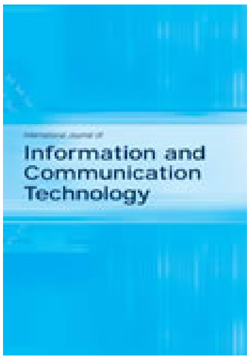


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Early warning system of cigarette process quality combined with intelligent sensing technology

Rongya Zhang, Mingchang Liu, Yaping Ma, Tingjie Bao, Peng Kuang, Wenbin Feng, Jiaxi Ni, Lingying He, Bin Yang, Chao Yang, Wu Wen, Ruifang Gu

DOI: 10.1504/IJICT.2025.10071852

Article History:

Received:	19 February 2025
Last revised:	08 May 2025
Accepted:	11 May 2025
Published online:	10 July 2025

Early warning system of cigarette process quality combined with intelligent sensing technology

Rongya Zhang, Mingchang Liu,
Yaping Ma and Tingjie Bao

Technology Center,
China Tobacco Sichuan Industrial Co., Ltd.,
Section 2, Chenglong Avenue, Longquanyi District,
Chengdu, Sichuan 610100, China
Email: rongyazhang@126.com
Email: liumc@sctobacco.com
Email: mayp@sctobacco.com
Email: baotj@sctobacco.com

Peng Kuang and Wenbin Feng

Shifang Cigarette Factory,
China Tobacco Sichuan Industrial Co., Ltd.,
No. 158, South Section of Yinghuashan Road,
Shifang, Sichuan 618400, China
Email: 18090793553@163.com
Email: 15884269507@163.com

Jiaxi Ni and Lingying He

Technology Center,
China Tobacco Sichuan Industrial Co., Ltd.,
Section 2, Chenglong Avenue, Longquanyi District,
Chengdu, Sichuan 610100, China
Email: 23008597@qq.com
Email: 137366167@gq.com

Bin Yang and Chao Yang

Shifang Cigarette Factory,
China Tobacco Sichuan Industrial Co., Ltd.,
No. 158, South Section of Yinghuashan Road,
Shifang, Sichuan 618400, China
Email: yangbin03@sctobacco.com
Email: 984303660@qq.com

Wu Wen and Ruifang Gu*

Technology Center,
China Tobacco Sichuan Industrial Co., Ltd.,
Section 2, Chenglong Avenue, Longquanyi District,
Chengdu, Sichuan 610100, China
Email: wenwu@sctobacco.com
Email: 15620452928@163.com
*Corresponding author

Abstract: In this paper, an early warning system of cigarette process quality combined with intelligent sensing technology is proposed to improve the quality of cigarette process production. A PCA multi-block modelling algorithm based on autoencoder feature extraction is proposed to extract autoencoder features from each sub-block, and the statistics of all sub-blocks are fused by Bayesian inference to make the monitoring results more intuitive. Compared with the traditional PCA and AE-PCA detection methods, the AE-MPCA algorithm proposed in this paper improves the abnormality detection accuracy of the drum leaf drying production process, and realises the accurate alarm of quality abnormalities, thus providing technical support for the early warning of subsequent cigarette process quality. In the subsequent process of cigarette process quality control, the application scope of intelligent sensing technology can be further improved to promote the effect of cigarette process quality control.

Keywords: intelligent perception; cigarettes; process quality; early warning.

Reference to this paper should be made as follows: Zhang, R., Liu, M., Ma, Y., Bao, T., Kuang, P., Feng, W., Ni, J., He, L., Yang, B., Yang, C., Wen, W. and Gu, R. (2025) 'Early warning system of cigarette process quality combined with intelligent sensing technology', *Int. J. Information and Communication Technology*, Vol. 26, No. 23, pp.1–23.

Biographical notes: Rongya Zhang received her PhD degree in Chemical Engineering from Tianjin University in 2019. She is currently a Senior Engineer with China Tobacco Si Chuan Industrial Company Ltd. Her research interests include cigarette technology, knowledge graphs, and digital transformation of cigarette technology.

Mingchang Liu received his Master's degree in Tobacco Technology from Zhengzhou Tobacco Research Institute in 2012, and is now a Senior Engineer of Sichuan China Tobacco Industry Co., Ltd. His research interests are tobacco technology.

Yaping Ma received her master's degree in tobacco technology from Zhengzhou Tobacco Research Institute in 2018, and is now an engineer of Sichuan China Tobacco Industry Co., Ltd. Her research interests are tobacco technology.

Tingjie Bao received his Bachelor's degree in Automation from Chengdu Institute of Engineering Information Technology in 2010. He is currently an Assistant Engineer at Sichuan Tobacco Industry Co., Ltd., with a research focus on cigarette technology.

Peng Kuang obtained his Bachelor's degree in Business Administration from Sichuan Agricultural University. He is currently an Assistant Engineer at the Shifang Cigarette Factory of Sichuan Tobacco Industry Co., Ltd. His research interests include cigarette manufacturing techniques, among others.

Wenbin Feng obtained his Bachelor's degree in Electrical Engineering and Automation from Sichuan University of Technology in 2014. He is currently an Assistant Engineer at the Shifang Cigarette Factory of Sichuan Tobacco Industry Co., Ltd. His research interests include cigarette manufacturing techniques, among others.

Jiaxi Ni graduated from Zhejiang University of Science and Technology in Hangzhou in 2019 with a Bachelor's degree in Packaging Engineering. She currently serves as an Analytical Tester in the Quality Supervision Department of the Technical Centre at China Tobacco Sichuan Industrial Co., Ltd., holding the position of Assistant Engineer. Her research focuses on cigarette product quality testing.

Lingying He graduated from Zhengzhou University of Light Industry in 1998 with a major in Tobacco Engineering. She is currently an engineer at China Tobacco Sichuan Industrial Co., Ltd., specialising in research on cigarette packaging technology.

Bin Yang obtained his Bachelor's degree in Automation from Southwest University of Science and Technology in 2003. He is currently an Engineer at the Shifang Cigarette Factory of Sichuan Tobacco Industry Co., Ltd. His research interests include cigarette manufacturing techniques, among others.

Chao Yang graduated from Chengdu University in 2015 with a degree in Automation. Currently working as an Assistant Engineer, his main research focuses on quality assurance in cigarette production processes and the control of production consumption during cigarette manufacturing.

Wu Wen received a Bachelor's degree in Tobacco Science and Engineering from the University of Science and Technology of China in 2007 and is now a Senior Engineer of Sichuan China Tobacco Industry Co., Ltd. His research interests are tobacco technology.

Ruifang Gu graduated from Shanghai University in 2006 with a Master's degree in Food Science. She is currently an Engineer at China Tobacco Sichuan Industrial Co., Ltd., specialising in research areas such as cigarette processing technology, cut tobacco and packaging process control, and quality supervision.

1 Introduction

In the tobacco manufacturing industry, from tobacco leaves to cut tobacco, cigarette silk making has formed a relatively perfect technological process. However, the traditional cigarette silk-making process requires high manpower, and there are many external factors affecting tobacco quality, which leads to the difficulty of tobacco quality control. The use of intelligent control can not only improve the accuracy of feeding, temperature control, material ratio and other work, but also effectively guarantee process quality through intelligent supervision. Therefore, it is very important to strengthen the research of intelligent control to promote the development of cigarette silk making technology.

In the process of cut tobacco manufacturing, intelligent control can monitor the working conditions of different processes in real-time during the processing, such as colour humidity, cutting effect, etc., and detect the problems existing in cut tobacco manufacturing in time, so as to reasonably adjust the related work. Under intelligent control, for cut tobacco with different quality requirements, intelligent control can combine preset values to make different evaluations and feedbacks on the quality of cut tobacco in different production lines, so as to ensure that the problems existing in the quality of cut tobacco in different production lines can be found in time. Compared with the traditional manual inspection, the inspection effect and inspection efficiency of the inspection work under intelligent control are better (Chong, 2023).

