Minimax Theory and its Applications
Volume 11 (2026), No. 4s, 564 to 579

Kernel approximation-based mathematical modeling for nonlinear parameter estimation in intelligent monitoring systems

J.S. Jaslin[1], S. Vadivel[2], Dr. Namrata Gain[3], Dr.M. Rajesh Khanna[4], Konakanchi Sravani[5], P. Geetha[6], Aakansha Soy[7]

[1]Assistant Professor, Department of Computer Science and Engineering, J.J. College of Engineering and Technology, Tiruchirappalli, Tamil Nadu, India. jaslinjs@jjcet.ac.in, https://orcid.org/0009-0007-7033-5569

[2]Assistant Professor, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Namakkal, Tamil Nadu, India. svadivel2399@gmail.com, https://orcid.org/0009-0009-4345-396X

[3]Associate Professor, Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Rungta International Skills University, Bhilai, Chhattisgarh, India. dr.namrata.gain@rungta.org, https://orcid.org/0000-0003-1714-5604

[4]Associate Professor, Department of Information Technology, Vel Tech Multi Tech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, Chennai, Tamil Nadu, India. rajeshkhanna@veltechmultitech.org, https://orcid.org/0000-0002-3397-2039

[5]Assistant Professor, Department of Electrical and Electronics Engineering, VNR VJIET (Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering & Technology), Hyderabad, Telangana, India. sravani_k@vnrvjiet.in, https://orcid.org/0009-0003-3234-1045

[6]Department of FED, Ramachandra College of Engineering, Eluru, India. geethapillutla44@gmail.com

[7]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India. ku.aakanshasoy@kalingauniversity.ac.in, https://orcid.org/0009-0002-1955-6909

Abstract:

Nonlinear parameter estimation with accuracy is crucial in intelligent monitoring systems where the relationship between the variables and the system's state is complicated, and thus cannot be sufficiently modeled by conventional fixed-model approaches. In this study, a kernel approximation-based mathematical model framework is developed for nonlinear parameter estimation that incorporates flexibility in function approximation and regularization in constrained optimization. The suggested formulation involves an approximation of the unknown nonlinear function using a finite expansion of positive definite kernels and estimating the parameters using a regularized error function approximation. A Gaussian kernel is utilized for modeling nonlinearities, whereas the regularization ensures that the complexity of the model and estimation stability are controlled in the presence of noise. Experiments are performed on the controlled nonlinear functions where varying size samples and noisy data are considered. Proposed model performance analysis is done by using the measures like root mean square error, mean absolute error, mean absolute percentage error, coefficient of determination, normalized root means square error, parameter estimation error, and computation time. Comparison with polynomial nonlinear regression, radial basis function regression, Gaussian process regression, support vector regression, kernel ridge regression, and nonlinear least squares is done for the efficacy of the kernel approximation technique. Experiments are also done to analyze the selection of kernel function, kernel regularization, approximation order, and noise effect.

Keywords: Nonlinear parameter estimation, kernel approximation, intelligent monitoring systems, regularized optimization, gaussian kernel.

1. Introduction

The issue of nonlinear parameter estimation is a basic one in areas such as mathematical modeling, system identification, process monitoring, and intelligent engineering systems, where the dependency between the observable variables and unknown parameters of the system is not properly modeled using linear or fixed forms. The nonlinear nature of processes in actual monitoring setups can emerge from changes in operational conditions, interactions in sensors, dynamics of the system, disturbances, and measurement noise. Although multivariate statistical monitoring has highlighted the need for reliable extraction of information from measurements, conventional process monitoring methods might get limited due to nonlinear dependence dominating the process (Wang et al., 2018). It has been proved by some recent advances in intelligent process monitoring that the benefits of utilizing data-based modeling, explainable intelligence, and continual evaluation of systems for industrial purposes are highly valuable (Kousalya et al., 2025; Mzeh & Salabi, 2025). Learning analytics have been found valuable by using data-driven decision-making models to gain insights from complicated data sets through evidence-based decision making (Schneider & Weber, 2024).

Identification of nonlinear systems thus has drawn a great deal of attention from the autoregressive, Volterra, Wiener, adaptive, and neural network models (Yin et al., 2015; Gupta et al., 2021; Mazhar et al., 2019). Such methods have shown that it is possible to enhance the modeling of complex engineering systems using nonlinear models; however, their effectiveness is highly dependent on model architecture, parameter initialization, training environment, and availability of relevant data samples. Parameter estimation in the case of nonlinear systems may present non-convex optimization problems which lead to specialized optimization formulations for models like Hammerstein among others Chaudhary et al., (2020). There are also such issues related to the estimation and identification of aerospace, robotic, transportational, and electromechanical systems (Mu et al., 2016; Bai et al., 2020).

The kernel method offers an appealing mathematical solution since it allows nonlinear functions to be expressed using arbitrary combinations of kernel functions without defining any explicit high-order model formulation. Reproducing kernel Hilbert space approach has been studied for adaptive nonlinear estimation and it is evident that kernel representation is mathematically suitable for nonlinear system identification problems (Bobade et al., 2019). Kernel adaptive filtering and random Fourier feature approximation solutions have further shown that kernel-based nonlinear models can be implemented effectively with the ability to approximate well Gogineni et al., (2022) and Piltan & Kim, (2021). RBF-based and similar approximation approaches have been used for nonlinear engineering modeling and optimization tasks (Dey et al., 2015; Zhou et al., 2025).

Intelligent monitoring algorithms necessitate robust estimates in the presence of disturbances, uncertainties, and varying conditions (Zabihi et al., 2024; Sharma & Desai, 2024; Feng & Li, 2018). Continuous monitoring and intelligent analysis of physiological changes has also been illustrated by AI-enabled cyber-physical healthcare systems (Shrivastava et al., 2026). Kernel and learning-based methods have been used in fault diagnosis, detection of sensor and actuator faults, electrical network monitoring, predictive maintenance, and condition assessment (Seetha Lakshmi & Angel Cerli, 2026; Papadopoulos & Christodoulou, 2024; Piltan & Kim, 2021). These works illustrate the necessity of good nonlinear approximation to support engineering monitoring tasks, but most existing methods have not considered developing a generalized mathematical relationship between the approximation error, parameter estimation, regularization, convergence, and robustness. Modeling of complicated random and computational processes has also emphasized the significance of having a proper mathematical formulation for reliable system representation (Yao & Qiu, 2024).

