APEM journal

**Advances in Production Engineering & Management**
Volume 17 | Number 3 | September 2022 | pp 326–338
https://doi.org/10.14743/apem2022.3.439

**ISSN 1854-6250**
Journal home: apem-journal.org
Original scientific paper

# A new approach for quality prediction and control of multistage production and manufacturing process based on Big Data analysis and Neural Networks

**Tian, S.[a], Zhang, Z.[b], Xie, X.[b,*], Yu, C.[c]**

[a]School of Basic Sciences, Shandong Institute of Petroleum and Chemical Technology, Dongying, P.R. China
[b]Beijing Jiaotong University, School of Economics and Management, Beijing, P.R. China
[c]School of Information and Electronics, Beijing Institute of Technology, Beijing, P.R. China

**A B S T R A C T**

As consumers care more and more about product quality, it is important to mine the deep correlations between production and manufacturing parameters and the evaluation of product quality through the analysis of industrial big data. The existing research of product quality prediction faces several major problems: the lack of diverse quality features, the poor tractability of abnormal parameters, the strong nonlinearity of parameters, the obvious sequential property of data, and the severe time lag of data. To solve these problems, this paper explores the quality prediction and control of multistage MP process (MPMP) based on big data analysis. Firstly, the prediction strategy and flow were specified for MPMP product quality prediction, and the features were extracted from MPMP product quality. After that, the MPMP product quality features were described in multiple dimensions, the attention mechanism was introduced to the prediction process. In addition, the recurrent neural network was improved, and an MPMP product quality prediction model was established on bidirectional long short-term memory (BiLSTM) network. Our model was compared with AdaBoost and XGBoost through experiments. The effectiveness of our model was demonstrated by the results of the appearance quality PQ1, and the area under the curve (AUC) for each process parameter. In general, our model is superior to other algorithms in the accuracy, mean accuracy, and precision of product quality prediction.

**A R T I C L E   I N F O**

*Keywords:*
Big data analysis;
Multistage production and manufacturing process (MPMP);
Quality prediction;
Machine learning;
Artificial neural network;
Recurrent neural network;
Bidirectional long short-term memory (BiLSTM)

**Corresponding author:*
xxie@bjtu.edu.cn
(Xie, X.)

*Article history:*
Received 9 May 2022
Revised 16 August 2022
Accepted 20 August 2022



## 1. Introduction

With the continuous advancement of manufacturing intelligence, the new generation information technology (IT) has been deeply integrated with intelligent, automated production technology, providing an important driver of the industrial transformation of modern manufacturing [1-7]. As products become increasingly personalized, customized, and complex, the production and manufacturing (MP) process tends to cover multiple stages, and the consumers care more and more about product quality [8-14].

The manufacturing Internet of things (IoT) collects and stores the changing parameter data of the MP process, forming the industrial big data [15-22]. Through big data analysis, it is possible to reveal the deep connections between the parameters and the evaluation of product quality, providing support to the optimization of manufacturing technology and the prediction of product quality in modern manufacturing.

In modern industries, information has been fully shared between production machines, smart subsystems, and mobile devices through advanced network technology [23, 24]. Ren *et al.* [25] proposed a soft measuring method for product quality prediction based on semi-supervised deep parallel factorization-machine (deepFM). Specifically, the process variables were discretized through data blocking; the deepFM was improved to extract quality information from different components, and obtain both high- and low-dimensional features. To solve the lack of process data and the strong nonlinearity and multiscale property of new intermittent processes, Chu *et al.* [26] presented a product quality prediction approach based on multi-scale kernel JYMKPLS (Joint-Y multi-scale kernel partial least squares) transfer learning. The merits of transfer learning and multiscale kernel learning were fully inherited by their approach. In addition, Lughofer *et al.* designed the strategies for online update and data removal of the model, aiming to adapt the transfer model continuous to new batch processes. Lughofer *et al.* [27] suggested that, in modern manufacturing facilities, high-quality production can be guaranteed through two basic stages. The first stage is to recognize the possible problems of product quality early on. The second stage is to take proper responses to the recognized problems. On this basis, Hao *et al.* put forward a holistic method to continuously implement the two stages under the prediction and maintenance framework of online production system, constructed a data-driven prediction model based on product quality standard, and realized the cyclic technological optimization in support of multistage functions. Hao *et al.* [28] developed an interaction model, which uses a linear model to characterize the influence of old production techniques and machines over quality degradation, and employs a random differential equation to capture the factors affecting quality degradation. Based on multiphase support vector regression (SVR), Zheng and Pan [29] proposed a soft measurement model for online quality prediction of liquid product concentration, and proved that the model is more efficient than the reported technologies. Melhem *et al.* [30] monitored the real-time quality of electronic products through the autocorrelation multivariate process, and came up with an online product quality prediction method, which forecasts the quality of products based on the correlations between product quality measurements in different steps.