On the premise of applying intelligent control, it is necessary to upgrade and improve the cigarette silk making system in an all-round way, build an information base and install intelligent control device. These will have a direct impact on the silk making process and further improve the level of silk making process. In recent years, some tobacco processing and manufacturing enterprises have built intelligent control systems, used modern technology to replace traditional technology, and used intelligent technology to discover the problems existing in traditional technology and improve the technology, thus comprehensively promoting the upgrading of silk making technology. With the continuous upgrading and development of intelligent control technology, control technology will further promote the modernisation of tobacco quality control (Du et al., 2023).

At present, intelligent control has been applied to the facilities and equipment of cigarette silk making. Through intelligent control of facilities and equipment, it can effectively ensure that operators know the operation of the equipment at the first time and realise automatic adjustment and control. Intelligent control of facilities and equipment is the prerequisite for achieving precise control. In the intelligent control of cigarette silk making field equipment, it is necessary to prepare the hardware facilities required for intelligent control, including on-site distributed control box, sensor detection mechanism equipment and power distribution circuit cabinet, etc., and install supporting intelligent control software. At the same time, it is necessary to use bus network or industrial Ethernet for information transmission, install intelligent control devices on the equipment, including induction devices and control devices, etc., and ensure that the intelligent control devices establish contact with the main system (Gang et al., 2021).

Intelligent control can not only improve various operational processes, but also play an important role in on-site quality supervision and management, and has been applied in tobacco enterprises. The intelligent control of supervision and management includes discovering the problems existing in each link of silk making by combining each induction device, and comprehensively analysing the problems found in the recent silk making work by using data. In the intelligent control of supervision and management, relevant preparations need to be made, including setting up monitoring servers and monitoring computers, Ethernet network equipment and application servers. The software platform is the basic supervision and management intelligent control. The supervision and management system adopts the client-server mode, which can establish the connection with the database of production and processing system, so as to realise the collection and analysis of production information and material scheduling information.

2 Related works

2.1 Laser scanning technology

The laser scanning volume calculation includes a hardware system and a software system. The hardware system is mainly a laser scanner for obtaining point cloud data on the object surface and a computer for data processing. The software system includes the combination of point cloud filter data pre-processing, 3D reconstruction and volume calculation algorithm. 'Clean and effective' point cloud data is the guarantee of the accuracy of 3D reconstruction and volume calculation. Using 3D laser scanning to obtain point cloud data has the characteristics of fast acquisition speed and high sampling frequency, but there are also some problems such as high redundancy, missing information and noise of point cloud data. This can affect the accuracy and quality of data, reduce the reliability of modelling or analysis, and thus affect subsequent analysis and application (Gong et al., 2023). The point cloud obtained by the sensor will inevitably be polluted by noise, including the inherent noise of the acquisition equipment, the reflection properties of the surface of the measured object, and special outliers. Therefore, it is necessary to perform a 'cleaning' operation on the original point cloud, that is, filter the point cloud, remove noise, outliers, and smooth the point cloud to improve the data quality to obtain an accurate point cloud suitable for further processing. Domain-based filtering techniques are the most common methods, which exploit the similarity of a point with its neighbourhood to determine the filtering position (Guo and Hu, 2022). Hu et al. (2024) designed a method based on domain connectivity to judge glitch noise by laser scanning building point cloud, and then used clustering method to further remove smaller dense noise. This shows that when selecting filtering methods, factors such as data characteristics, noise types, application requirements, etc. need to be considered. Usually, better results can be achieved by using multiple filtering methods comprehensively or adjusting parameters. Since sensors can only scan within their limited field of view, point cloud data may come from different sensors, devices, or points in time when scanning large targets. The main purpose of registering these datasets into the same coordinate system is to achieve spatial consistency between point cloud data collected from different sources or at different times, so that they can be compared, fused, analysed or subsequently applied in the same coordinate system. Most existing registration methods alternately perform the two processes of correspondence search and transformation estimation until the set projection error is minimum (Liu, 2023). At present, common registration methods include iterative nearest point (ICP), feature matching, feature descriptor, transformation model, etc. Liu et al. (2023) used a laser scanner to obtain the three-dimensional point cloud information of the workpiece and the fixture, and used the iterative nearest point algorithm based on particle swarm search to register the coordinates with high accuracy, so as to establish the three-dimensional spatial coordinates of the workpiece and the machine tool.

Point cloud volume calculation is to estimate the volume of an object through point cloud data. The basic idea of point cloud volume calculation is to map the points in point cloud data to three-dimensional space, and then estimate the volume on this basis. At present, many volume calculation methods have been studied, including slice method, projection method, mesh method and convex hull method (Mu et al., 2022). Different calculation methods will have different applicable scenarios and accuracy, but no matter what kind of volume calculation method, its core is based on the idea of integration.

2.2 Stereo vision technology

In addition to laser detection technology, the application of stereo vision technology as a non-contact measurement method in the field of material volume measurement is an important development direction in the field of automation and intelligent measurement in recent years. Stereoscopic vision, also known as binocular vision, refers to the use of two or more cameras to shoot the same scene from different angles. By simulating the human binocular vision mechanism, it can recover depth information and calculate three-dimensional information of objects. The core challenge is to determine the best method to map the difference between images to the difference of the detection environment (Rui et al., 2023).

The core of stereo vision measurement system is stereo matching algorithm, which calculates the position of these points in three-dimensional space by analysing the corresponding points in the pictures taken by two cameras. New methods and technologies to solve this problem are developing every year, and there is a trend of improvement in accuracy and time consumption. This process involves complex geometric transformations and optical principles, and requires accurate camera calibration to ensure the accuracy of measurement results (Wang et al., 2022a). Stereo vision systems usually include hardware components such as cameras, lenses, lighting equipment, computer processing units, and software algorithms such as camera calibration, image processing, stereo matching, and 3D reconstruction. In the application of material volume measurement, stereo vision technology can handle various objects with complex shapes and irregular surfaces, and has the advantages of low cost, simple operation and easy popularisation (Wang and Zhang, 2023). Wang et al. (2022b) used binocular vision technology to obtain the image of the material and obtain the three-dimensional coordinate information of the belt conveyor under no-load and full-load scenes, and proposed a method to improve the accuracy of material volume calculation based on pixel coordinates. The test results show that the improved material volume calculation accuracy reaches more than 95%. Wang et al. (2024) proposed an accurate measurement method of droplet volume based on stereo vision, which obtains high-resolution images through stereo vision, and proposed an accurate binocular droplet image segmentation algorithm to segment images with different ambiguity. According to the contour obtained by binocular droplet image segmentation algorithm, the droplet is reconstructed and the droplet volume is calculated. The experimental results show that the measurement accuracy of this algorithm is ± 3%. Wang et al. (2021) proposed a method to detect the volume of mouse feed based on binocular stereo vision. The coordinates of these dense points are formed into a point cloud, and then the volume of the point cloud is calculated by projection method, and finally the volume of mouse feed is obtained. Moreover, the experimental validation is performed using the stereo matching dataset provided by the Middlebury evaluation platform. The results show that the average error between the calculated volume and the actual volume is 7.12%. In stereo vision measurement system, its core is stereo matching algorithm between images, which is also a research hotspot in recent years. However, the application of stereo vision technology in material volume measurement also faces some challenges. For example, factors such as changes in lighting conditions, differences in camera viewing angles, and reflections from object surfaces can affect the accuracy of measurements. Therefore, researchers are working hard to continuously optimise the stereo matching algorithm and improve the robustness and adaptability of the system (Wei et al., 2024).

3 Monitoring method of cigarette production process based on AE and PCA

The production process of leaf silk has strong coupling characteristics among multi-variables. Because univariate statistical process monitoring does not consider the correlation between process variables, and the data are diversified, univariate statistical methods are increasingly unable to meet the needs of process monitoring.

