Neuro-ICA Based Process Monitoring Strategy for Fault Detection in Steel Billets Manufacturing Unit

Bhagwan Kumar Mishra[1], Sanjay Kumar[2], Pallavi Sindhu[3], Anupam Das[4], Saroj Kumar[5]*
[1,2]Department of Mechanical Engineering, ISBM College of Engineering, Nande, Pune-412115.Maharastra, India.
[3]Department of Mechanical Engineering, Government Engineering College Sheikhpura, Bihar, India
[4,5]Department of Mechanical Engineering, National Institute of Technology Patna, Patna, Bihar, India

Abstract. The proposed monitoring strategy addresses the challenges associated with non-linear process behavior through the application of the Neural Network Fitting (NNF) Model technique, while non-Gaussian process characteristics are effectively managed using the I^2 control chart. The study focuses on developing a monitoring strategy that employs statistical techniques while accounting for nonlinear as well as non-normal or non-Gaussian data. The proposed statistical monitoring strategy employs a multivariate I^2 control chart for non-normal data, while data nonlinearity is addressed using a Neural Network Fitting model. Quality characteristics of steel billets are initially processed through a Neural Network model to mitigate nonlinear patterns. The fully or partially linearized data are then evaluated using the I^2 control chart, with its control limits determined through the Bootstrap procedure. Observations identified as out-of-control are classified as faults, and their detection prompts the implementation of suitable corrective measures.

1 Introduction

Statistical Process Monitoring (SPM) is a real-time methodology for evaluating system performance by collecting process data, detecting deviations, and signaling abnormal behavior to ensure good quality eventual product. SPM [1-5] strategies are broadly categorized into univariate and multivariate approaches, depending on the number of process characteristics under observation. Univariate SPM monitors a single characteristic at a time, while multivariate SPM simultaneously evaluates two or more correlated characteristics.

In univariate SPM, traditional Shewhart, EWMA (Exponentially Weighted Moving Average), and CUSUM(Cumulative Sum) control charts are commonly employed, whereas multivariate SPM typically utilizes Hotelling's T^2 [6-10] and Chi-square charts. When many highly correlated characteristics must be monitored, projection-based techniques such as PCA (Principal Component Analysis) and PLSR (Partial Least Squares Regression) are commonly applied. However, PCA and PLSR based multivariate SPM approaches [11-15] are inherently suited to linear processes, as they assume linear relationships among variables. Applying such linear projection methods to nonlinear processes may result in misleading outcomes, since real-world processes rarely exhibit strict linearity. To address nonlinearity, several nonlinear extensions have been proposed, including KPCA (Kernel Principal Component Analysis) [16-19] and NNF (Neural Network Fitting) based models [20-24].

* Corresponding author:bhagwanm.phd17.me@nitp.ac.in

Multivariate projection-based methods such as PCA and PLSR are widely applied in process monitoring strategies for systems with many quality variables. However, traditional MSPM approaches assume that the data associated with PFC observations follow a normal distribution. In practice, this assumption often fails in many industrial processes, where large-scale and complex datasets exhibit non-Gaussian behavior. Relying on normality in such cases may therefore yield misleading results. To address this issue, recent research has focused on non-Gaussian variants of MSPM strategies for fault detection and diagnosis in complex manufacturing systems. These include ICA (Independent Component Analysis), SVDD (Support Vector Data Description), and GMM (Gaussian Mixture Models). Among these, ICA has been extensively applied as a non-Gaussian alternative, as it decomposes original variables into statistically Independent Components (ICs) and performs fault detection.

ICA [25-30] is a multivariate, non-Gaussian technique that transforms statistical data into statistically independent components, which are represented as linear combinations of the observed variables. Unlike PCA, which achieves independence only up to second-order correlations by enforcing orthogonality, ICA captures higher-order dependencies, thereby extracting more meaningful information from non-Gaussian data. SVDD [31-34], in contrast, requires no distributional assumptions and identifies a minimum hypersphere enclosing most of the data, which is then used to determine whether new observations remain within control limits. GMM [35-38] provide a probabilistic framework for non-Gaussian MSPM by representing process data as a mixture of local linear models, with parameters estimated through the iterative expectation–maximization (EM) algorithm. Most of the statistical multivariate non-Gaussian techniques, ICA technique and its variants are widely used for the process monitoring. In contrast to other non-Gaussian techniques, ICA technique has simple model structure and easier mathematical formulation. In contrast, GMM techniques display difficulties in the process of finding optimal number of local linear models that are required to characterize the data under consideration. Similarly, SVDD technics poses difficulties for selecting an appropriate kernel function, and the said technique is perceptive to producing false alarms in fault detection owing to the imposition of overly stringent control limits.