The existing research of product quality prediction faces several major problems: the lack of diverse quality features, the poor tractability of abnormal parameters, the strong nonlinearity of parameters, the obvious sequential property of data, and the severe time lag of data. To solve these problems, this paper explores the quality prediction and control of multistage MP process (MPMP) based on big data analysis. Section 2 specifies the prediction strategy and flow for MPMP product quality prediction. Section 3 introduces recursion to the principal component analysis (PCA) based on kernel functions, and extracts the features of MPMP product quality. Section 4 describes the MPMP product quality features in multiple dimensions, and incorporates the attention mechanism to prediction process. In addition, the recurrent neural network was improved, and an MPMP product quality prediction model was established on bidirectional long short-term memory (BiLSTM) network. Our model was compared with AdaBoost and XGBoost through experiments. The effectiveness of our model was demonstrated by the results of the appearance quality $PQ1$, and the area under the curve (AUC) for each process parameter.

## 2. Quality prediction framework

Fig. 1 shows the MPMP product quality prediction strategy. Wireless sensing, industrial IoT, and big data analysis are adopted to capture and store the PM states and production technology parameters in different MP stages, aiming to facilitate the big data analysis. Fig. 2 shows the flow of MPMP product quality prediction. The specific steps of the prediction are as follows.

Firstly, the historical MPMP data are combined with the data on manufacturing technology, MP experience, and abnormal product quality into a real-time MPMP and product quality database. Next, the MPMP product quality is monitored in real time, referring to the product quality indices. After that, the MPMP data in the current stage and previous stages are mined through big data analysis. On this basis, a product quality prediction model is established, and a strategy is developed to optimize abnormal production technology parameters.

The data from the historical MP and product quality database are preprocessed before being used to predict whether the product is qualified, update the product quality rule library, and optimize the abnormal production technology parameters.