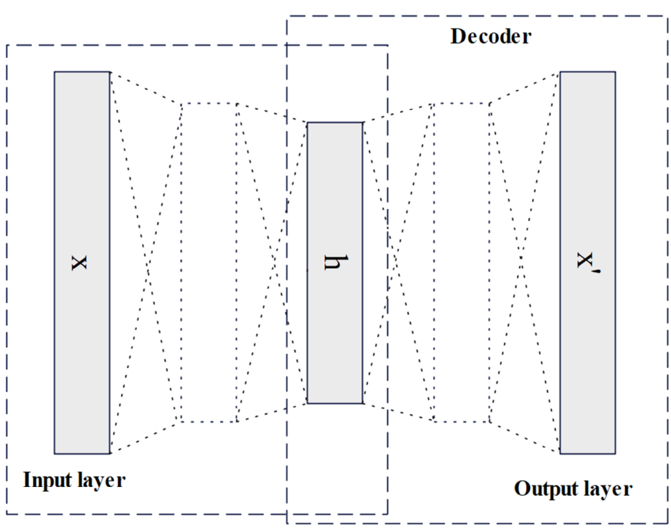
The multi-block modelling strategy solves the problem of inaccurate monitoring of global modelling in complex industrial processes, and also takes into account the huge advantages of deep learning in data nonlinear feature extraction.

3.1 Algorithm principle

Feature extraction is a key step in process monitoring. In fact, cigarette production process data is nonlinear and non-Gaussian. Traditional PCA method can express the linear features of data well, but it does not have strong nonlinear expression ability.

Autoencoder is an unsupervised learning algorithm. This network has the advantage of nonlinear dimensionality reduction, which can extract deep nonlinear features of data and maximise the representation of original data information. As shown in Figure 1, the network includes a three-layer network structure: an input layer, an intermediate hidden layer and an output layer. The network is completely data-driven, does not require prior knowledge, and can encode and decode data independently of labels. Its network output layer and input layer have the same size. A complete autoencoder contains an encoder, a decoder, and a loss function. The encoder compresses the input information x through the neural network to extract the important features h of the data, and then the decoder decompresses it to obtain the reconstructed data x'. The model can transmit it in reverse by minimising the reconstruction error of the input and output data, update the network parameters, gradually improve the accuracy of the model, and make the output as close to the input as possible to ensure better expression of the hidden layer.

Figure 1 AE network structure



The intermediate hidden layer obtained by the encoder learns the latent representation of the original data. The latent feature is the essence of the input data. For the input x, the encoding process is expressed as (Xu et al., 2023):

$$h = f\left(W_e x + b_e\right) \tag{1}$$

The decoding process of hidden layer features is expressed as:

$$x' = f\left(W_d x + b_d\right) \tag{2}$$

In the formula, f is the activation function, W_e and W_d are the weight matrices of the hidden layer and the output layer, respectively, and b_e and b_d are the deviations of the hidden layer and the output layer, respectively. AE iteratively optimises the parameter set $C = \{W_e, W_d, b_e, b_d\}$ through the back propagation algorithm until the reconstruction error is minimised, so that the output of the network is as close to the input as possible.

Commonly used activation functions are Sigmoid, Tanh, ReLU and their variants, and the Sigmoid and PReLU activation functions are used.

Principal component analysis is the most commonly used statistical method to solve multivariate problems. From the perspective of dimensionality reduction, it replaces the original variables with a new set of variables. The new data is the combination of the original data and contains the maximum amount of information of the original data. From a mathematical point of view, PCA projects data into principal component space and residual space, removes redundant information, and extracts irrelevant latent variables through orthogonal linear transformation to represent data changes. When monitoring modelling, for a standardised data matrix $X \in R^{m\times n}$, n is the number of samples and m is the number of variables. The matrix X can be decomposed by PCA into the following form (Xu et al., 2022):

$$X = TP^T + E \tag{3}$$

$$T = XP \tag{4}$$

$$E = X\left(I - PP^T\right) \tag{5}$$

In the formula, T is the principal component score matrix, P is the load matrix, and E is the residual matrix. The original data space is decomposed into principal component space and residual space, and the number k of principal components that can represent 85% of the information of the original data is retained in this paper.

For a new sample $x \in R^m$, the $1 \times k$-dimensional principal component score vector t, the estimated value $\tilde{x}$ and the residual e of x can be calculated.

$$t = P^T x \tag{6}$$

$$\tilde{x} = Pt = PP^T x \tag{7}$$

$$e = x - \tilde{x} = \left(I - PP^T\right)x \tag{8}$$

Using the above information for statistical hypothesis testing, including the hotelling-T^2 statistics in the principal component space and the Q statistics in the residual space, it can be determined whether there are abnormal working conditions in the production process.

The T^2 and Q statistic is calculated as follows:

$$T^2 = tS^{-1}t^T = xPS^{-1}P^T x^T \tag{9}$$

$$Q = ee^T = x\left(I - PP^T\right)x^T \tag{10}$$

In the formula, t is the $(1 \times K)$-dimensional principal component score vector, the diagonal matrix $S = diag(\lambda_1, \lambda_2, \ldots, \lambda_R)$ is composed of the first k eigenvalues in the covariance matrix of the modelling dataset X, e is the residual vector, and the T^2 and Q control limits are usually obtained by F distribution and weighted χ^2 distribution, respectively (Zhong et al., 2023):

$$T_\partial^2 - \frac{k\left(n^2 - 1\right)}{n(n-k)} F_{k,n-k,\partial} \tag{11}$$

$$Q_\partial \sim g\chi^2_{h,\partial} \tag{12}$$

$$g = \frac{v}{2m} \tag{13}$$

$$h = \frac{2m^2}{v} \tag{14}$$

In the formula, $F_{k,n-k,\partial}$ represents the F distribution under the n–k condition of confidence level a and freedom level k, m and v represent the mean and variance of the Q statistic of the modelled data respectively. In this paper, the confidence level is taken as 0.99. When both statistics T and Q are within the control limits T_∂^2 and Q_∂, it indicates that the production process is in a normal state, otherwise it is considered that an abnormal alarm has occurred in the process.

Anomaly detection rate (FDR) and false alarm rate (FAR) are defined as follows:

$$FDR = \frac{Effective\ number\ of\ alarms}{Total\ number\ of\ abnormal\ samples} \tag{15}$$

$$FAR = \frac{Number\ of\ false\ alarms}{Total\ number\ of\ normal\ samples} \tag{16}$$

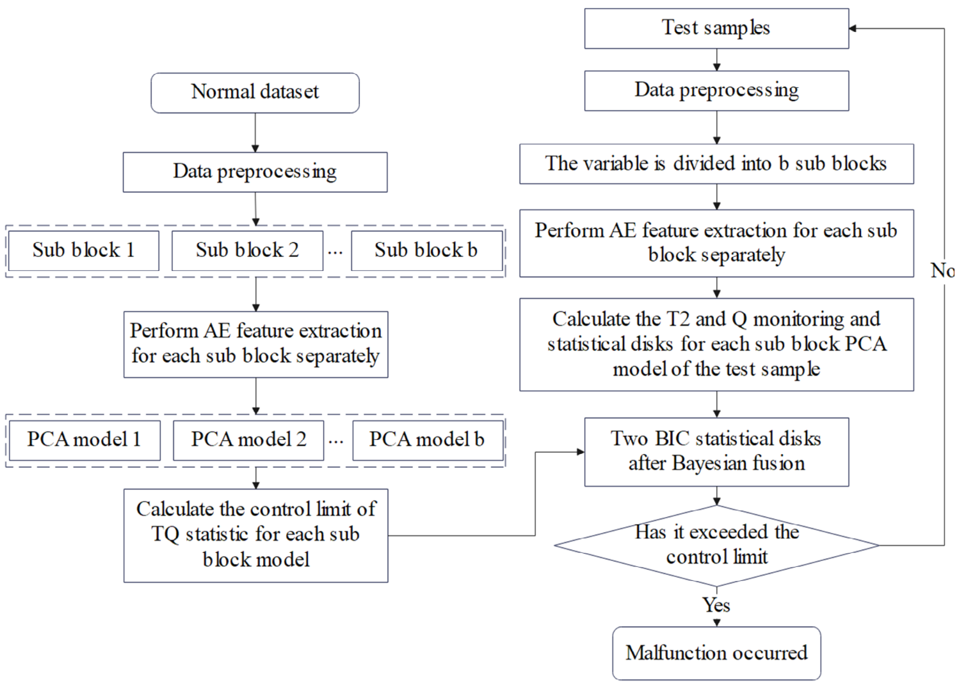
3.2 Process monitoring method based on AE-MPCA

A large amount of process data obtained online in actual industrial large-scale production has the characteristics of nonlinearity and strong coupling between variables. Global modelling can't extract local characteristics of the process, which leads to the reduction of detection accuracy of abnormal working conditions. In order to solve this problem, firstly, the variables are divided into b sub-blocks with appropriate rules, and multiple local models are established to improve the monitoring effect. Common blocking methods mainly include unsupervised clustering, mutual information, etc. However, these methods are based on data characteristics and do not consider the knowledge of process mechanism, which will lead to differences between the blocking results and the actual process. Therefore, this paper starts from the process principle of drum drying machine and implements a reasonable blocking strategy for the process in combination with the actual process.