In neural network–based monitoring strategies, unsupervised NNF models are typically employed to transform nonlinear datasets into partially or fully linear representations. An NNF model comprises of three layers are viz. an input layer, a hidden layer, and an output layer. The nonlinear process data is first received by the input layer, after which the hidden layer employs a log-sigmoid activation function to extract variable characteristics. Subsequently, the output layer performs a linear transformation to convert the nonlinear input into a linearized dataset appropriate for process monitoring.

The present study aims to develop a nonlinear and non-Gaussian statistical process monitoring strategy for a SBMU (Steel Billets Manufacturing Unit). An unsupervised NNF model was employed to transform the nonlinear dataset into a partially or fully linear form. The resulting linearized data were then used to design a non-Gaussian monitoring framework based on the ICA technique. Fault detection was carried out using an I^2 control chart constructed from ICA scores to identify out-of-control observations. The monitoring focused on quality characteristics corresponding to the chemical composition of the steel billets produced in the SBMU.

The novelty of the current work lies in the introduction of a two-stage monitoring framework designed to handle nonlinear and non-Gaussian process data. In the first stage, the NNF model is employed to convert nonlinear data into partially or fully linear forms. In the second stage, an ICA-scores-based I^2 control chart is used to detect out-of-control observations, which are characterized as faults. This novel approach effectively overcomes the limitations of existing monitoring techniques such as PCA, PLSR, ICA, SVDD, GMM, and others.

2 NNF Model and I^2 Control Chart

The NNF model, also know as a Feedforward Neural Network, is used to transform data exhibiting nonlinear behavior into partially or fully linearized data. The linearized data obtained from the NNF model are then used to develop an ICA scores based non-Gaussian process monitoring strategy, which overcomes the limitations of traditional MSPM techniques such as PCA, PLSR, and ICA. The NNF model is employed to transform nonlinear datasets into linear representations suitable for statistical monitoring. Structurally, the NNF model consists of an input layer (represents the independent variables), a hidden layer (represents intermediate layers between input and output layer, also called transformation layer) and an output layer (class prediction layer).

Prior to training, appropriate initial values of weights and biases must be assigned, which depend on the network architecture. In the case of multilayer networks, these parameters are typically initialized with small random values in the range of –0.5 to 0.5. The input layer serves as the entry point, feeding the nonlinear data into the network. The hidden layer captures and analyzes the relationships among variables using the log-sigmoid activation function, a common choice in multilayer networks trained with the back-propagation algorithm. This activation function accepts inputs spanning from negative to positive infinity and transforms them into outputs constrained within the range [0, 1], as defined by the following expression:

$$a = \frac{1}{1+e^{-n}} \tag{1}$$

Where $n = f(wx+b)$, x is a quality characteristic, w is a weight matrix, b is a bias.

The output layer performs the transformation of the nonlinear dataset into a partially or fully linearized form by employing a linear transformation activation function. This step ensures that the processed data is suitable for subsequent statistical monitoring and analysis.

After data linearization, the ICA technique is employed to extract multiple ICs score matrices. These individual score matrices are then combined into a single comprehensive matrix for further analysis. For fault detection, I^2 control chart is used which is a non-normal variant of Hoteling's T^2 control chart. The control limits of the I^2 chart are estimated using a bootstrap-based procedure to ensure robustness under non-Gaussian conditions. The mathematical description of ICA technique is presented below.

Let the quality characteristics variables are $X_1, X_2, X_3, X_4, \ldots\ldots. X_d$, these quality variables can be represented as a linear combination of m (≤d) of unknown ICs with zero mean. The expression below represents the relationship between the quality variables and the extracted ICs.

$$X = AS \tag{2}$$

Where,

$X = [X_1, X_2, \ldots\ldots, X_d]$ is the measured quality characteristics vector.