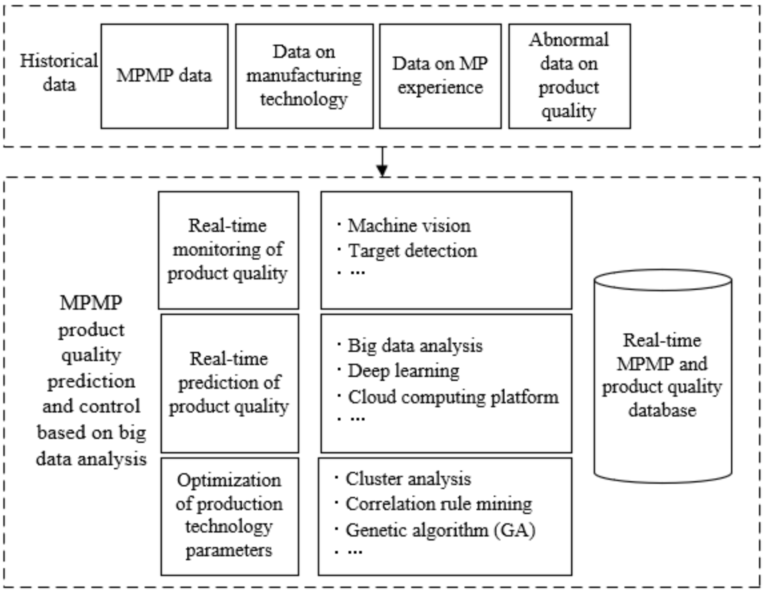



**Fig. 1** MPMP product quality prediction strategy

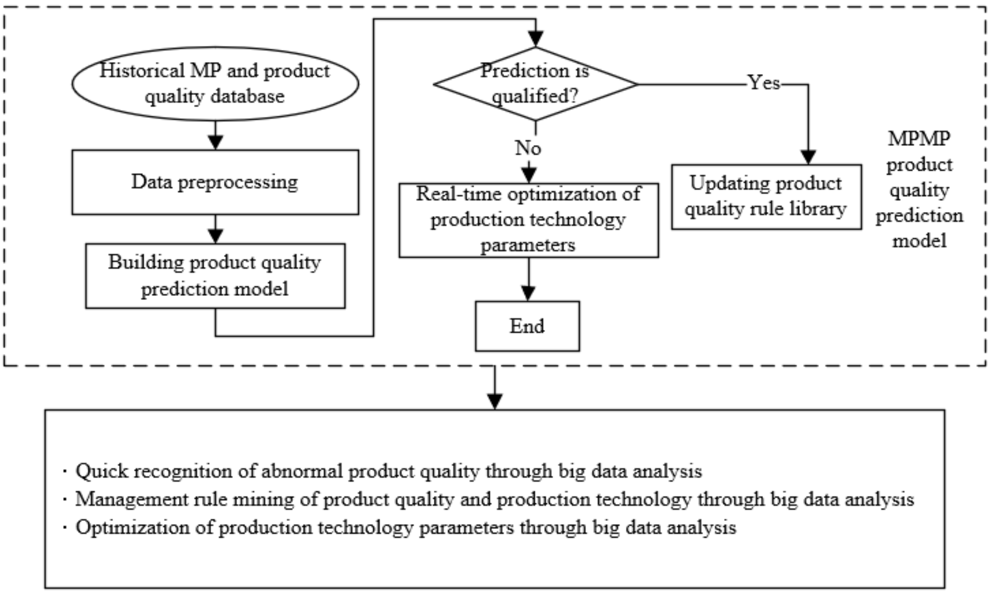



**Fig. 2** Flow of MPMP product quality prediction

## 3. Feature extraction

MPMP product quality data contains a rich information of MP process, and many redundant factors. The variables in the data are strongly correlated, and cannot be directly used to build the MPMP product quality prediction model. To solve the problem, it is necessary to extract low-dimensional features from the data to obtain the most representative data of the MPMP, thereby reducing the modeling difficulty and computing complexity.

The recursive update of principal components was introduced to the PCA based on kernel functions. In this way, the model can be updated without needing overall computing. The model computing becomes less complex, and satisfies the real-time online requirement of MPMP product quality prediction. Fig. 3 explains the flow of extracting MPMP product quality features.

The kernel-based PCA recursive update algorithm mainly has four steps:

Step 1: The original MPMP product quality data are mapped from low-dimensional feature space to high-dimensional feature space, producing the reconstructed data $\psi(a_i)$.

Step 2: The kernel matrix $L(a_i, a_j) = \{\psi(a_i), \psi(a_j)\}$ is constructed, and the inner product of vectors in the original data space is converted into kernel functions.

Step 3: In the high-dimensional feature space, the kernel functions obtained in the previous steps are subjected to PCA, and the corresponding eigenvalues and eigenvectors are calculated.

Step 4: The MPMP product quality data in the new stage are introduced to the kernel space through recursive solving.

Step 5: The kernel-based PCA model of product quality is updated recursively, and the dynamic PCA features of MP product quality are extracted from the new stage.

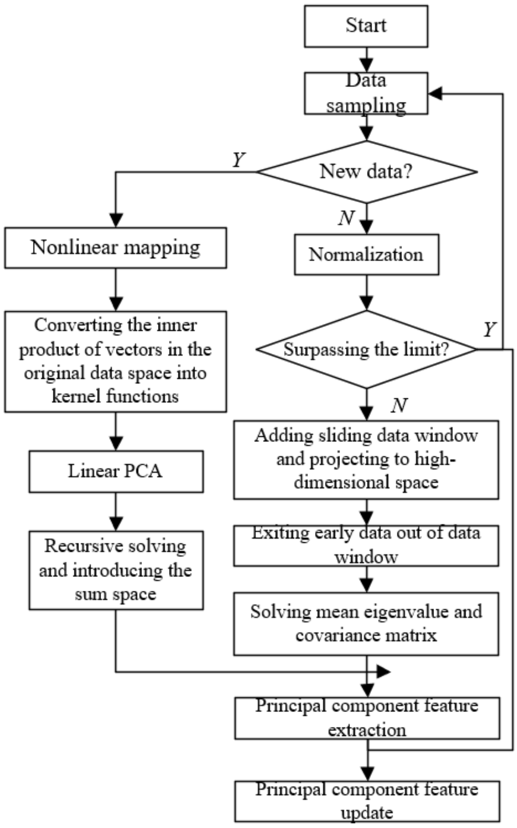



**Fig. 3** Flow of MPMP product quality feature extraction

Let $a_i$ be the n-dimensional MPMP product quality variable obtained through m samplings under the normal working condition; $E^n$ be the original MPMP product quality data; $G$ be the high-dimensional feature space; $\psi(a_i) \in G \subseteq E^n$ be the product quality data mapped nonlinearly from $E^n$ to $G$. Then, we have