Under the above multi-block modelling framework, the flow of AE-MPCA process monitoring algorithm is shown in Figure 2.

Firstly, the hidden features of each sub-block are extracted by autoencoder, and the hidden features are used as new observation variables to establish a sub-PCA model. Finally, the monitoring results of multiple subspaces are fused by Bayesian inference (BI) to obtain the overall monitoring index (Zhu et al., 2021).

Figure 2 Flow chart of AE-MPCA process monitoring



The specific steps are as follows:

3.2.1 *Offline modelling*

1 A normal dataset X after standardisation.

2 According to the actual process principle of the drum drying machine, the process variables are divided into b sub-blocks from the angle of material and machine.

3 Each sub-block is trained with an autoencoder model, and the latent feature h of each sub-block are extracted according to equation (1), and b feature datasets $h_1, h_2, \ldots h_b$ are obtained. The feature data represents the original data information to the maximum extent.

4 $h_1, h_2, \ldots h_b$ is used as the new dataset for each sub-block. According to equation (3), the PCA model of each sub-block can be expanded as:

$$X_i = T_i P_i^T + E_i, i \in [1, b] \tag{17}$$

According to equation (11) and equation (12), the control limits $T_{i,lim}^2$ and $Q_{i,ion}$ of each sub-PCA model statistic are calculated.

3.2.2 Online monitoring

1. For new test samples, a standardised dataset X is obtained through data pre-processing, and the process variables are divided into b sub-blocks using the same block strategy
2. Based on the trained autoencoder model, the hidden features h of each sub-block are extracted. The hidden features of each block are used as a new observation dataset to establish b sub-PCA models, and the monitoring statistics T^2 and Q of the first PCA model are calculated:

$$T_i^2 = x_i P_i S_i^{-1} P_i^T x_i^T \tag{18}$$

$$Q_i = x_i \left(I - P_i P_i^T \right) x_i^T \tag{19}$$

3. Each sub-block of multi-block modelling has monitoring results, and for more sub-blocks, an overall monitoring index cannot be obtained intuitively, which is not conducive to the final decision of abnormal working conditions. Moreover, the statistics and control limits of each sub-block are different, so it is difficult to directly fuse these results. BI strategy can construct the overall monitoring index. Taking T^2 statistics as an example, the failure probability of a sample X in the first sub-block can be expressed as follows:

$$P_{T^2}\left(F \middle| X_{test,i}\right) = \frac{P_{T^2}\left(X_{test,i} \middle| F\right) P_{T^2} * (F)}{P_{T^2}\left(X_{test,i}\right)} \tag{20}$$

$$P_{T^2}\left(X_{test,i}\right) = P_{T^2}\left(X_{test,i} \middle| N\right) P_{T^2}(N) + P_{T^2}\left(X_{test,i} \middle| F\right) P_{T^2}(F) \tag{21}$$

The specific expression of conditional probability $P_{T^2}\left(X_{test,i} \middle| N\right)$ and $P_{T^2}\left(X_{test,i} \middle| N\right)$ is as follows:

$$P_{T^2}\left(X_{test,i} \middle| N\right) = e^{-T_i^2 / T_{i,\lim}^2} \tag{22}$$

$$P_{T^2}\left(X_{test,i} \middle| F\right) = e^{-T_{i,\lim}^2 / T_i^2} \tag{23}$$

In the formula, $x_{test,i}$ represents the test sample in the 1st sub-block, T_i^2 represents the T^2 statistic of the i^{th} sub-block, $T_{i,lim}^2$ represents the T^2 control limit of the i^{th} sub-block, N and F represent normal and abnormal situations, $P_{T^2}(N)$ is the prior probability of a normal sample, which represents the confidence level α, $P_{T^2}(F)$ represents the confidence degree $1-\alpha$, which is the prior probability of an abnormal sample. Then, the detection results of all sub-blocks are fused, and the obtained BIC_{T^2} statistic is the global monitoring result.

$$BIC_{T^2} = \sum_{i=l}^{m} \left\{ \frac{P_{T^2}\left(X_{test,i} \middle| F\right) P_{T^2}\left(F \middle| X_{test,i}\right)}{\sum_{i=1}^{m} P_{T^2}\left(X_{test,i} \middle| F\right)} \right\} \tag{24}$$

Similarly, when the Q statistic is Bayesian fused, the failure probability of sample X_{test} in the first sub-block can be expressed as follows:

$$P_Q\left(F \middle| X_{test,i}\right) = \frac{P_Q\left(X_{test,i} \middle| F\right) P_Q(F)}{P_Q\left(X_{test,i}\right)} \tag{25}$$

$$P_Q\left(X_{test,i}\right) = P_Q\left(X_{test,i} \middle| N\right) P_Q(N) + P_Q\left(X_{test,i} \middle| F\right) P_Q(F) \tag{26}$$

The specific expression conditional probability $P_Q(X_{test,i}|N)$ and $P_Q(X_{test,i}|F)$ is as follows:

$$P_Q\left(X_{test,i} \middle| N\right) = e^{-Q_i/Q_{i,lim}} \tag{27}$$

$$P_Q\left(X_{test,i} F\right) = e^{-Q_{i,lim}/Q_t} \tag{28}$$

In the formula, $x_{test,i}$ represents the test sample in the i^{th} sub-block, Q_l represents the Q statistic of the i^{th} sub-block, $Q_{i,lim}$ represents the Q control limit of the i^{th} sub-block, N and F represent normal and abnormal situations, $P_Q(N)$ is the prior probability of normal samples, which represents the confidence level α, $P_Q(F)$ represents the confidence degree $1-\alpha$, which is the prior probability of abnormal samples. Then, the detection results of all sub-blocks are fused, and the obtained BICQ statistics are the global monitoring results.

$$BIC_Q = \sum_{i=l}^{b} \left\{ \frac{P_Q\left(X_{test,i} \middle| F\right) P_Q\left(F \middle| X_{test,i}\right)}{\sum_{i=1}^{b} P_Q\left(X_{test,i} \middle| F\right)} \right\} \tag{29}$$

The control limit of the fused statistics is $1-\alpha$, and when both BIC statistics are within the control limit, the process is considered to be in a normal state, otherwise it is considered that the process has an abnormal alarm.

4 Test study

4.1 Test methods

Intelligent control can not only improve various operational processes, but also play an important role in on-site quality supervision and management, and has been applied in tobacco enterprises. The intelligent control of supervision and management includes discovering the problems existing in each link of silk making by combining each induction device, and comprehensively analysing the problems found in the recent silk making work by using data. In the intelligent control of supervision and management, relevant preparations need to be made, including setting up monitoring servers, monitoring computers, Ethernet network equipment and application servers. Software platform is the foundation of supervision and management of intelligent control. The supervision and management system adopts the client-server mode, which can establish the connection with the database of production and processing system, so as to realise the collection and analysis of production information and material scheduling information.