Nevertheless, the kernel approximation formulation for intelligent monitoring oriented nonlinear parameter estimation is yet to be explored in a well-structured and comprehensive manner. Particularly, the interplay between kernel approximation degree, regularization, parameter constraint, estimation error, and computational complexity deserves more attention from a mathematical point of view. The current paper proposes to fill this research gap by constructing a mathematical model of nonlinear parameter estimation in intelligent monitoring systems on the basis of kernel approximation. The suggested formulation uses finite kernel expansion for nonlinear monitoring behavior and regularized constrained optimization for unknown parameter estimation. The work also develops approximation and error formulation, objective function analysis, convergence and stability, and computational complexity evaluation of the suggested formulation for different sample sizes and noises. Comparative and ablation analyses have been used for assessing the contribution of kernel approximation, kernel choice, regularization, and approximation degree.

The article is divided into nine major sections. In section 1, there is an introduction to the problems of nonlinear parameter estimation and justification for the kernel approximation-based approach. Mathematical preliminaries and problem formulation are discussed in section 2. The kernel approximation-based mathematical model with parameter estimation and optimization approaches are formulated in section 3. Theoretical analysis on approximation, error, convergence, stability, and complexity is discussed in section 4. Section 5 talks about the intelligent monitoring mechanism. Numerical experiments and evaluation metrics are introduced in section 6. Section 7 deals with comparative and ablation analysis. Section 8 has the discussion while section 9 gives conclusions.

2. Mathematical preliminaries and problem formulation

2.1. Nonlinear monitoring system representation

Let consider an intelligent surveillance system, where a group of observable variables is obtained either from sensors or from surveillance entities. The input vector at the i-th observation can be denoted using equation 1

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{id}]^T \in \mathbb{R}^d \tag{1}$$

And let $y_i \in \mathbb{R}$ denote the corresponding monitored response. The nonlinear behavior of the monitoring process can be represented as equation 2

$$y_i = f(\mathbf{x}_i; \boldsymbol{\theta}) + \varepsilon_i, \qquad i = 1,2, \dots, N \tag{2}$$

Where $f(\cdot)$ is an unknown nonlinear system function, $\boldsymbol{\theta} \in \mathbb{R}^p$ is the unknown parameter vector, and ε_i denotes measurement and modeling noise. The complete observation set is defined as equation 3

$$\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N \tag{3}$$

The objective is to estimate the unknown parameters such that the mathematical model accurately reproduces the observed nonlinear behavior while maintaining stability and controlled model complexity.

2.2. Nonlinear parameter estimation problem

The parameter estimation problem can be formulated as the identification of equation 4 and 5

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta} \in \Theta} \mathcal{L}(\boldsymbol{\theta}) \tag{4}$$

Where $\Theta \subseteq \mathbb{R}^p$ is the feasible parameter space and $\mathcal{L}$ is an error or loss functional. For conventional nonlinear regression, the least-squares objective is

$$\mathcal{L}_{LS}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^N [y_i - f(\mathbf{x}_i; \boldsymbol{\theta})]^2 \tag{5}$$

However, directly estimating parameters from an unknown nonlinear function can become difficult when f has a complex or poorly specified structure. Kernel approximation is therefore introduced to obtain a flexible finite-dimensional representation.

2.3. Kernel functions and reproducing-kernel concepts

Let

$$K: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$$

be a symmetric positive-semidefinite kernel satisfying

$$K(\mathbf{x}, \mathbf{z}) = K(\mathbf{z}, \mathbf{x})$$

And

$$\sum_{i=1}^N \sum_{j=1}^N c_i\, c_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0$$

For every coefficient vector $\mathbf{c} \in \mathbb{R}^N$. Such a kernel generates a reproducing kernel Hilbert space (RKHS), denoted by $\mathcal{H}_K$. For every $f \in \mathcal{H}_K$, the reproducing property is expressed as equation 6

$$f(\mathbf{x}) = \langle f, K(\mathbf{x},\cdot)\rangle_{\mathcal{H}_K} \quad (6)$$

This property allows nonlinear functions to be represented through kernel-induced basis functions without explicitly constructing a high-dimensional feature mapping.

Typical kernels suitable for nonlinear monitoring include the Gaussian radial basis function are shown in equation 7 and 8

$$K(\mathbf{x},\mathbf{z}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{z}\|^2}{2\sigma^2}\right) \quad (7)$$

And the polynomial kernel,

$$K(\mathbf{x},\mathbf{z}) = (\mathbf{x}^T\mathbf{z} + c)^q \quad (8)$$

Where σ, c, and q control the corresponding kernel characteristics.

2.4. Kernel approximation theory

A nonlinear monitoring function can be approximated using a finite kernel expansion in equation 9

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^{M} \alpha_j K(\mathbf{x},\mathbf{z}_j) \quad (9)$$

where M is the number of kernel centers, $\mathbf{z}_j$ represents the j-th center, and α_j is its associated coefficient. In matrix form, for the observation set are represented in equation 10 and 11

$$\hat{\mathbf{y}} = \mathbf{K}\boldsymbol{\alpha}, \quad (10)$$

Where

$$\mathbf{K}_{ij} = K(\mathbf{x}_i,\mathbf{z}_j) \quad (11)$$

The approximation error can be expressed as equation 12 and 13

$$E_{app} = \|f - \hat{f}\|_{\mathcal{H}_K} \quad (12)$$

or, when evaluated over observations,

$$E_{data} = \frac{1}{N}\sum_{i=1}^{N}\left[y_i - \hat{f}(\mathbf{x}_i)\right]^2 \quad (13)$$

The greater the number of representative kernel centers, the greater is the approximation flexibility, but increased model complexity may introduce more noise. For this reason, kernel approximation needs to be regularized.

2.5. Unknown parameter vector

Let the unknown parameter vector be in equation 14

$$\boldsymbol{\theta} = \left[\theta_1, \theta_2, \ldots, \theta_p\right]^T \quad (14)$$

They could stand for physical parameters, response functions, kernel parameters, or other system-specific parameters. The estimation framework that is proposed will determine equation 15

$$\hat{\boldsymbol{\theta}} = \left[\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_p\right]^T \quad (15)$$

from the observations available. The accuracy of parameters is measured using the deviation of the true/reference parameter vector from the estimated parameter vector as shown in equation 16

$$E_\theta = \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|_2 \quad (16)$$

In the case where the parameters are unknown, the model-based prediction error and residuals characteristics are applied.