$$\psi: E^n \rightarrow G \quad a \rightarrow \psi(a) \tag{1}$$

Let $\lambda_\psi$ be the mean of the data mapped to space $G$ from the original product quality samples. The covariance matrix $D_G$ of the MP product quality samples in the new stage after the mapping can be calculated by:

$$D_G = \frac{1}{m}\sum_{i=1}^{m}\left(\psi(a_i) - \lambda_\psi\right)\left(\psi(a_i) - \lambda_\psi\right)^T = \frac{1}{m}\sum_{i=1}^{m}\psi(a_i)\psi(a_i)^T \tag{2}$$

where, $\lambda_\psi$ can be calculated by:

$$\lambda_\psi = (1/m)\sum_{i=1}^{m}\psi(a_i) \tag{3}$$

Suppose the mapping values of the original product quality samples in space $G$ are zero-centered. Then, $\lambda_\psi$ is equal to zero. To obtain the corresponding eigenvalue $\mu$ and eigenvector $U$, $D_G$ should go through eigenvalue decomposition in space $G$:

$$\mu U = D_G U \tag{4}$$

Solving the inner product of each product quality sample, i.e., pre-multiplying Eq. 4 with $\psi(a_l)$, $l$ = 1, 2, ..., $m$:

$$\mu(\psi(a_l)\cdot U) = (\psi(a_l)\cdot D_G U) \tag{5}$$

According to the reproducing kernel theory, all the eigenvectors $U$ corresponding to nonzero eigenvalues $\mu$ are all distributed in the tensor space of {$\psi(a_1)$, $\psi(a_2)$, ..., $\psi(a_m)$}. Then, the coefficient vector $\beta_i$($i$ = 1, 2, ..., $m$) should satisfy:

$$U = \sum_{i=1}^{m}\beta_i\,\psi(a_i) \tag{6}$$

Combining Eqs. 5 and 6:

$$\mu\sum_{i=1}^{m}\beta_i\langle\psi(a_i),\psi(a_l)\rangle = \frac{1}{m}\sum_{i=1}^{m}\beta_i\left\langle\psi(a_i)\bullet\sum_{j=1}^{n}\psi(a_j)\right\rangle\bullet\langle\psi(a_j),\psi(a_i)\rangle \tag{7}$$

Let $L_{ij} = L(a_i, a_j)$ = {$\psi(a_i)$, $\psi(a_j)$} be an $n \times n$-order kernel matrix. That is, the kernel functions of the two variables in the input space can characterize the inner product between the variables in the feature spaceThe specific form of $\psi$ should be avoided in the computing process. Substituting $L_{ij}$ into Eq. 7, the characteristic equation can be obtained:

$$m\mu L\beta = L^2\beta \Rightarrow m\mu\beta = L\beta \tag{8}$$

Solving Eq. 8, it is possible to obtain nonzero eigenvalues $\mu_2$($\mu_1 \geq \mu_2 \geq ... \mu_m$) and the corresponding eigenvectors $\beta^l$($l$ = 1, 2, ..., $m$). There exists ($u_l$,$u_l$) = 1 for normalizing eigenvector $U_l$. Thus, it can be derived from Eq. 7 that $\mu(\beta^l - \beta^l) = 1$. Then, the $l$-th principal component $p_l$ mapped from original MPMP product quality sample $a$ to space $G$ can be expressed as:

$$p_l = \langle u_l,\psi(a)\rangle = \sum_{i=1}^{m}\beta_i^l\langle\psi(a_i)\bullet\psi(a)\rangle = \sum_{i=1}^{m}\beta_i^l L(a,a_i) \tag{9}$$