Figure 3 Main process links of cigarette silk making

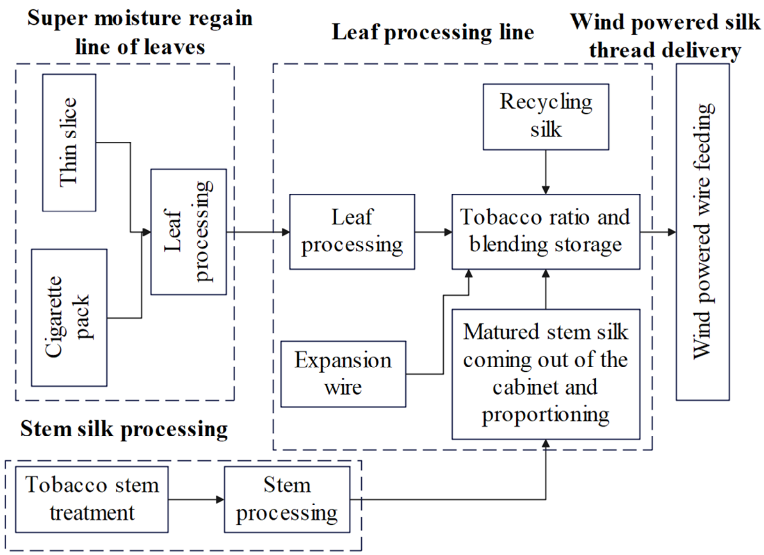
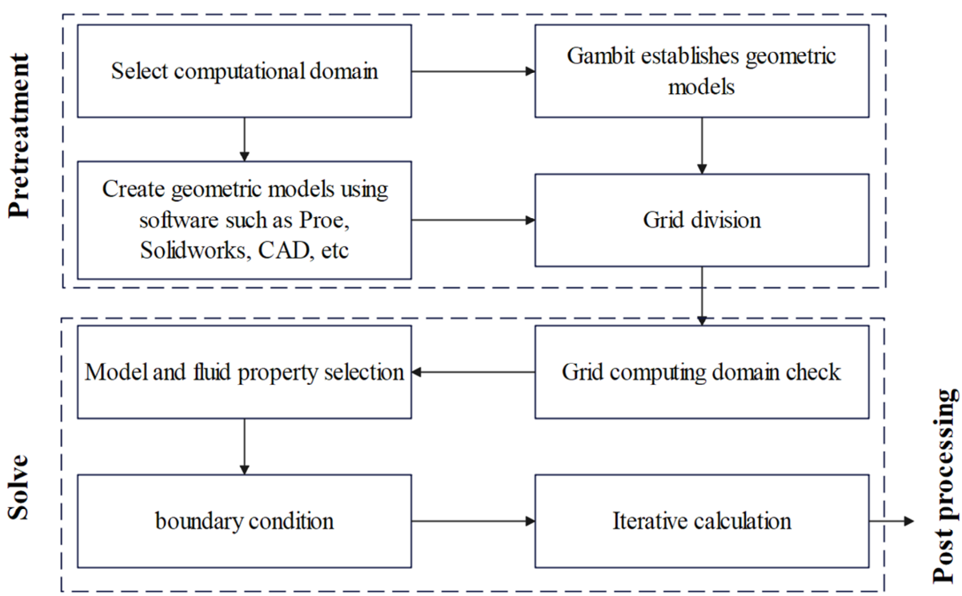


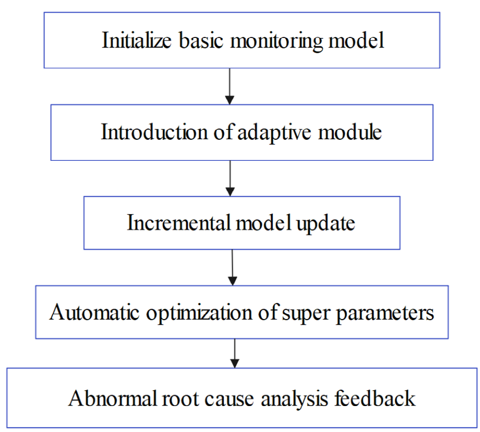
Figure 4 Simulation flow chart



Because the cigarette process is complicated, this paper takes the drying process of cigarette straight filament process as an example to analyse. Computer can effectively simulate various field distributions in drum dryer. At present, the common numerical simulation tool for drying materials in drum is computational fluid dynamics (CFD). The drying process of cut tobacco in drum is simulated by Fluent, and the distribution of temperature and humidity field in drum is analysed according to the simulation results. The whole process is generally divided into three parts (pre-processing, solution and post-processing), as shown in Figure 4, which is the flow chart of simulation calculation. Pre-processing mainly uses Solidworks, Proe and other software (or Gambit) to draw and model the actual device, and then meshes the built structural model. The solution is mainly divided into checking the grid computing domain, inputting relevant physical property data and selecting the calculation model, and iterative calculation after confirming the boundary conditions. Post-processing is mainly to process and analyse the data of the calculation results.

Because the air flow and cut tobacco mainly exchange heat and mass in the inner cylinder, this simulation only studies the inner cylinder flow field of the cylinder. In addition, due to the thin thickness of the copying plate, the quality of the mesh will be affected in the later mesh division. Proper thickening of the copying plate thickness will not obviously affect the flow field analysis. In order to improve the mesh quality, the copying plate and guide plate with a thickness of 2 mm are changed to 6 mm.

Figure 5 Architecture design of adaptive learning technology

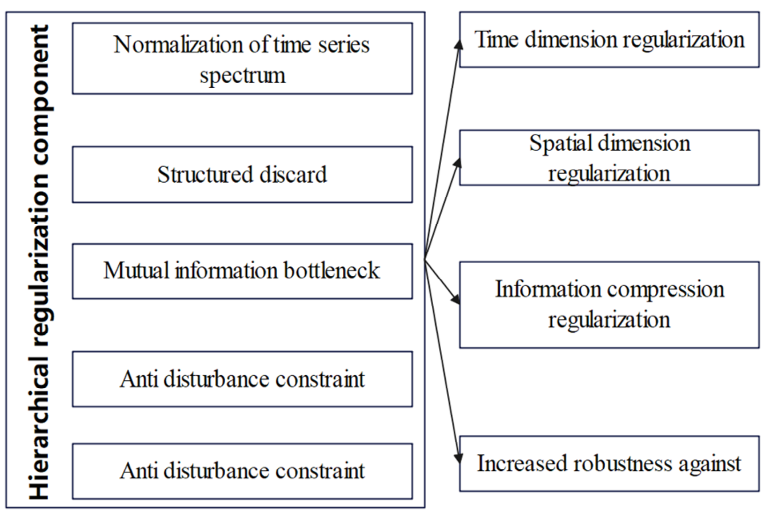


As shown in Figure 5, combined with the current cutting-edge technologies in the field of industrial process monitoring, the following adaptive learning technology solutions that can be embedded in the original monitoring system are designed, focusing on solving the problems of insufficient manual parameter adjustment and model generalisation ability:

The adaptive model is a super parameter adaptive structure based on meta learning (Figure 6). The meta learning controller is constructed by flexible control analysis through dynamic PCA dimension selection. By evaluating the contribution of each principal component to the reconstruction error, the reserved principal component score is automatically adjusted Establish a double-layer time window mechanism (short-term window: 50 batches, long-term window: 500 batches) to automatically balance the impact of new and old data By embedding an adaptive learning module, the model can maintain stable monitoring performance within the range of ± 15% of process parameter drift, while reducing the dependence on the experience of domain experts. In the actual deployment, the synergy effect of each module is verified through historical data simulation, and then the online debugging is gradually implemented.

The advanced regularisation technology for dynamic industrial data is introduced (Figure 6). The proposed regularisation technology system enhances the generalisation ability of industrial process monitoring model through multi-dimensional constraint mechanism. Based on the traditional spectral normalisation, the causal time series spectral normalisation introduces an exponential attenuation factor (γ = 0.9), and dynamically adjusts the Lipschitz constant of the weight matrix to maintain the temporal causality; structured random discarding improves the traditional dropout method and adopts block_size = 3 to enhance the decoupling ability of feature space and reduce information loss 48; the mutual information bottleneck regularisation is represented by the variational approximation constraint hidden layer, and the dynamic balance of feature retention and redundancy elimination (β = 0.1) is used to improve the robustness of the model to noise interference; The anti gradient matching technology forces the model to learn a smoother decision boundary by constraining the gradient consistency of normal samples and anti samples (λ = 0.05).