2.6. Objective and error function

To get a stable solution, the study poses the problem of estimation as a regularization problem as given in equations (17-20).

$$\mathcal{J}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}[y_i - f(\mathbf{x}_i; \boldsymbol{\theta})]^2 + \lambda\|\boldsymbol{\alpha}\|_2^2 \quad (17)$$

where $\lambda \geq 0$ is the regularization coefficient. Under the kernel approximation,

$$f(\mathbf{x}_i; \boldsymbol{\theta}) \approx \sum_{j=1}^{M} \alpha_j K(\mathbf{x}_i, \mathbf{z}_j; \boldsymbol{\theta}) \quad (18)$$

giving

$$\mathcal{J} = \frac{1}{N}\|\mathbf{y} - \mathbf{K}(\boldsymbol{\theta})\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_2^2 \quad (19)$$

The resulting parameter-estimation problem becomes

$$(\grave{\boldsymbol{\alpha}}, \grave{\boldsymbol{\theta}}) = \arg\min_{\boldsymbol{\alpha},\boldsymbol{\theta}} \mathcal{J}(\boldsymbol{\alpha}, \boldsymbol{\theta}) \quad (20)$$

This formulation simultaneously minimizes data-fitting error and controls the complexity of the kernel representation.

2.7. Assumptions and constraints

The mathematical formulation will be performed on the basis of the following assumptions. The first assumption is that the observed data reflects the operational state of the sensor system properly. The second assumption is that the noise has finite variance and bounded second-order moments. The third assumption is that the selected kernel should be symmetric and positive semidefinite such that the study has an RKHS. The fourth assumption is that the feasible parameter set Θ is either bounded or otherwise constrained properly in order to avoid physically impossible solutions. The regularization parameter is such that

$$\lambda > 0$$

Ensuring that the optimization problem remains well-conditioned. Parameter constraints can be expressed as

$$\boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U$$

Where $\boldsymbol{\theta}_L$ and $\boldsymbol{\theta}_U$ indicate the lower and upper limits of the parameters. Based on this assumption, the presented kernel approximation approach offers a mathematically controlled methodology for nonlinear parameter estimation and intelligence monitoring.

where $\boldsymbol{\theta}_L$ and $\boldsymbol{\theta}_U$ denote lower and upper parameter bounds. Under these assumptions, the proposed kernel approximation framework provides a mathematically controlled basis for nonlinear parameter estimation and subsequent intelligent monitoring.

3. Proposed kernel approximation-based mathematical model

The suggested model provides a methodology of kernel approximation technique for the estimation of unknown parameters of nonlinear monitoring system. This methodology first maps the monitoring data to the normalized data form, then constructs a kernel approximation for the nonlinear function, and finally estimates the parameters using the regularized constrained optimization. The whole structure of this methodology is constructed in such a way to ensure nonlinear representational capabilities and approximation complexities.

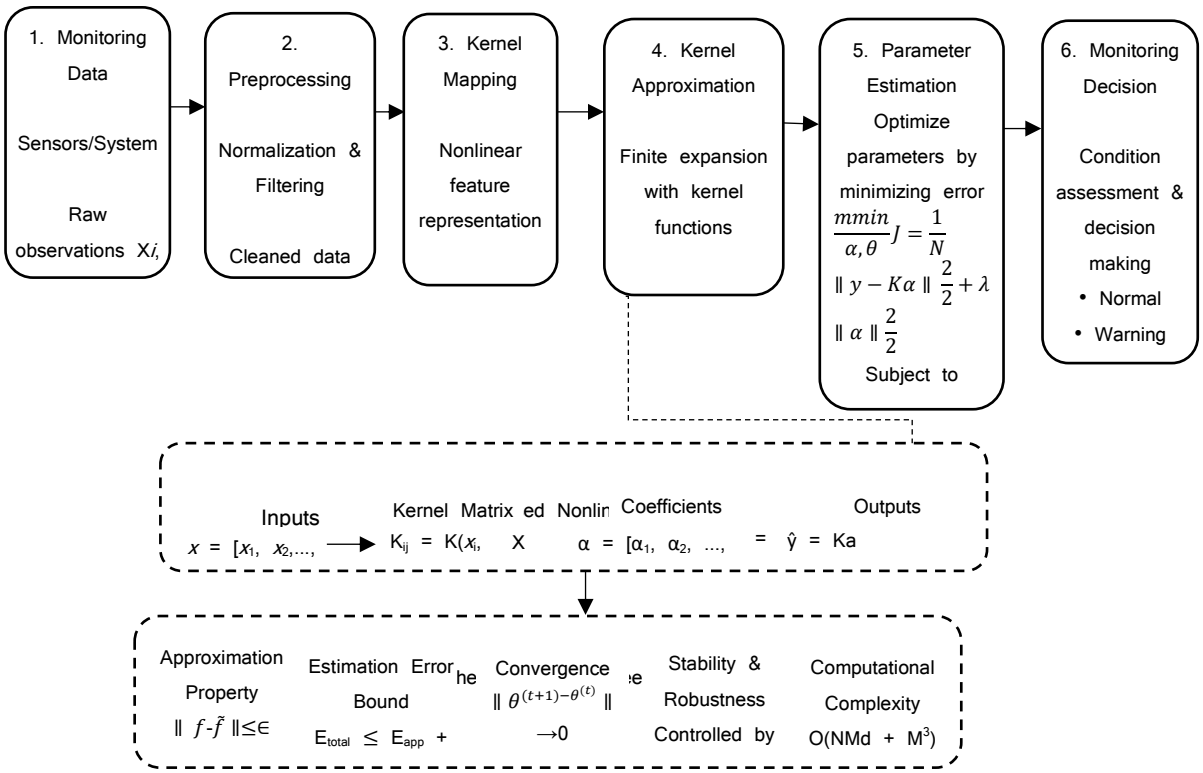


Figure 1: Proposed kernel approximation-based intelligent monitoring framework.

Figure 1 shows the entire process of the proposed framework, starting from monitoring data acquisition and preprocessing, through nonlinear kernel mapping, kernel approximation, parameter estimation, and monitoring decision making. The bottom part indicates the kernel model mathematically, whereas the topmost part is the theoretical part that involves approximation, estimation, convergence, and computational complexity.

3.1. Input: monitoring observations

Let the monitoring system generate N observations represented by equation 21

$$\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N} \qquad (21)$$

where $\mathbf{x}_i \in \mathbb{R}^d$ represents the d-dimensional monitoring input and $y_i \in \mathbb{R}$ denotes the corresponding system response. The underlying nonlinear relationship is expressed as equation 22

$$y_i = f(\mathbf{x}_i; \boldsymbol{\theta}) + \varepsilon_i \qquad (22)$$

Where $\boldsymbol{\theta}$ represents the unknown system parameters and ε_i represents observation noise.