The principal component direction of $a$ in space $G$ is the projection of $\psi(a)$ on new eigenvector $U$. After being added to the data window, the MP data $a_{m+1}$ of a new stage should be mapped to space $G$. Let $\psi(a_{m+1})$ be the nonlinearly mapped value of $a_{m+1}$ from low-dimensional feature space to high-dimensional feature space. Then, the mapped mean eigenvalue $v'_\psi$ can be recursively computed by:

$$\lambda_\psi' = \frac{1}{m+1}\psi([A,a_{m+1}])S_{m+1} = \frac{m}{m+1}\lambda_\psi + \frac{1}{m+1}\psi(a_{m+1}) \tag{10}$$

where, $S$ = [1, 1, ..., 1] ∈ $E^{m+1}$. The covariance matrix $D'_G$ can be recursively solved by:

$$\begin{aligned}
D_G' &= \frac{1}{m}\overline{\psi}([A,a_{m+1}])\overline{\psi}([A,a_{m+1}])^T\\
&= \frac{1}{m}\sum_{i=1}^{m}\left(\psi(a_i)-\lambda_\psi'\right)\left(\psi(a_i)-\lambda_\psi'\right)^T + \frac{1}{m}\sum_{i=1}^{m}\left(\psi(a_{m+1})-\lambda_\psi'\right)\left(\psi(a_{m+1})-\lambda_\psi'\right)^T\\
&= \frac{1}{m}\sum_{i=1}^{m}\left[\psi(a_i)-\lambda_\psi+\frac{1}{m+1}\lambda_\psi+\frac{1}{m+1}\psi(a_{m+1})\right]\times\left[\psi(a_i)-\lambda_\psi+\frac{1}{m+1}\lambda_\psi+\frac{1}{m+1}\psi(a_{m+1})\right]^T\\
&+\frac{1}{m}\left[\frac{m}{m+1}\psi(a_{m+1})-\frac{m}{m+1}\lambda_\psi\right]\times\left[\frac{m}{m+1}\psi(a_{m+1})-\frac{m}{m+1}\lambda_\psi\right]^T
\end{aligned} \tag{11}$$

$$
\begin{aligned}
&= \frac{1}{m}\sum_{i=1}^{m}\left(\psi(a_i)-\lambda_\psi\right)\left(\psi(a_i)-\lambda_\psi\right)^T + \frac{1}{m+1}\sum_{i=1}^{m}\left(\psi(a_{m+1})-\lambda_\psi\right)\left(\psi(a_{m+1})-\lambda_\psi\right)^T \\
&= \frac{m-1}{m}D_G + \frac{1}{m+1}\left(\psi(a_{m+1})-\lambda_\psi\right)\left(\psi(a_{m+1})-\lambda_\psi\right)^T \\
&= \frac{m-1}{m}\left[\sqrt{\frac{m-1}{m}}\overline{\psi}(a)\sqrt{\frac{m}{m^2-1}}\left(\psi(a_{m+1})-\lambda_\psi\right)\right] \times \left[\sqrt{\frac{m-1}{m}}\overline{\psi}(a)\sqrt{\frac{m}{m^2-1}}\left(\psi(a_{m+1})-\lambda_\psi\right)\right]^T
\end{aligned}
$$

To complete the eigenvalue decomposition of $v'_\psi$ and $D'_G$, linear PCA should be employed to extract the principal component features of the MP product quality in the new stage. Therefore, this paper introduces the recursive algorithm to the kernel-based PCA, eliminating the need to compute the entire principal component model again after the input of the MP product quality data of the new stage. In this way, our model can effectively reduce the computing load of real-time prediction.

## 4. Methodology

### 4.1 Multidimensional description of quality variables and attention mechanism

Fig. 4 describes MPMP product quality features in multiple dimensions. To illustrate the time-variation of product quality data, this paper evaluates product quality from three dimensions: production technology, product quality, and time. The production technology dimension examines the level of production technology from the accuracy of dimensions, shape, and position. The product quality dimension evaluates product quality from appearance quality, performance index, and life index. The time dimension is generated from the real-time MPMP product quality dataset.

The global decision-making of production control is underpinned by the online detection of key processes of each MP stage, and the information interaction and sharing of the time dimension. This paper extracts the samples from the time series of MPMP product quality, and constructs a multidimensional MPMP product quality prediction model, which is capable of tracing and optimizing the abnormal production technology parameters. Fig. 5 shows the flow of real-time tracking and optimization of abnormal production technology parameters.

Under the different features of production technology parameters, the product quality features of different MP processes are affected to different degrees. To this end, the attention mechanism was innovatively introduced to MPMP product quality prediction. Firstly, the MP product quality of the new stage was analyzed against the quality evaluation criteria. Then, the similarity between product quality features of adjacent stages was calculated. After that, the authors computed the correlation of the product quality features in the current MP stage with the change law of different production technology parameters, and obtained the corresponding similarity score. On this basis, the abnormal technology parameters were valued for the new MP stage, thereby completing the prediction of MPMP product quality.