Figure 6 Regularisation technique

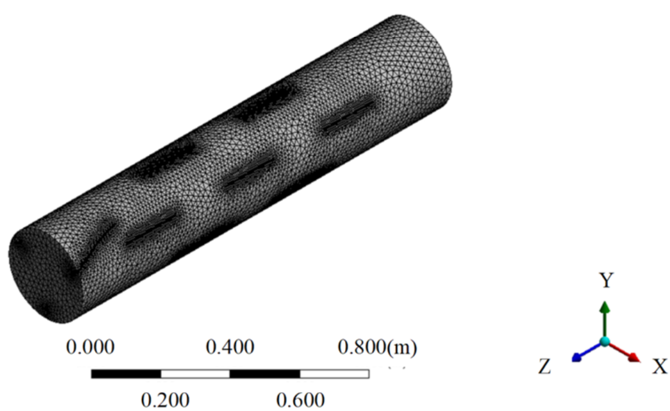


Based on the monitoring requirements of the drum drying process, an integrated learning mechanism is introduced into the ae-mpca framework to design a multi-dimensional robustness enhancement scheme. This integrated learning model creatively integrates heterogeneous self encoder and dynamic regularisation technology to build a multi-level robustness enhancement framework for the monitoring requirements of industrial processes. Through the parallel reasoning of heterogeneous model pool (five types of differentiated ae-mpca, such as uniform/non-uniform blocking, noise injection, timing expansion, etc.), combined with the gradient lifting tree dynamic weighted fusion module, the multi-dimensional feature complementarity and decision coordination are realised; The hierarchical regularisation system (causal time series normalisation, structural discarding, mutual information bottleneck and confrontation gradient matching) is introduced to optimise the constraint model from four dimensions: time dependence, feature decoupling, information compression and decision smoothing; Combined with embedded memory optimisation (block model loading alternately) and parameter adaptive mechanism (regularisation strength closed-loop control).

The introduction of interpretable artificial intelligence (Xai) technology into ae-mpca integrated monitoring framework significantly improves the transparency and reliability of monitoring results through multi-dimensional interpretation mechanism. The scheme dynamically quantifies the contribution weight of key process parameters (such as cylinder temperature and wind speed) to the reconstruction error by combining the shap value, generates local decision boundary comparison samples by using lime, and constructs a causal map to depict the abnormal propagation path between parameters, supporting the tracing of the abnormal root from the three levels of feature importance, individual sample and temporal correlation; At the same time, a lightweight real-time interpretation engine (delay ≤ 35 ms) is designed, which adapts to different user needs through two-level interpretation strategies (basic trend chart and expert level hidden layer analysis), and cooperates with confidence quantitative evaluation and three-dimensional traceability matrix (model version - data distribution – interpretation report).

The geometric model uses regular tetrahedron to divide the face and volume to obtain the mesh, and the divided mesh model is shown in Figure 7. The cell size is set to 4mm, and the number of meshes and cells after division are 1012123 and 1356231, respectively.

Figure 7 Mesh model of drum cylinder (see online version for colours)



This paper selects 20 batches of historical data of a tobacco company for simulation research. The sampling interval is 10s, and the time is from June 6, 2024 to July 20, 2024. Through data screening, this paper selects 11 normal batches of data for monitoring modelling, selects another 5 batches of normal operating data to add disturbances as test samples, and reselects 3 batches of normal operating data from May 24 to May 25, 2024 to add disturbances as test samples. After pre-processing the three groups of data, we get the modelling dataset (7,346 × 13) and test samples (2,614 × 13) and (2,316 × 13).

The test environment is as follows: CPU: i9-13000H; RAM: 32.00 GB; Python version 39, Pytorch-CPU version 1.11. 0, and PCA uses Matlab software platform.

Perturbations need to be added to the test set before data normalisation can be done. Three variables, namely, the opening of steam valve, the temperature of cylinder wall and the opening of moisture discharge damper, which have great influence on the moisture content of leaf silk are selected to add amplitude interference. Different variables have different dimensions and different fluctuation forms, which determines the amplitude of added interference. The specific form of disturbance is shown in Table 1:

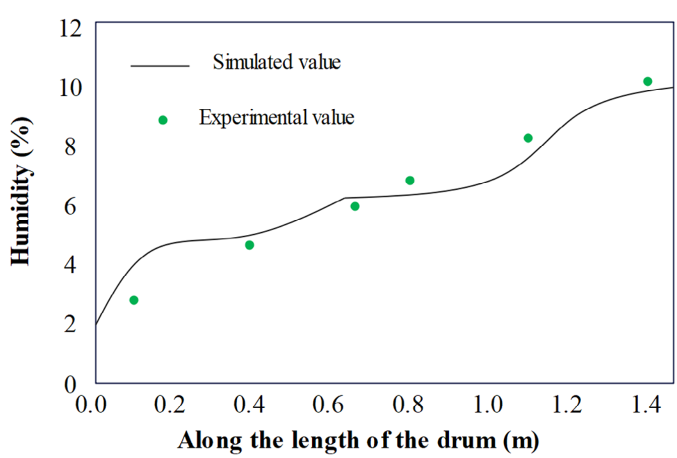
Table 1 Fault description

Fault no.	*Description*
Fault1	Step disturbance signal with amplitude of 2.0% is added to drum drying _ steam valve opening
Fault2	Ramp interference signal with amplitude of 0.008 is added to drum drying _ steam valve opening
Fault3	Step interference signal with amplitude of 0.2% is added to drum drying _ cylinder wall temperature
Fault4	Ramp interference signal with amplitude of 0.001 is added to drum drying _ cylinder wall temperature
Fault5	Step interference signal with an amplitude of 50% is added to the opening of the drum drying _ moisture discharge damper opening
Fault6	Ramp interference signal with an amplitude of 0.1 is added to the opening of the drum drying _ moisture discharge damper opening

4.2 Results

Iterative calculation is adopted. Firstly, the single-phase simulation is carried out under the condition of only hot air. After about 500 steps of calculation, the discrete phase is opened, and the convergence of 3,000 steps is calculated. The calculation results are processed and analysed. In order to verify the accuracy of the simulation model, the simulated humidity field and the experimental value are compared and analysed. As shown in Figure 8, the simulated humidity at the central axis of the drum and the humidity at five points at the central axis detected by the test under the same working condition are close to the overall trend of the simulated value and the experimental value, which proves that the simulation results are reliable.

Figure 8 Comparison of simulation values and experimental values (see online version for colours)



During the drying process of the drum leaf filament, after the leaf filament is expanded, heated and humidified online, the wet material enters the drum through the feed port, rolls forward under the rotation of the roller, and all the steam enters the drum copying plate through the rotary joint to heat the leaf filament. The other way enters the heat exchanger of the hot air system, and the hot air enters the drum from the front chamber and fully contacts the leaf filaments, so that the leaf filaments are evenly dried, and the whole process is circulated in an orderly manner. Therefore, the drying process is the result of the interaction between leaf silk and expansion drying equipment, and it can also be said that it is the mutual coupling of material and machine. In order to realise the refined modelling of the process, the whole leaf silk drying system is divided into sub-blocks of import and export materials and sub-blocks of expansion drying system, and the local information of the two links is well considered. The blocking results are shown in Table 2.