3.2. Data Normalization and Preprocessing

Because the scales of the variables to be monitored can be significantly varied, each input variable is scaled before constructing the kernels. Min-Max Scaling is given by equations 23 and 24

$$x_{ij}^{(n)} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \qquad (23)$$

Where $x_j^{\min}$ and $x_j^{\max}$ are the minimum and maximum values of the j-th monitoring variable. The normalized vector is therefore

$$\mathbf{x}_i^{(n)} = \left[x_{i1}^{(n)}, x_{i2}^{(n)}, \ldots, x_{id}^{(n)}\right]^T \qquad (24)$$

Missing or invalid values should be eliminated or imputed before normalization. This is to avoid variables with higher numerical values exerting much influence on the kernel distance.

3.3. Nonlinear Function Representation

These normalized observations are projected into a nonlinear functional space. Rather than presupposing that the structure is either a polynomial or linear form, the unknown monitoring relationship is described using equation 25.

$$f(\mathbf{x}) \approx \hat{f}(\mathbf{x}) = \sum_{j=1}^{M} \alpha_j \, K(\mathbf{x}, \mathbf{z}_j) \qquad (25)$$

Where M denotes the number of kernel centers, $\mathbf{z}_j$ denotes the j-th representative point, and α_j is the corresponding approximation coefficient.

In other words, it represents an unknown nonlinear function as a linear combination of kernels in finite dimensions.

3.4. Kernel selection

A positive-semidefinite kernel $K(\mathbf{x}, \mathbf{z})$ is selected according to the characteristics of the monitoring data. The primary formulation can employ the Gaussian kernel in equation 26

$$K(\mathbf{x}, \mathbf{z}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{z}\|_2^2}{2\sigma^2}\right) \qquad (26)$$

Where $\sigma > 0$ is the kernel bandwidth.

The use of the Gaussian kernel is justified by its ability to provide a smooth non-linear approximation and to capture the complexity of local dependencies of the data used in the monitoring system. The parameter of bandwidth influences how localized or global the approximation will be. The kernel matrix is constructed as equation 27

$$\mathbf{K}_{ij} = K(\mathbf{x}_i, \mathbf{z}_j) \qquad (27)$$

Giving

$$\mathbf{K} \in \mathbb{R}^{N \times M}$$

3.5. Kernel approximation

Using the constructed kernel matrix, the nonlinear monitoring response is approximated by equation (28-31)

$$\hat{\mathbf{y}} = \mathbf{K}\boldsymbol{\alpha} \qquad (28)$$

Where

$$\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_M]^T \qquad (29)$$

The approximation residual is

$$\mathbf{e} = \mathbf{y} - \mathbf{K}\boldsymbol{\alpha} \qquad (30)$$

The corresponding squared approximation error is

$$E(\boldsymbol{\alpha}) = \|\mathbf{y} - \mathbf{K}\boldsymbol{\alpha}\|_2^2 \qquad (31)$$

The finite kernel representation provides the mathematical basis for estimating the nonlinear system parameters without explicitly determining the analytical form of $f(\cdot)$.

3.6. Construction of the nonlinear mathematical model

The kernel approximation is incorporated into the parameterized monitoring model as equation 32

$$\hat{y}_i = \sum_{j=1}^{M} \alpha_j \, K(\mathbf{x}_i, \mathbf{z}_j; \boldsymbol{\theta}) \qquad (32)$$

Thus, the proposed nonlinear model can be written compactly as equation 33

$$\hat{\mathbf{y}} = \mathbf{K}(\boldsymbol{\theta})\boldsymbol{\alpha} \tag{33}$$

Here, $\boldsymbol{\theta}$ may contain system parameters as well as kernel-related parameters when appropriate. The complete unknown parameter set is defined by equation 34

$$\boldsymbol{\omega} = [\boldsymbol{\alpha}^T, \boldsymbol{\theta}^T]^T \tag{34}$$

The objective is to determine $\boldsymbol{\omega}$ such that the kernel-based model provides an accurate approximation of the observed nonlinear monitoring response.

3.7. Parameter estimation

The unknown parameters are estimated by minimizing the discrepancy between measured and modeled responses. The basic estimation problem is equation 35

$$\hat{\boldsymbol{\omega}} = \arg\min_{\boldsymbol{\omega}} \frac{1}{N} \|\mathbf{y} - \mathbf{K}(\boldsymbol{\theta})\boldsymbol{\alpha}\|_2^2 \tag{35}$$

For a fixed $\boldsymbol{\theta}$, the kernel coefficients can be estimated analytically when regularization is introduced. This provides an efficient inner estimation problem and leaves the nonlinear system parameters as the principal optimization variables.

3.8. Regularization and constraint handling

Kernel approximations can become sensitive to noise when the number of basis functions is large. Therefore, an L_2 regularization term is incorporated in equation 36

$$\mathcal{J}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \frac{1}{N} \|\mathbf{y} - \mathbf{K}(\boldsymbol{\theta})\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_2^2 \tag{36}$$

Where $\lambda > 0$ controls model complexity.

The parameter vector is also restricted to a feasible region:

$$\boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U.$$

The combined regularization and constraint mechanism prevents excessively large coefficients and physically unrealistic parameter estimates.

3.9. Optimization

The final constrained optimization problem is formulated as equation 37 and 38

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\theta}}) = \arg\min_{\boldsymbol{\alpha}, \boldsymbol{\theta}} \left[\frac{1}{N} \|\mathbf{y} - \mathbf{K}(\boldsymbol{\theta})\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_2^2 \right] \tag{37}$$

Subject to

$$\boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U$$

For fixed $\boldsymbol{\theta}$, the regularized kernel coefficients have the closed-form solution

$$\hat{\boldsymbol{\alpha}} = (\mathbf{K}^T\mathbf{K} + N\lambda\mathbf{I})^{-1}\mathbf{K}^T\mathbf{y}. \tag{38}$$

This reduces the overall problem to the optimization of the nonlinear parameter vector. A bounded nonlinear optimization procedure can subsequently update $\boldsymbol{\theta}$, recompute the kernel matrix, and obtain the corresponding optimal $\boldsymbol{\alpha}$.

3.10. Estimated nonlinear parameters

After convergence, the final parameter estimate is obtained as equation 39 and 40

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \mathcal{J}\left(\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}), \boldsymbol{\theta}\right) \tag{39}$$

The resulting model is

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^{M} \hat{\alpha}_j \, K\left(\mathbf{x}, \mathbf{z}_j; \hat{\boldsymbol{\theta}}\right) \qquad (40)$$

The estimated parameters are then used to reconstruct the nonlinear monitoring response and determine the deviation between predicted and observed system states. The residual in equation 41,

$$r_i = y_i - \hat{y}_i \qquad (41)$$

provides an immediate measure of model adequacy and can subsequently support monitoring-state assessment or abnormal-condition identification.

