maximum Versoria Criterion Algorithm for Nonlinear Sparse System Identification*," IEEE Signal Processing Letters, (2022).

[26] Xiang, Liu., Song, C., and Pang, Z., "*Kernel recursive maximum correntropy with variable centre*," Signal Processing, Vol. 191, (2022), pp. 108364.

[27] Zhang, Y., Zhang, D., Han, Z., &Jiang, P., "*A Joint Denoising Learning Model for Weight Update Space–Time Diversity Method*," Remote Sensing, Vol. 14, No. 10, (2022), pp. 2430.

[28] Pattanaik, R. K., & Mohanty, M. N., "*Nonlinear System Identification Using Robust Fusion Kernel-Based Radial basis function Neural Network*," International Conference on Emerging Smart Computing and Informatics (ESCI*)* IEEE, (2022), pp. 1-5.

[29] Porikli, F., "*Constant time O (1) bilateral filtering*," Conference on Computer Vision and Pattern Recognition, *IEEE*, (2008), pp. 1-8.

[30] Ring, M., & Escoffier, B. M., "*An approximation of the Gaussian RBF kernel for efficient classification with SVMs*," Pattern Recognition Letters, Vol. No. 84, (2016), pp.107-113.

[31] ALTINER, İ. B., Doğan, M., & Daşdemir, J. "*Adaptive output tracking of distributed parameter systems"*. Turkish Journal of Electrical Engineering and Computer Sciences, Vol. 30, No. 3, (2022), pp. 518-530.

[32] Patra, J. C., & Kot, A. C. "*Nonlinear dynamic system identification using Chebyshev functional link artificial neural networks"*. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics*)*, Vol. 32, No. 4, (2002), pp. 505-511.

[33] Baddoo, P. J., Herrmann, B., McKeon, B. J., & Brunton, S. L. "*Kernel learning for robust dynamic mode decomposition: linear and nonlinear disambiguation optimization."* Proceedings of the Royal

Society A, 2022, Vol. 478, No. 2260, pp. 20210830.

[34] Sheikhlar, Z., Hedayati, M., Tafti, A. D., & Farahani, H. F. "*Fuzzy Elman Wavelet Network: Applications to function approximation, system identification, and power system control. Information Sciences"*, Vol. 583, (2022), pp. 306-331.

[35] Hung, G. C., KAMARUDIN, M., Lee, I. E., & Tan, S. F. Identification, and mitigation of non-line-of-sight path effect using repeater for hybrid ultra-wideband positioning and networking system. *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol. 30, No. 3, (2022), pp. 1084-1096.

[36] Liu, Y., Sun, C., & Jiang, S., "*A reduced Gaussian kernel least-mean-square algorithm for nonlinear adaptive signal processing*," Circuits, Systems, and Signal Processing, Vol. 38, No. 1, (2019), pp. 371-394.

[37] Libera, A. D., Carli, R., & Pillonetto, G., "*A novel multiplicative polynomial kernel for volterra series identification*," IFAC-Papers on Line, Vol. 53, No. 2, (2020), pp. 316-321.

[38] Mohapatra, S. K., Behera, S., & Mohanty, M. N. "*A comparative analysis of cardiac data classification using support vector machine with various kernels"*. International Conference on Communication and Signal Processing *(ICCSP)* IEEE. In (2020), pp. 515-519.

[39] Taylor, J. S., & Cristianini, N., "*Kernel methods for pattern analysis,"* Cambridge university press, (2004).

[40] Khanna, G. D., Bhattacharyya A., Hassanien, S., Ella,Aboul ., Sameer, A., Jaiswal, A., Optimization Algorithm-Based "*Artificial Neural Network Control of Nonlinear Systems*," Advances in Intelligent Systems and Computing International Conference on Innovative Computing and Communications Vol. 1165 , (2021), pp.1007-1015.
http://doi:10.1007/978-981-15-5113-0_85.

[41] Kumpati, S. N., & Kannan, P., "Identification and control of dynamical systems using neural networks." IEEE Transactions on neural networks, Vol. 1, No.1, (1990), pp. 4-27.

[42] Shen, H., Yu, F., Kong, X., Mokbel, A. A. M., Wang, C., & Cai, S. "*Dynamics study on the effect of memristive autapse distribution on Hopfield neural network. Chaos":* An Interdisciplinary Journal of Nonlinear Science, Vol. 32, No. 8, (2022), pp. 083133.

[43] Merabti, H., & Massicotte, D. *"FPGA based implementation of a genetic algorithm for ARMA model parameters identification."* In *Proceedings of the 24th edition of the great lake's symposium on VLSI* (2014), pp. 95-96.

[44] Akpan, V. A., Chasapis, D., & Hassapis, G. D. "*FPGA Implementation of Neural Network-Based AGPC for Nonlinear F-16 Aircraft Auto-pilot Control: Part 1– Modeling, Synthesis, Verification and FPGA-in-Loop Co-Sim"*. American Journal of Embedded Systems and Applications, Vol. 9, No. 1, (2022), pp. 6-36.

Follow This Article at The Following Site:

Pattanaik. R. K., Sarfraz. M., Mohanty. M. N., Nonlinear System Parameterization and Control using Reduced Adaptive Kernel Algorithm. IJIEPR. 2022; 33 (4) :1-16
URL: http://ijiepr.iust.ac.ir/article-1-1572-en.html