$A = [A_1.A_2\ldots\ldots, A_m]$ is the mixing vector.

$S = [S_1, S_2, \ldots\ldots, S_m]$ is the independent component vector.

The expression to estimate the independent component matrix is represented below.

$$Y = WX \tag{3}$$

Where, Y represents ICs and W represents the demixing matrix

In the ICA technique, the whitening transformation is used to remove the cross-correlation between quality variables. The expression of whitening transformation is given below

$$Z = Q\,X \tag{4}$$

Where,
$Q = \wedge^{-1/2}\, U^T$ is represents the expression of whitening matrix whereas $\wedge$ represents the diagonal matrix of eigenvalues and U represents the orthogonal matrix of eigenvectors. Both $\wedge$ and U are derived from the Eigen-decomposition of the covariance matrix $E\,(X\,X^T) = U \wedge U^T$. Equation (4) can be further converted as

$$Z = Q\,X = Q\,A\,S = B\,S \tag{5}$$

Where,
$B_i = Q_i\, A_i$ is the expression of an orthogonal matrix, and which is given that $E\,(Z_i\, Zi^T) = B_i\, E\,(S_i\, S^T_i) = B_i\, B^T_i = I.$
After that independent component Y is calculated from equation (5), which is further expressed as

$$Y = B^T Z = B^T QX \tag{6}$$

Equations (3) and (6) are used to derive the relation between W and B, which can be expressed as

$$W = B^T Q \tag{7}$$

Afterwards, for new observations deterministic part ($Y_{d\ new}$) of independent components is evaluated. L_2 norm method was employed to estimate the deterministic part of demixing matrix (W_d).

$$Y_{d\,new} = W_d\, X_{new} \tag{8}$$

In which, X new is represented as the matrix of new observations

3 Case Overview

The case study considered in this work involves a SBMU (Steel Billets Manufacturing Unit) located in eastern India. Steel billets are semi-finished products, generally square or rectangular in cross-section, with relatively small thickness and variable lengths. The SBMU under consideration operates through three main stages viz. EAR (Electric Arc Furnace), LRF (Ladle Refining Furnace) and CCM (Continuous Casting Machine). The EAF, a widely adopted technology in modern steelmaking, facilitates the rapid melting of ferrous scrap and direct reduced iron by converting electrical energy into heat through an electric arc formed between graphite electrodes and the scrap charge. This process allows for precise control of melting temperature and chemical composition. The molten steel is subsequently transferred to the LRF, which serves as a critical secondary metallurgical unit. Positioned after the EAF, the LRF enables accurate control of temperature, alloy composition, and overall steel cleanliness through operations such as arc heating, gas purging, and slag conditioning, and alloying additions. In the final stage, solidification process of refined molten steel is initiated by transferring it into water-cooled copper mold in the CCM. The partially solidified strand is continuously withdrawn, guided by rollers, and cooled further with secondary water sprays until complete solidification is achieved. The resulting billets are then cut to the required length, providing a consistent and high-quality characteristics for downstream rolling operations.

For process monitoring, the quality characteristics of the steel billets produced by the SBMU were considered. These characteristics corresponded to the percent chemical composition of the billets, which directly influences product quality and downstream performance. In this study, seven quality characteristics were selected, as detailed in Table1. Observations for these seven variables were systematically collected and subsequently utilized for the development and implementation of the proposed monitoring strategy.

Table 1. Quality characteristics of the SBMU.

Sr.No	Quality Characteristics	Description
1	C	Chemical composition (in %) of carbon in the still billets
2	N_2	Chemical composition (in %) of carbon in the still billets(ppm)
3	S	Chemical composition (in %) of sulphur in the still billets
4	Al	Chemical composition (in %) of aluminum in the still billets
5	Si	Chemical composition (in %) of silicon in the still billets
6	P	Chemical composition (in %) of phosphorus in the still billets
7	Mn	Percentage composition of manganese in the still billets

4 Analysis

A total of four stages were involved in the analysis process. Initially, quality characteristics data associated with steel billets were collected and then outlier removal technique was employed for the removal of outliers. After this treatment, seventy valid observations were retained for subsequent analysis. The first stage focused on assessing the linearity of the quality characteristics. Linearity testing was performed by constructing fitted line plots between pairs of quality characteristics and calculating the R^2 (Coefficient of Determination). The R^2 coefficient serves as a quantitative indicator of the degree of linear association between two variables. For illustration, the compositions of silicon and phosphorus in the billets were selected to evaluate linearity. The fitted line plot between these two characteristics yielded an R^2 value of 0.0 percent, indicating the absence of a linear relationship and thereby confirming the nonlinear nature of the dataset.