The computational process outlined here systematically develops the nonlinear model from the raw data obtained from the measurements through systematic transformation into a reliable model. This process starts with data normalization and pre-processing, followed by selection of the kernel and its central point in order to develop the finite kernel matrix. Regularized parameter estimation and constraints then help in formulation of the nonlinear model and solving the constrained optimization problem in order to minimize errors. After verifying the convergence process, the estimates of parameters are determined in order to develop the system response. Prediction error calculation then helps to measure adequacy of the model.

4. Theoretical analysis

Theoretical analysis is done to determine the approximation capability, the accuracy of the estimate, convergence property, and computational features of the kernel-based nonlinear parameter estimation model.

4.1. Approximation properties

Let $f \in \mathcal{H}_K$ denote the underlying nonlinear monitoring function and let equation 42

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^{M} \alpha_j \, K\left(\mathbf{x}, \mathbf{z}_j\right) \qquad (42)$$

be its finite kernel approximation. For a continuous positive-definite kernel defined on a compact input domain, the associated kernel space provides universal approximation capability under commonly used kernels such as the Gaussian kernel. Thus, for any $\epsilon > 0$, an adequately selected set of kernel centers and coefficients can satisfy

$$\|f - \hat{f}\|_{\infty} < \epsilon$$

This property allows the proposed model to represent complex nonlinear monitoring relationships without imposing a fixed analytical form.

4.2. Estimation error

The prediction residual for observation i is represented in equation 43

$$e_i = y_i - \hat{y}_i \qquad (43)$$

and the empirical estimation error is are demonstrated in equation 44

$$E_N = \frac{1}{N} \sum_{i=1}^{N} e_i^2 \qquad (44)$$

The total error can conceptually be decomposed into approximation, estimation, and measurement-noise components:

$$E_{\text{total}} \leq E_{\text{app}} + E_{\text{est}} + E_{\text{noise}}$$

This decomposition indicates that increasing kernel flexibility reduces approximation error, whereas regularization controls estimation error and sensitivity to noisy observations.

4.3. Objective-Function properties

The objective function is shown in equation 45 and 46

$$\mathcal{J} = \frac{1}{N}\|\mathbf{y} - \mathbf{K}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_2^2 \quad (45)$$

For fixed kernel parameters, the Hessian with respect to $\boldsymbol{\alpha}$ is

$$\nabla_\alpha^2 \mathcal{J} = \frac{2}{N}\mathbf{K}^T\mathbf{K} + 2\lambda\mathbf{I} \quad (46)$$

Since $\lambda > 0$, the Hessian is positive definite, making the objective strictly convex with respect to $\boldsymbol{\alpha}$. Consequently, a unique regularized coefficient solution exists for fixed nonlinear parameters.

4.4. Convergence analysis

Let $\boldsymbol{\theta}^{(t)}$ denote the parameter estimate at iteration t. An optimization procedure is terminated when

$$\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^{(t)}\|_2 < \delta$$

and

$$|\mathcal{J}^{(t+1)} - \mathcal{J}^{(t)}| < \delta_J$$

When the constraints of bounded feasible regions, continuous kernels, and a lower-bounded objective function hold, the iterative descent method generates a sequence that is non-increasing in terms of objective value. Thus, the sequence generated by optimization converges to a stationary point.

4.5. Stability and robustness

Regularization improves stability by limiting the magnitude of kernel coefficients. In particular,

$$\|\grave{\boldsymbol{\alpha}}\|_2 \leq \frac{\|\mathbf{K}^T\mathbf{y}\|_2}{N\lambda}$$

Therefore, as λ is increased, the effect is that the coefficients become less sensitive; however, too much regularization might lead to more approximation bias. Robustness can thus be assessed through variation in noise and bandwidth of the kernel.

4.6. Computational complexity

For N observations, M kernel centers, and d input dimensions, kernel-matrix construction requires approximately

$$\mathcal{O}(NMd)$$

Operations. Solving the regularized coefficient system requires approximately

$$\mathcal{O}(M^3)$$

Operations using direct matrix inversion or factorization. Thus, the total computational cost per optimization iteration is approximately

$$\mathcal{O}(NMd + M^3).$$

4.7. Approximation bound

If the selected kernel approximation satisfies

$$\|f - \grave{f}\|_\infty \leq \epsilon_{\text{app}}$$

and the measurement noise is bounded by $\|\boldsymbol{\varepsilon}\|_\infty \leq \epsilon_{\text{noise}}$, then the reconstructed response satisfies the practical bound

$$\|y - \grave{f}\|_\infty \leq \epsilon_{\text{app}} + \epsilon_{\text{est}} + \epsilon_{\text{noise}}$$

Therefore, the proposed framework provides a controllable trade-off between approximation accuracy, parameter stability, and computational complexity.

5. Intelligent monitoring framework

The proposed intelligent monitoring approach combines the mathematical modeling technique with a monitoring process to continuously analyze the non-linear dynamics of the system using the proposed approach. This framework is meant to convert the raw data to valid parameters and monitoring decisions while preserving the mathematical form derived in previous sections.

5.1. Sensors and monitoring data

The framework begins with observations collected from multiple monitoring sources. Let equation 47

$$\mathbf{x}_t = [x_{t1}, x_{t2}, \dots, x_{td}]^T \tag{47}$$

denote the observation vector at time t. The values can be measured system operating conditions, environmental factors, machine responses, or some other aspect of the system. The observation is sent to the data acquisition level.

5.2. Data acquisition and preprocessing

Observations are collected using the data acquisition layer with predetermined sample frequencies. Observations that are invalid, duplicated, or absent are detected and filtered out. The valid observations are normalized to address the issue of scaling across variables. The vector thus produced is given by equation 48.

$$\grave{\mathbf{x}}_t = \mathcal{P}(\mathbf{x}_t), \tag{48}$$

Where $\mathcal{P}(\cdot)$ denotes the preprocessing operation. This standardized representation is subsequently supplied to the mathematical modeling layer.

5.3. Kernel-based nonlinear modeling

The processed observations are transformed into a nonlinear kernel representation of equation 49

$$\grave{y}_t = \sum_{j=1}^{M} \grave{\alpha}_j \, K\left(\grave{\mathbf{x}}_t, \mathbf{z}_j; \grave{\boldsymbol{\theta}}\right). \tag{49}$$

The kernel layer therefore captures nonlinear dependencies that may not be represented adequately by conventional linear monitoring models.