Table 2 Block results

Block1	*Block2*
Import and export material sub-block	Expansion drying system sub-block
Leaf filament expansion _ inlet material instantaneous flow rate	Leaf filament expansion _ steam flow rate
Leaf filament expansion _ inlet material moisture	Drum drying _ cylinder wall temperature
Leaf filament expansion _ outlet temperature	Drum drying _ steam pressure
Drum drying _ outlet material temp	Drum drying _ steam valve opening
Drum drying _ outlet moisture	Drum drying _ hot air steam flow
	Drum drying _ hot air temperature
	Drum drying _ moisture drain damper opening
	Drum drying _ moisture discharge negative pressure

Figure 9 Fluctuation of quality indicators (see online version for colours)

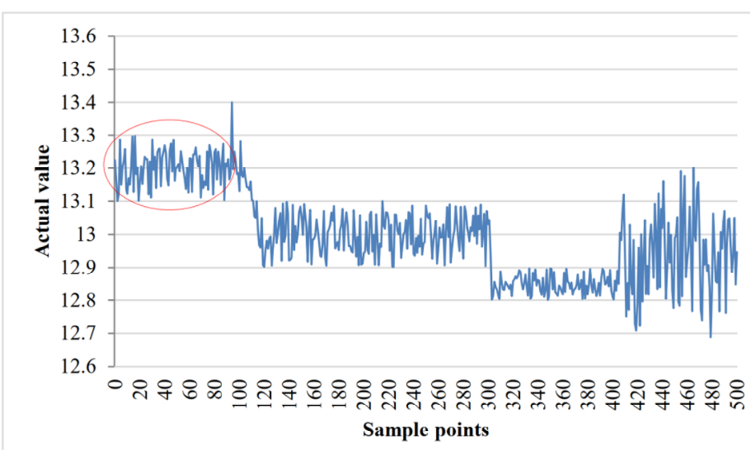
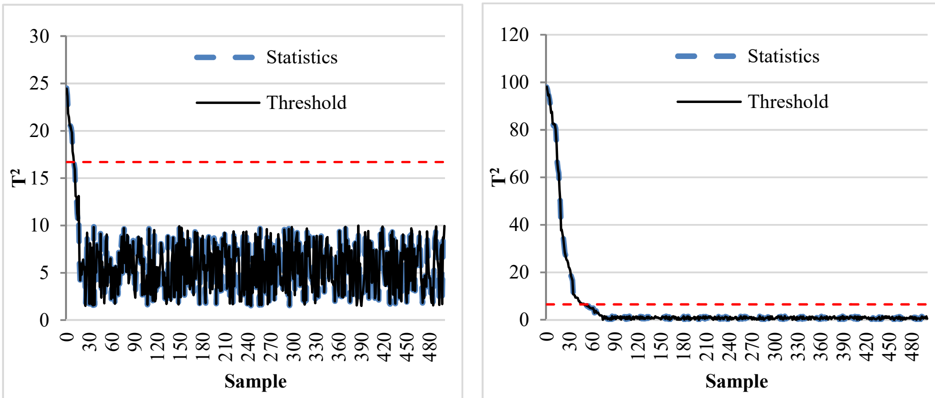
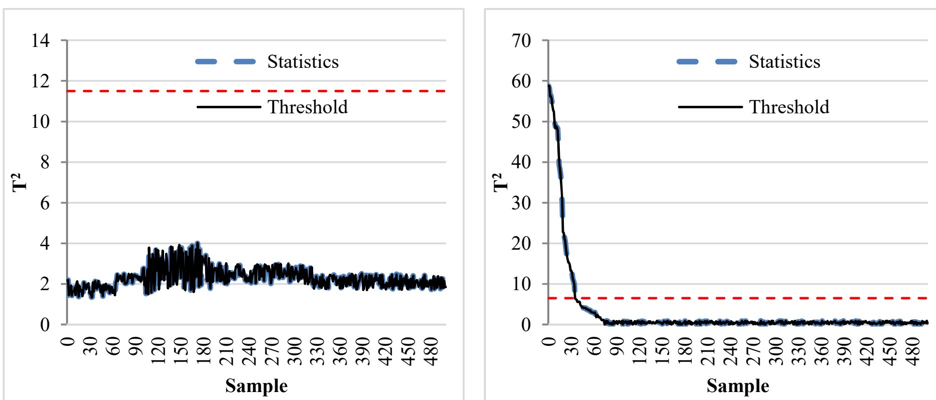


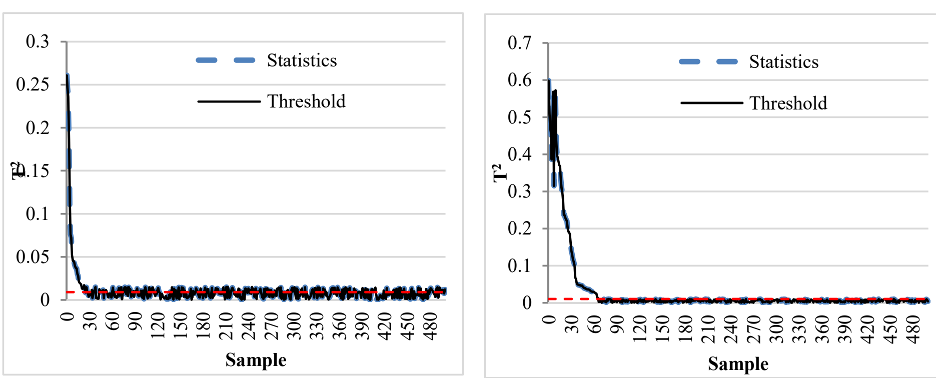
Figure 10 Process monitoring diagram of three methods (a) PCA (b) AE-PCA (c) AE-MPCA (see online version for colours)



(a)



(b)



(c)

In order to further verify the monitoring effect of AE-MPCA algorithm, three process monitoring models are used to monitor two actual abnormal working conditions, and the alarm rate (FDR) for each abnormal working condition is calculated and a visual monitoring diagram is drawn for comparative analysis. By querying the historical data of drying production of leaf silk in this enterprise, there is a case shown in Figure 9. There is moisture fluctuation in the production process, as shown by the curve in the red circle. The abnormal fluctuation is characterised by the fact that the state is different from other data distributions within a certain time step, which directly affects the quality of leaf silk. In order to improve the quality qualification rate of the whole leaf silk and ensure stable production, it is necessary to identify it for operators to judge it in time.

The process monitoring algorithm proposed in this paper is used to verify that the number of principal components of global PCA is 6, the number of principal components of AE-PCA is 3, and the number of principal components of two sub-blocks is 3 and 5 respectively. The specific experimental results are shown in Table 3.

Table 3 Comparison results of anomaly detection rate FDR (%) of the three methods

	Q	*T2*
PCA	32.18	2.39
AE-PCA	76.60	0.00
AE-MPCA	90.75	0.00

In order to better demonstrate the effectiveness of the AE-MPCA method compared to the comparative method, the monitoring charts of the three methods for two cases are listed in Figure 10. In the figure, the red dashed line represents the control limit of 0.99 confidence, and the solid line represents the and statistics of the test set. The blue and black solid lines are used to distinguish the abnormal sample points and the normal sample points of the test set. Both statistical indicators are within the control limit, indicating that the production process is in a normal state, otherwise it is considered that the process has an abnormal alarm.

4.3 Analysis and discussion

Figure 8 shows the cross-sectional humidity field with x = 0. It can be seen that the ambient humidity of the whole humidity field in the cylinder gradually increases in the horizontal direction (along the z-axis), and the humidity in the field gradually increases from 0 to 3.81% from the inlet end to about 0.2 m along the cylinder. There is a humidity block at the wall under the cylinder, which is quite different from the surrounding humidity. The reason may be that cut tobacco is deposited at the bottom of the cylinder. The colour block gradient of humidity in the field changes obviously between 0.2 m and 1.1 m along the horizontal direction of the drum (along the z axis), and the humidity in the field increases along the drum, and increases from top to bottom in the longitudinal direction (along the y axis), and the humidity near the lower side of the drum wall is higher. It can be seen that the environment in the cylinder is in the drying stage of increasing speed at about 0.4 m ~ 0.75 m, and the water on the surface of cut tobacco quickly volatilises to the environment in the cylinder. In addition, 0.75 m ~ 1.1 m is in the decreasing stage, and the moisture inside the cut tobacco gradually flows to the surface of the cut tobacco and evaporates into the air until the humidity in the field along the drum 1.1 m ~ 1.5 m decreases slightly, showing a stable trend, but the humidity at both ends near the drum wall at the outlet is slightly higher than that at the middle. In this section, the inner cylinder environment is also in the deceleration stage, the heat and mass transfer of air flow and cut tobacco tend to be stable, and the humidification rate of the inner cylinder environment decreases to zero.