In the second stage of analysis, NNF (Neural Network Fitted) model was used for the transformation of nonlinear data into partial or fully linear one. Following this transformation, linearity testing was repeated, yielding an R^2 value of approximately 75 percent, which indicates a reasonably linear relationship. Figure 1 represented the schematic diagram of the NNF model.

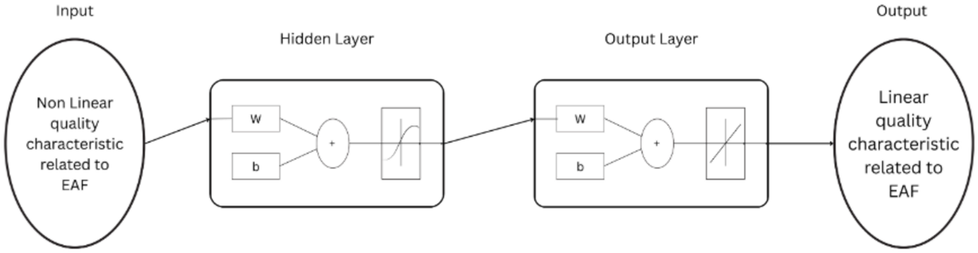

Fig. 1. Block diagram of NNF model.

In the third stage, MCLT (Multivariate Central Limit Theorem) was carried out to test the normality relation of the transformed data by the NNF model. The MCLT was carried out a hypothesis test and test result represents whether the transformed data is normally distributed or Gaussian distributed. Figure 2 presents the probability plot of the mean values of the quality characteristics. The plot clearly indicates noticeable deviations from normality. Moreover, the corresponding p-value obtained from the hypothesis test was less than 0.005, thereby leading to the rejection of the null hypothesis that assumes normally distributed data.

In the fourth stage, an I^2 control chart was constructed using the scores of ICs which are extracted by the ICA decomposition. The control limits for the I^2 chart can be estimated using KDE (Kernel Density Estimation) or the Bootstrap procedure. In this study, the Bootstrap method was preferred due to its relative simplicity, whereas KDE is computationally intensive and requires critical decisions regarding kernel function, smoothing parameters, and space points. Out of 70 observations, 46 were utilized to establish the control limit, while the remaining 24 observations were used as new observations. Figure 3 demonstrates the I^2 control chart for the 24 new observations. The Upper Control Limit (UCL) value of 9.368 on the present control chart represents the maximum threshold used to monitor process variation in the process monitoring strategy. Any value above this threshold is characterized as a fault or an out-of-control observation, while values below the threshold indicate that the ongoing process is under control.

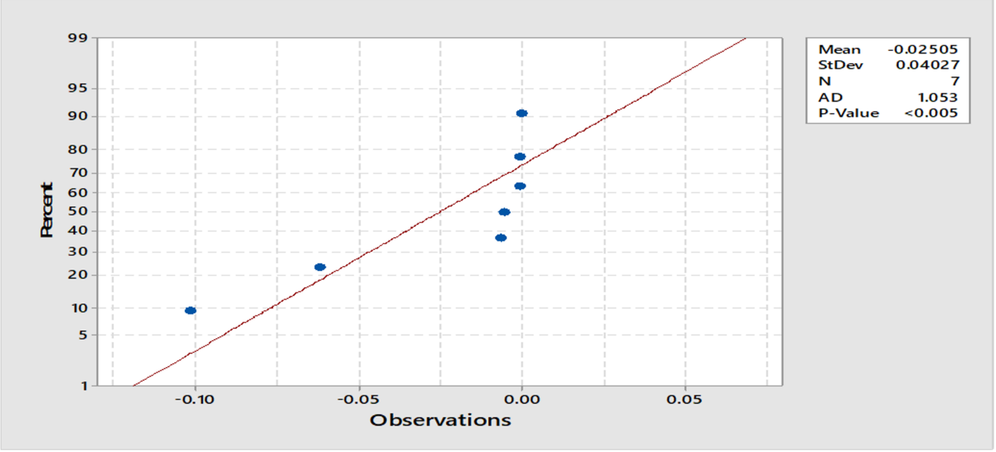

Fig. 2. Probability of the mean value of quality characteristics associated with SBMU.