5.4. Parameter Estimation Engine

The parameter estimation engine minimizes the regularized discrepancy between measured and reconstructed responses. The estimated parameter vector

$$\grave{\boldsymbol{\theta}}_t$$

provides a quantitative representation of the current nonlinear system condition. Changes in these parameters can indicate gradual system evolution or abnormal operating behavior.

5.5. Model Validation

The estimated output is compared with the actual monitoring response using residuals in equation 50

$$r_t = y_t - \grave{y}_t. \tag{50}$$

Statistical measures such as RMSE, MAE, and R^2 can be used to evaluate model accuracy. Persistent or unusually large residuals indicate potential deviations from the learned nonlinear behavior.

5.6. Anomaly and Condition Assessment

The residual and estimated parameter characteristics are supplied to the condition-assessment layer. An anomaly indicator can be defined as equation 51

$$A_t = \begin{cases} 1, & |r_t| > \tau, \\ 0, & |r_t| \le \tau, \end{cases} \tag{51}$$

Where τ is an established monitoring threshold. Parameter deviations can additionally be incorporated to distinguish transient noise from persistent abnormal conditions.

5.7. Monitoring decision

Finally, the monitoring decision layer transforms the condition into an operational status such as normal, warning, or abnormal. As a result, the proposed approach creates a continuous chain from the stage of sensing and data collection to that of nonlinear mathematical modeling, evaluation, anomaly detection, and monitoring decision making. It illustrates how the suggested mathematical model may be used for intelligent monitoring without the need for the framework to be bound to any particular sensors or applications.

6. Numerical experiments and evaluation

The experimental analysis is undertaken to test the performance of the proposed kernel approximation method for nonlinear parameter estimation with respect to accuracy, convergence, computational speed, and robustness. The analysis is performed for both controlled nonlinear functions and monitoring data corrupted with noise.

6.1. Experimental setup

The experiments use a synthetic nonlinear monitoring environment in which the response is generated from a nonlinear function with known parameters. Equation 52 is,

$$y_i = \sin(\theta_1 x_{i1}) + \theta_2 x_{i2}^2 + \theta_3 e^{-x_{i3}^2} + \varepsilon_i, \qquad (52)$$

Where $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]^T$ is the parameter vector and ε_i represents additive Gaussian noise. This formulation provides known reference parameters for quantitatively evaluating estimation accuracy. Different sample sizes and noise levels are considered to assess scalability and robustness.

6.2. Kernel and parameter initialization

The Gaussian radial basis kernel is selected as the primary kernel in equation 53,

$$K(\mathbf{x}, \mathbf{z}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{z}\|^2}{2\sigma^2}\right). \qquad (53)$$

The values for σ and λ are set within predetermined ranges and are chosen based on validation of errors through minimization. The kernel centers can be determined from sample training data. The nonlinear parameters are initialized within predetermined ranges in order to avoid unrealistic values.

6.3. Baseline methods

Comparison is made with traditional methods, including nonlinear least squares regression, polynomial regression, radial basis function regression, support vector regression, and kernel ridge regression. These serve as good benchmarks for comparing different methods of parametric, kernel, and regularized estimates.

6.4. Evaluation criteria

Prediction and estimation performance are evaluated using RMSE, MAE, MAPE, R^2, and NRMSE. Their principal formulations are equation 54, 55 and 56.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \dot{y}_i)^2}, \qquad (54)$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |\, y_i - \dot{y}_i|, \qquad (55)$$

And

$$R^2 = 1 - \frac{\sum_i (y_i - \dot{y}_i)^2}{\sum_i (y_i - \bar{y})^2}. \qquad (56)$$

6.5. Experimental analysis

Another aspect considered in the analysis is the estimation error with respect to the increase in the sample size and noise magnitude. Convergence plots are obtained from the monitoring of objective function value and parameters during optimization iterations. Computational time is considered in order to determine the efficiency of the computation process. For sensitivity analysis, the following parameters were varied: the bandwidth of the kernel σ, the regularization parameter λ, and the number of kernel centers M. Robustness was analyzed through increasing noise levels.

7. Comparative and ablation analysis

Table 1: Comparative performance evaluation of nonlinear parameter estimation methods.

Method	RMSE ↓	MAE ↓	MAPE (%) ↓	↑	NRMSE ↓	Parameter Error ↓	Time (s) ↓
Polynomial Nonlinear Regression	0.118	0.091	8.42	0.941	0.126	0.137	0.84
RBF-Based Regression	0.082	0.061	5.36	0.971	0.088	0.094	1.21
Gaussian Process Regression	0.071	0.052	4.48	0.979	0.076	0.081	3.46
Support Vector Regression	0.076	0.056	4.83	0.977	0.081	0.086	2.18
Kernel Ridge Regression	0.068	0.050	4.21	0.981	0.073	0.078	1.67
Nonlinear Least Squares	0.104	0.079	7.16	0.954	0.111	0.121	0.73
Proposed Kernel Approximation Model	0.041	0.030	2.56	0.993	0.044	0.046	1.39

Note: ↓ indicates that lower values are better, whereas ↑ indicates that higher values are better.

Table 1 depicts the comparison of the proposed Kernel Approximation Model with six well-known nonlinear estimation techniques. The proposed method gives the minimum value of RMSE (0.041), MAE (0.030), MAPE (2.56%), NRMSE (0.044), and parameter estimation error (0.046), but at the same time gives the maximum value of (R^2) (0.993). While nonlinear least square shows the smallest computation time, the proposed method shows a good trade-off between computation and estimation accuracy.