The reason for the change in Figure 9 is that during the production process, the moisture at the inlet of the leaf silk did not change obviously, and the outlet temperature of the leaf silk from the previous process gradually increased. During this period, the temperature of the cylinder wall was low, the difference between the temperature of the leaf silk and the ambient temperature in the drum became smaller, and the water evaporation of the leaf silk was slow, resulting in the increase of the moisture at the outlet of the leaf silk. After being regulated to the 60th sampling point, the system gradually returned to stability.

In Table 3, the three algorithms can detect the process abnormal fluctuations of two batches of data to varying degrees, and from PCA to AE-PCA and AE-MPCA, the alarm rate of abnormal data gradually increases. The AE-MPCA algorithm has the highest alarm rate in both cases.

As can be seen from Figure 10, the alarm rate of the traditional PCA method for case 2 is 91.07%, and after AE feature extraction, the alarm rate reaches 92.86%. Although a relatively large range of alarms has been achieved, the proposed AE.MPCA method can still increase the alarm rate to 98.21%, and accurately alarm the later stage of abnormal working conditions.

The above examples show that compared with the traditional PCA and AE-PCA detection methods, the AE-MPCA algorithm proposed in this paper improves the abnormality detection accuracy of drum leaf drying production process, and realises accurate alarm for quality abnormalities. Early detection of abnormal phenomena in actual production is the premise of early prevention and early solution. This block modelling method can lay a foundation for operators to accurately locate the causes of abnormal working conditions from material and machine in the next step, further provide accurate production status information, minimise the influence of abnormal working conditions on the quality of leaves on the production line as much as possible, and help to ensure the stability of production process and improve the stability of product quality.

5 Conclusions

With the increasing complexity of production equipment, the drying system of tobacco industry is increasing, which leads to the probability of failure in the production process increasing. At the same time, the difficulty of control of the system increases, so the diagnosis of failure and the control of the system are widely concerned. This paper proposes a PCA multi-block modelling algorithm based on autoencoder feature extraction, extracts autoencoder features for each sub-block, and regards the feature information as a new observation variable to establish a PCA monitoring model. The results show that the monitoring results are more intuitive by fusing the statistics of all sub-blocks through Bayesian reasoning. Moreover, compared with the traditional PCA and AE-PCA detection methods, the AE-MPCA algorithm proposed in this paper improves the anomaly detection accuracy of the drum leaf drying production process and realises accurate alarm for quality anomalies.

When training autoencoder, it is often necessary to adjust parameters and overfitting phenomenon. Therefore, how to improve the model so that it can adaptively adjust parameters and be better suitable for anomaly detection based on drum leaf drying process can be further studied.

Acknowledgements

Rongya Zhang and Mingchang Liu contributed equally to this work.

Declarations

Authors declare that they have no conflict of interest.

References

Chong, J. (2023) 'An intelligent detection approach for smoking behavior', *International Journal of Cognitive Informatics and Natural Intelligence* (*IJCINI*), Vol. 17, No. 1, pp.1–18.

Du, S., **, G., Chen, X., Wang, X., Ai, B., Wu, S., ... and Zhao, W. (2023) 'Applications of electronic nose technology in the tobacco industry: a comprehensive review', *Measurement Science and Technology*, Vol. 35, No. 3, pp.32002–32014.

Gang, S., **aoming, W., Junfeng, W., Shufang, L., Zhuowen, L. and Wang, X. (2021) 'Analysis of the information management system in the manufacturing process of cigarette enterprises using fuzzy AHP', *Journal of Intelligent and Fuzzy Systems*, Vol. 40, No. 4, pp.8257–8267.

Gong, H., Hu, R., Ding, X., Pang, S. and Che, Y. (2023) 'Intelligent large-scale flue-cured tobacco grading based on deep densely convolutional network', *Scientific Reports*, Vol. 13, No. 1, pp.11119–11130.

Guo, D. and Hu, Q. (2022) 'Design of multi-indicator integrated testing system for tobacco intelligent silk production line', *Journal of Intelligent and Fuzzy Systems*, Vol. 42, No. 3, pp.2615–2627.

Hu, Y., Dong, G., Wang, B., Liu, X., Wen, J., Dai, M. and Wu, Z. (2024) 'NSGA-II algorithm-based automated cigarette finished goods storage level optimization research', *Advanced Control for Applications: Engineering and Industrial Systems*, Vol. 6, No. 4, pp.e171–e183.

Liu, X. (2023) 'Analysis on the development of cigarette packaging in the era of intelligence', *Journal of Electronic Research and Application*, Vol. 7, No. 3, pp.1–6.

Liu, X., Li, J., Wang, H., Jia, W., Yang, J. and Guo, Z. (2023) 'Design of an optimal scheduling control system for smart manufacturing processes in tobacco industry', *IEEE Access*, Vol. 11, No. 1, pp.33027–33036.

Mu, L., Bi, S., Yu, S., Liu, X. and Ding, X. (2022) 'An intelligent moisture prediction method for tobacco drying process using a multi-hierarchical convolutional neural network', *Drying Technology*, Vol. 40, No. 9, pp.1791–1803.

Rui, J., Qiu, D., Hou, S., Rong, J., Qin, X., Fan, J., ... and Zhu, C. (2023) 'Deep learning-based intelligent control of moisture at the exit of blade charging process in cigarette production', *Applied Mathematics and Nonlinear Sciences*, Vol. 9, No. 1, pp.12–22.

Wang, C., Zhao, J., Yu, Z., **e, S., Ji, X. and Wan, Z. (2022a) 'Real-time foreign object and production status detection of tobacco cabinets based on deep learning', *Applied Sciences*, Vol. 12, No. 20, pp.10347–10360.

Wang, L. and Zhang, C. (2023) 'Artificial intelligence and big data in the production process to optimise the parameters of the cut tobacco-making process', *International Journal of Information Technology and Management*, Vol. 22, Nos. 3–4, pp.315–334.

Wang, W., Zhang, L., Wang, J. and Long, Z. (2022b) 'Locating defects and image preprocessing: deep learning in automated tobacco production', *Journal of Sensors*, Vol. 2022, No. 1, pp.6797207–6797218.

Wang, X., Li, H., Wang, Y., Fu, B. and Ai, B. (2024) 'Intelligent detection and odor recognition of cigarette packaging paper boxes based on a homemade electronic nose', *Micromachines*, Vol. 15, No. 4, pp.458–470.

Wang, X., Wang, X., Wu, J., Zheng, K., Pang, Y. and Gang, S. (2021) 'Research on quality traceability of cigarette by combining PDCA quality cycle with information strategy based on fuzzy classification', *Journal of Intelligent and Fuzzy Systems*, Vol. 40, No. 4, pp.8217–8226.

Wei, J., Wang, Z., Li, S., Wang, X., Xu, H., Wang, X., ... and Mei, C. (2024). Prediction modeling of cigarette ventilation rate based on genetic algorithm backpropagation (GABP) neural network', *EURASIP Journal on Advances in Signal Processing*, Vol. 2024, No. 1, pp.25–37.

Xu, J., Chen, Y., Yan, H. and Yang, X. (2023) 'Optimizing electromagnetic cigarette heaters using PSO-NSGA II algorithm: an effective strategy to improve temperature control and production rate', *Applied Artificial Intelligence*, Vol. 37, No. 1, pp.2222257–2222270.

Xu, Z., Wang, C., Wang, X. and Zhang, J. (2022) 'Robotics for control strategy optimization and simulation in automatic cigarette sorter', *Computers and Electrical Engineering*, Vol. 99, No. 1, pp.107801–107812.

Zhong, L., Rao, W., Zhang, X., Zhang, Z. and Krolczyk, G. (2023) 'Operation parameters optimisation of a machine swarm using artificial intelligence', *Elektronika ir Elektrotechnika*, Vol. 29, No. 5, pp.79–85.

Zhu, M., Wu, K., Zhou, Y., Wang, Z., Qiao, J., Wang, Y., ... and Zi, W. (2021) 'Prediction of cooling moisture content after cut tobacco drying process based on a particle swarm optimization-extreme learning machine algorithm', *Math. Biosci. Eng.*, Vol. 18, No. 2, pp.2496–2507.