As evident from figure 3, observation number 5 and 13 falls outside the control limit, indicating an out-of-control condition that necessitates further diagnosis to identify and eliminate the underlying cause. Figure 4 represents the process flow diagram of NNF-ICA based process monitoring strategy for the fault's detection in SBMU.

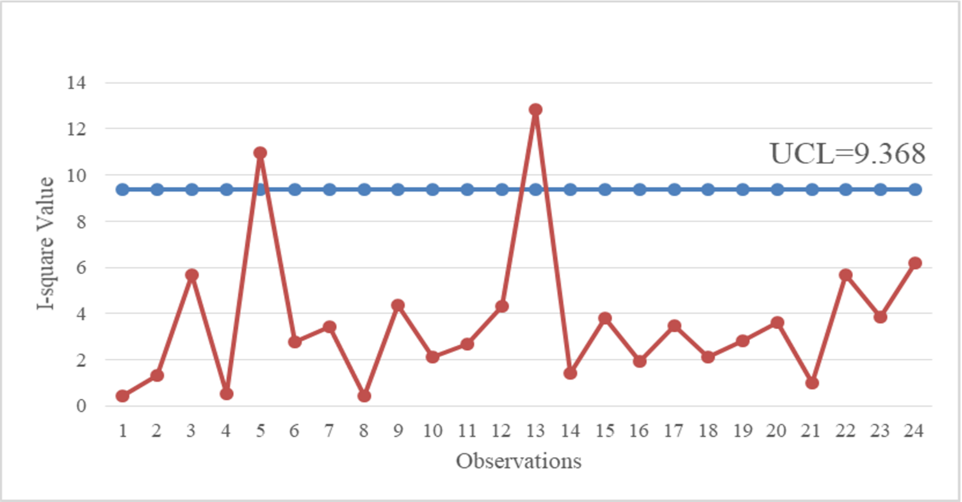

Fig. 3. I^2 control chart for new observations.