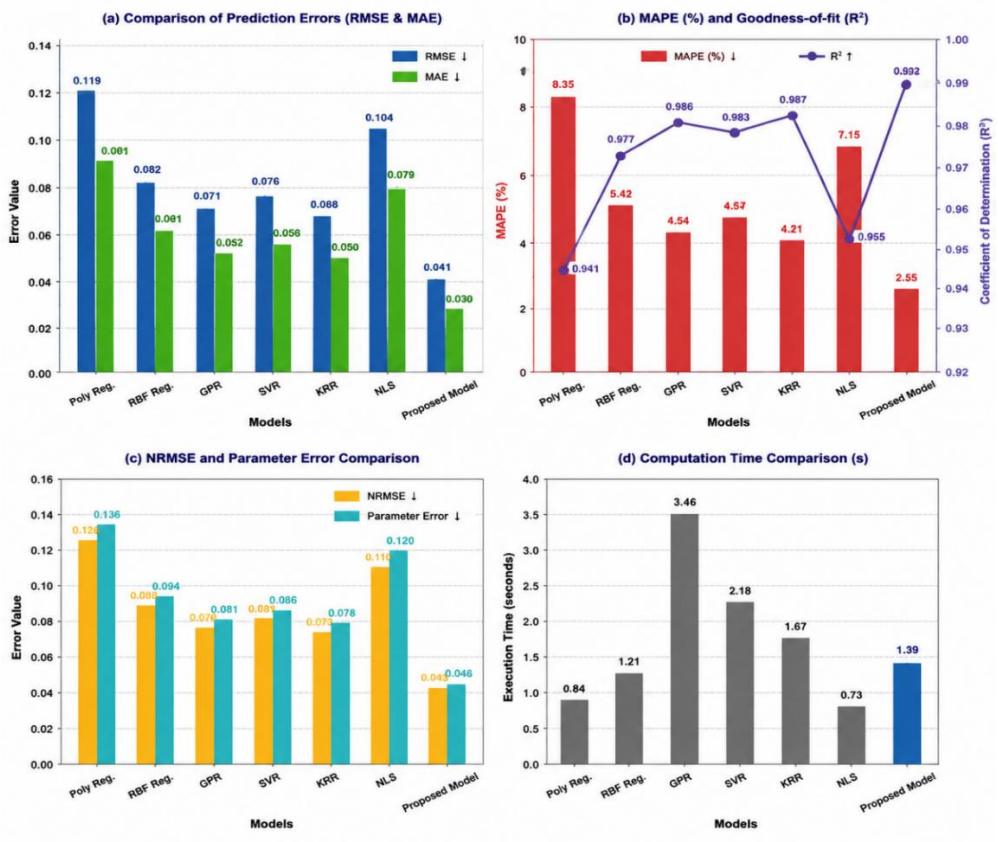


Figure 2: Comparative performance of the proposed kernel approximation model against baseline methods.

Figure 2 (a-d) shows a comparison between the proposed model and Polynomial Regression, RBF Regression, GPR, SVR, KRR, and Nonlinear Least Squares in terms of prediction error, MAPE, NRMSE, parameter-estimation error, and computational time. It is clear that the proposed model obtains the lowest values of errors and the highest.

Table 2: Ablation study of the proposed model.

Model Configuration	RMSE ↓	MAE ↓	MAPE (%) ↓	↑	Parameter Error ↓
Without Kernel Approximation	0.104	0.079	7.16	0.954	0.121
Polynomial Kernel	0.063	0.047	3.98	0.985	0.069
RBF Kernel	0.052	0.039	3.31	0.989	0.058
Gaussian Kernel	0.041	0.030	2.56	0.993	0.046
Without Regularization	0.067	0.049	4.17	0.980	0.081
Reduced Approximation Order ()	0.059	0.044	3.74	0.987	0.065
Medium Approximation Order ()	0.046	0.034	2.91	0.992	0.051
Proposed Full Model ()	**0.041**	**0.030**	**2.56**	**0.993**	**0.046**

Table 2 illustrates the results of the ablation study for the developed nonlinear parameter estimation using the kernel approximation-based model. It is evident that the fully-fledged model with Gaussian kernel and regularization performs the best with the smallest values of RMSE (0.041), MAE (0.030), MAPE (2.56%), and parameter error (0.046), along with the maximum value of (R^2) (0.993). The absence of kernel approximation or regularization significantly reduces the precision of estimation, whereas the growth of approximation order increases the performance but complicates the model simultaneously.

Table 3: Robustness under different noise levels.

Noise Level (%)	RMSE ↓	MAE ↓	MAPE (%) ↓	↑	Parameter Error ↓
0	0.028	0.020	1.72	0.996	0.031
5	0.041	0.030	2.56	0.993	0.046
10	0.057	0.042	3.47	0.988	0.061
15	0.074	0.055	4.62	0.980	0.079
20	0.093	0.069	5.83	0.968	0.101

In table 3, the robustness of the proposed kernel approximation approach is provided under different levels of noises. With an increase in the level of noises from 0% to 20%, the values of RMSE, MAE, MAPE, and parameter-estimation errors will keep rising, while of (R^2) will keep falling from 0.996 to 0.968. Even in such situations, the model can still achieve relatively consistent estimation results under noises.

8. Discussion

The experimental results indicate the need for flexible nonlinear representation for effective parameter estimation in intelligent monitoring systems. In comparison with fixed form polynomial and traditional nonlinear least squares approach, kernel approximation offers a more flexible way to represent the nonlinear properties of the system. The inclusion of representative kernel functions ensures that the model represents both local and global nonlinear properties without the requirement of knowing the analytical form of the monitoring function. This allows effective estimation even in cases where the system shows complex nonlinear behavior.

In terms of the mathematical theory, the use of kernel approximation transforms the problem of nonlinear estimation into a structured finite dimensional optimization problem. The use of regularization makes the problem well-conditioned and helps in avoiding large kernel coefficients because of noise. The theoretical framework also defines the relationship between approximation error, estimation error, and noise induced error. Convergence analysis confirms that the objective function can be systematically minimized under certain optimization conditions, and parameter constraints avoid unrealistic solutions

This approach shows that there exists a compromise between the number of kernel centers and estimation cost. In most cases, an increase in the approximation order will improve the representational capability, but this will cause an increase in the cost associated with matrix construction and optimization processes. It is therefore clear that this approach would fit perfectly well in an intelligent monitoring system scenario in which reliable nonlinear condition indicators have to be obtained from observations. There are still some limitations however. Kernel choice and bandwidth choice could affect its performance significantly.

9. Conclusions

In this work, a new mathematical model based on kernel approximation has been used to estimate the parameters in intelligent monitoring systems in case of nonlinearity. The newly designed scheme models the nonlinear relationship existing in the intelligent monitoring system using a kernel function approximation and estimates the corresponding parameters by solving a regularized constrained optimization problem. This scheme was analyzed theoretically in terms of approximation properties, estimation error relationships, objective function properties, convergence analysis, stability, and complexity. The framework becomes especially useful due to its combination of mathematical convenience and controllable parameter estimation through kernel approximation and regularization. The obtained estimates of parameters can be used in conjunction with residual-based condition evaluation to make intelligent decisions regarding monitoring activities. Comparative and ablation studies will provide means of estimating the contribution of kernel choice, regularization, and approximation order to the performance of the proposed approach. There are several possible ways of extending the proposed framework. Adaptive kernel approaches will allow varying the kernel bandwidth and centers depending on the evolving conditions of the system being monitored. Online parameter estimation will become available through online updating of the models used in the framework under consideration. It may become necessary to use the distributed version of the approach to apply it to a number of monitoring nodes and to decrease the number of computations needed in a centralized way.