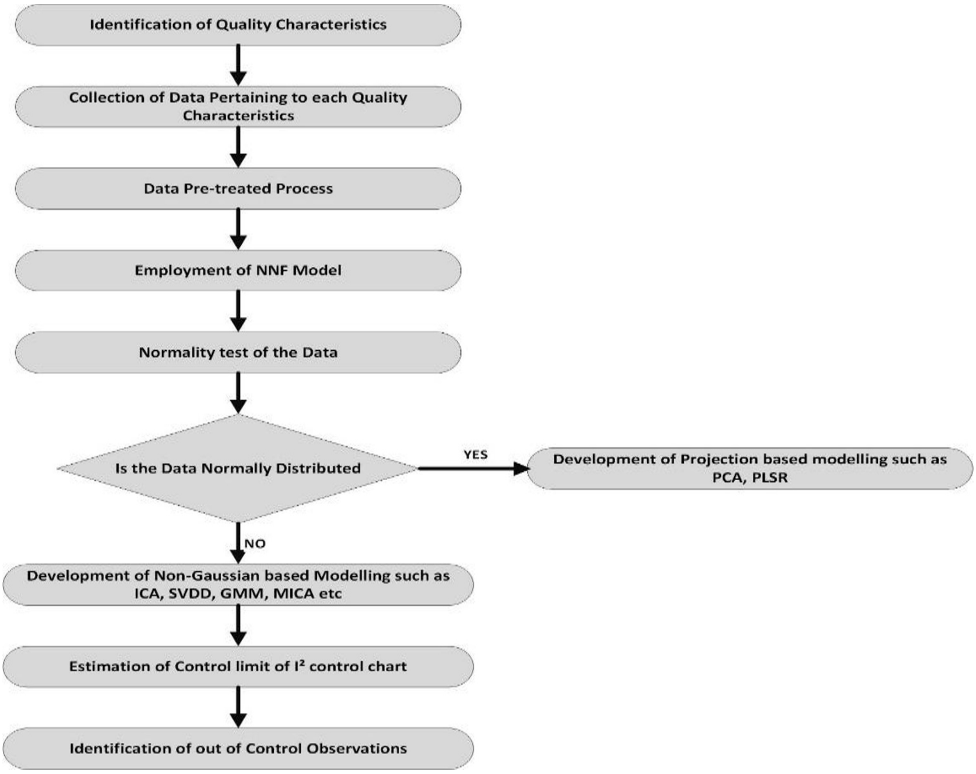

Fig. 4. Process flow diagram of NNF-ICA based process monitoring strategy.

5 Conclusion

In the current study, hybrid process monitoring comprises with NNF model and multivariate non-Gaussian statistical monitoring. The NNF model was used to transform nonlinear data forms into partially or fully linear data forms. To validate the proposed strategy, quality characteristics from the LRF stage of a SBMU were analyzed. The NNF model successfully achieved partial linearization of the quality characteristic observations, with the extent of transformation quantified by the R^2 value. Thereafter the linearly transformed data was used to extract ICs scores by ICA decomposition technique. The scores of ICs were used for the construction of an I^2 control chart via Bootstrap procedure. The newly plotted observations on the constructed I^2 chart revealed an out-of-control point, indicating the need for further diagnosis by process engineers. This demonstrates that the proposed strategy effectively detects abnormal process behavior in cases where the monitored characteristics exhibit nonlinear and non-normal relationships.

References

1. Z. Jalilibal, M.H.A. Karavigh, A. Amiri, M.B.C. Khoo, Run rules schemes for statistical process monitoring: A literature review, *Qual. Technol. Quant. Manag.* **20**, 21–52 (2023)

2. A. Melo, M.M. Câmara, J.C. Pinto, Data-driven process monitoring and fault diagnosis: A comprehensive survey, *Processes* **12**, 251 (2024)

3. R. Asadi, A. Queguineur, O. Wiikinkoski, H. Mokhtarian, T. Aihkisalo, A. Revuelta, I.F. Ituarte, Process monitoring by deep neural networks in directed energy deposition, *Rob. Comput. Integr. Manuf.* **87**, 102710 (2024)

4. D. Zago, Statistical process monitoring: A general framework for statistical process monitoring in Julia, *J. Stat. Softw.* **113**, 1–45 (2025)

5. R.L. Mason, J.C. Young, Multivariate statistical process control, in: *International Encyclopedia of Statistical Science* (Springer, Berlin, 2025), pp. 1645–1647.

6. T.J. Fontalvo-Herrera, R. Herrera, J. Zambrano, Three-phase method to assess the logistics service using Six Sigma metrics and Hotelling's T^2 chart, *Int. J. Product. Qual. Manag.* **35**, 17–39 (2022)

7. E. Sari, R. Widyasari, Application of Hotelling T^2 control diagram for drinking water quality, *Mathline J. Mat.* **9**, 273–284 (2024)

8. K. Rashid, A comparison between the MEWMA and Mahalanobis distance control chart, *Kurdistan J. Appl. Res.* **6**, 94–104 (2021)

9. D.H. Kadir, R.O. Yahya, A.M. Khudhur, T.H. Ali, Data de-noise for multivariate T^2 and S-charts using multivariate wavelets, *Zanco J. Human Sci.* **28**, 343–360 (2024)

10. S. Rakuff, R. Mikail, F. Sutton, M. Rutkevicius, Multiple principal component statistics for bearing condition monitoring, in: *Proc. IEEE SDEMPED* (2025), pp. 1–6.

11. M. Greenacre, P.J.F. Groenen, T. Hastie, A. Iodice d'Enza, A. Markos, E. Tuzhilina, Principal component analysis, *Nat. Rev. Methods Primers* **2**, 100 (2022)

12. B.M.S. Hasan, A.M. Abdulazeez, A review of principal component analysis algorithm, *J. Soft Comput. Data Min.* **2**, 20–30 (2021).