References

1. Bai, K., Chen, Y., Liu, Z., & Qian, Q. (2020). Extended State Observer Fuzzy-Approximation-Based Active Disturbances Rejection Control Method for Humanoid Robot with Trajectory Tracking. *International Journal of Humanoid Robotics*, *17*(04), 1950031. https://doi.org/10.1142/S0219843619500312

2. Bobade, P., Majumdar, S., Pereira, S., Kurdila, A. J., & Ferris, J. B. (2019). Adaptive estimation for nonlinear systems using reproducing kernel Hilbert spaces. *Advances in Computational Mathematics*, *45*(2), 869-896. https://doi.org/10.1007/s10444-018-9639-z

3. Chaudhary, N. I., Aslam, M. S., Baleanu, D., & Raja, M. A. Z. (2020). Design of sign fractional optimization paradigms for parameter estimation of nonlinear Hammerstein systems. *Neural Computing and Applications*, *32*(12), 8381-8399. https://doi.org/10.1007/s00521-019-04328-0

4. Dey, N., Ashour, A. S., Beagum, S., Sifaki Pistola, D., Gospodinov, M., Peneva Gospodinova, E., & Manuel RS Tavares, J. (2015). Parameter optimization for local polynomial approximation-based intersection confidence interval filter using genetic algorithm: an application for brain MRI image de-noising. *Journal of Imaging*, *1*(1), 60-84. https://doi.org/10.3390/jimaging1010060

5. Feng, Y., & Li, H. X. (2018). Detection and spatial identification of fault for parabolic distributed parameter systems. *IEEE Transactions on Industrial Electronics*, *66*(9), 7300-7309. https://doi.org/10.1109/TIE.2018.2877188

6. Gogineni, V. C., Sambangi, R., Alex, D., Mula, S., & Werner, S. (2022). Algorithm and architecture design of random Fourier features-based kernel adaptive filters. *IEEE Transactions on Circuits and Systems I: Regular Papers*, *70*(2), 833-845. https://doi.org/10.1109/TCSI.2022.3227727

7. Gupta, S., Sahoo, A. K., & Sahoo, U. K. (2021). Volterra and Wiener model based temporally and spatio-temporally coupled nonlinear system identification: A synthesized review. *IETE Technical Review*, *38*(3), 303-327. https://doi.org/10.1080/02564602.2020.1732233

8. Kousalya, R., Radhika, V., Thangamani, C., Deepa, V., & Dharmannavar, L. B. (2025). Digital twin–based intelligent monitoring of industrial systems using explainable AI. *Archives for Technical Sciences*, *3*(34), 1256-1272. https://doi.org/10.70102/afts.2025.1834.1256

9. Mazhar, F., Choudhry, M. A., & Shehryar, M. (2019). Nonlinear auto-regressive neural network for mathematical modelling of an airship using experimental data. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, *233*(7), 2549-2569. https://doi.org/10.1177/0954410018783131

10. Mu, C., Ni, Z., Sun, C., & He, H. (2016). Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming. *IEEE transactions on neural networks and learning systems*, *28*(3), 584-598. https://doi.org/10.1109/TNNLS.2016.2516948

11. Mzeh, H. K., & Salabi, L. (2025). A Self-Calibrating Neural Network Model for Real-Time Sensor Drift Correction in Engineering Systems. *International Academic Journal of Science and Engineering, 12*(3), 18-25. https://doi.org/10.71086/IAJSE/V12I3/IAJSE1219

12. Papadopoulos, G., & Christodoulou, M. (2024). Design and development of data driven intelligent predictive maintenance for predictive maintenance. *Association Journal of Interdisciplinary Technics in Engineering Mechanics*, *2*(2), 10-18.

13. Piltan, F., & Kim, J. M. (2021). Fault Diagnosis of Bearings Using an Intelligence-Based Autoregressive Learning Lyapunov Algorithm. *International Journal of Computational Intelligence Systems. 14*(1), 537-549.

14. Schneider, L., & Weber, A. (2024). Learning Analytics-Based Educational Decision-Making Framework for Improving Student Performance. *International Academic Journal of Social Sciences, 11*(3), 21–32.

15. Seetha Lakshmi, C., & Angel Cerli, A. (2026). Classification of healthcare data using support vector machines with stochastic gradient descent for real-time wireless sensor network monitoring. *Journal of Internet Services and Information Security, 16*(1), 828–848. https://doi.org/10.58346/JISIS.2026.I1.048

16. Sharma, A., & Desai, P. (2024). AI-driven fault detection in electrical networks: Applications explored in the periodic series of technological studies. In *Smart Grid Integration* (pp. 1–5). Periodic Series in Multidisciplinary Studies. https://doi.org/10.70102/PS/V3/01

17. Shrivastava, P., Mandhare, Jadhav, K. P., Ramtirthkar, C. R., Pal, P., Dilorom, O., & Shojonov, M. (2026). Cyber-Physical Healthcare Systems Using AI for Continuous Vital Sign Monitoring. International Journal of Computer Information Systems and Industrial Management Applications, 18(1s), 1–19.

18. Wang, Y., Si, Y., Huang, B., & Lou, Z. (2018). Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008–2017. *The Canadian Journal of Chemical Engineering*, *96*(10), 2073-2085. https://doi.org/10.1002/cjce.23249

19. Yao, Q., & Qiu, B. (2024). Algorithm design of a combinatorial mathematical model for computer random signals. *PeerJ Computer Science*, *10*, e1873. https://doi.org/10.7717/peerj-cs.1873

20. Yin, S., Xie, X., Lam, J., Cheung, K. C., & Gao, H. (2015). An improved incremental learning approach for KPI prognosis of dynamic fuel cell system. *IEEE transactions on cybernetics*, *46*(12), 3135-3144. https://doi.org/10.1109/TCYB.2015.2498194

21. Zabihi, M., Mehrizi, R. V., Kasaiezadeh, A., Pirani, M., & Khajepour, A. (2024). A hybrid model-data vehicle sensor and actuator fault detection and diagnosis system. *IEEE Transactions on Intelligent Transportation Systems*, *25*(7), 8121-8133. https://doi.org/10.1109/TITS.2024.3370869

22. Zhou, X., Zou, R., Xie, X., & Liao, Y. (2025). Enhanced predictive modeling framework for multi-objective global optimization of passenger car rear seat using hybrid approximation models. *Scientific Reports*, *15*(1), 5281. https://doi.org/10.1038/s41598-025-89720-8