13. X. Chang et al., Prediction of coal body structure using PCA and logging curves, *Nat. Resour. Res.* **34**, 1023–1044 (2025)

14. I.A. Sadiq, Performance of partial least squares regression under multicollinearity, in: *RSS Nigeria Annual Conf. Proc.* (2025), pp. 318–328.

15. N. Er-rahmany et al., QSAR model using PLS/RF for corrosion inhibitors, *Next Res.* **2**, 100148 (2025)

16. Y. Han et al., Fault monitoring using adaptive kernel PCA integrating grey relational analysis, *Process Saf. Environ. Prot.* **157**, 397–410 (2022)

17. B.K. Gwashavanhu, A.J. Oberholster, S.P. Heyns, PCA vs. KPCA for turbine blade analysis, *Eng. Struct.* **318**, 118712 (2024)

18. Z. Wang, T. van der Laan, M. Usman, Self-adaptive quantum kernel PCA, *Adv. Sci.* **12**, 2570102 (2025)

19. M.T.H. Kaib et al., Data size reduction for nonlinear process monitoring using KPCA, *Expert Syst. Appl.* **274**, 126975 (2025)

20. B.K. Mishra, A. Das, A non-linear SP monitoring strategy: A case study, *Mater. Today Proc.* **41**, 1107–1109 (2021)

21. K.R. Storrs et al., Diverse deep neural networks predict human IT cortex, *J. Cogn. Neurosci.* **33**, 2044–2064 (2021)

22. C. Shi et al., ANN-based shelf life prediction: A review, *Crit. Rev. Food Sci. Nutr.* **64**, 12009–12024 (2024)

23. S. Bhatnagar et al., Chemical reaction neural networks for calorimetry data, *J. Power Sources* **628**, 235834 (2025)

24. M. Qureshi et al., Forecasting cardiovascular mortality using ANN, *BMC Public Health* **25**, 34 (2025)

25. A. Tharwat, Independent component analysis: An introduction, *Appl. Comput. Inform.* **17**, 222–249 (2021)

26. A. Hyvärinen, I. Khemakhem, H. Morioka, Nonlinear ICA for disentanglement, *Patterns* **4**, (2023)

27. E. Cerrillo-Cuenca, M. Sepúlveda, Z. Guerrero-Bueno, ICA for rock painting analysis, *J. Archaeol. Sci.* **125**, 105269 (2021)

28. G.T. Alencar, R.C. dos Santos, A. Neves, Fault recognition in transmission systems using ICA and CNN, *Electr. Power Syst. Res.* **229**, 110105 (2024)

29. F. Ricci, L. Bardone, S. Goldt, Feature learning from non-Gaussian inputs via ICA, *arXiv preprint*, arXiv:2503.23896 (2025)

30. B.K. Mishra, A. Das, A stage-wise non-Gaussian SPC strategy, *Eng. Res. Express* **5**, 015069 (2023)

31. K. Li et al., Deep SVDD for hyperspectral anomaly detection, *IEEE Trans. Geosci. Remote Sens.* **60**, 1–16 (2022)

32. Q. Wu, W. Lu, X. Yan, Process monitoring of nonlinear uncertain systems using stacked autoencoder–SVDD, *Appl. Soft Comput.* **129**, 109570 (2022)

33. A.Beck, O. Noam . A note on the support vector data description problem. Oper Res Lett. ,107357.(2025)

34. S.P. Fatemi et al., Distributed data storage using decision trees and SVM for IoT, *Sustain. Comput. Inform. Syst.* **46**, 101134 (2025)

35. M. Chen, S.J. Ludtke, Deep learning-based Gaussian mixture model for cryo-EM, *Nat. Methods* **18**, 930–936 (2021)

36. K. Rezaee et al., Deep learning-based real-time crowd anomaly detection, *Pers. Ubiquit. Comput.* **28**, 135–151 (2024)

37. A. Entezami et al., Early warning of structural damage using manifold learning, *Mech. Syst. Signal Process.* **224**, 111984 (2025)

38. Y. Li et al., Unsupervised embedding for process monitoring in repetitive systems, *IISE Trans.* **57**, 724–739 (2025)