The core parameters of the attention mechanism include Key, Query, and Value. Among them, Key corresponds to Value, and is used to query for and compute the similarity of product quality features between adjacent stages. Query is the query during the execution of the attention mechanism. Value is the MPMP product quality data receiving attention and being selected. The attention mechanism can be defined by:

$$ATT(Key, Query, Value) = Softmax\left(SIM(Key, Query)\right) * Value \tag{12}$$

Let $q_i$ be the output tensor of the upper layer series at position $i$. If the upper layer is a bidirectional neural network in the model, then the Key at position $i$ can be denoted by $u_i$:

$$u_i = tan(\omega q_i + r) \tag{13}$$

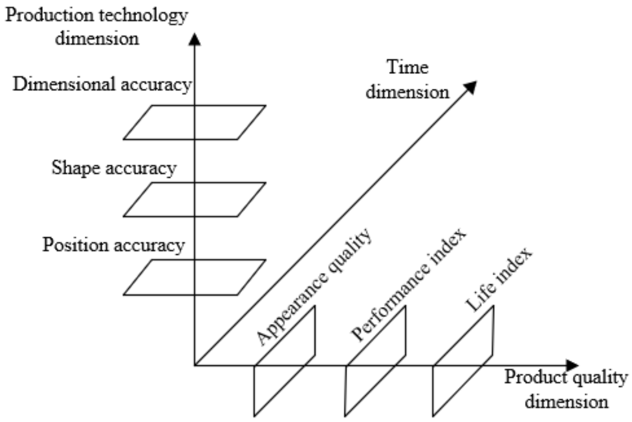



**Fig. 4** Multidimensional description of MPMP product quality features

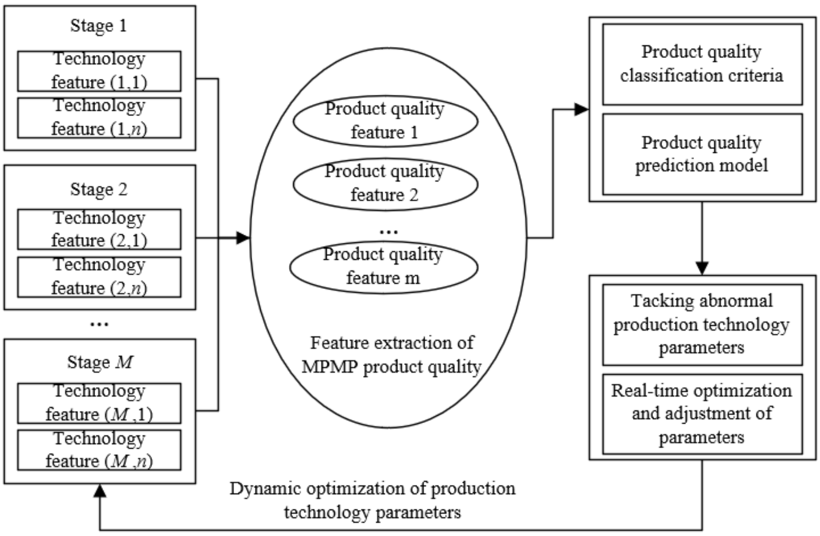



**Fig. 5** Real-time tracking and optimization of abnormal production technology parameters

Eq. 13 shows that $u_i$ can be derived from Value via the fully-connected layer. The attention weight $\beta_i$ can be calculated by:

$$\beta_i = \frac{o^{u_i{}^T}d}{\sum_i o^{u_i{}^T}d} \tag{14}$$

The output of the attention layer can be calculated by:

$$SC = \sum_i \beta_i\,(a)u_i(a) \tag{15}$$

The Query of that attention layer can be adjusted and updated through random initialization and model training:

$$\omega_p(i) = Soft\,max\left(\gamma_p SIM\left(a_p, N_p(i)\right)\right) \tag{16}$$

$$\omega_p(i) = Soft\,max\left(\left(\sum_l \omega_p\ (l)t_p(i+l)\right)^{\alpha_p}\right) \tag{17}$$

The output tensor $f_i$ can be expressed as:

$$q_i = CON(\vec{q}_i, \overleftarrow{q}_i) \tag{18}$$

### 4.2 Construction of product quality prediction model

The LSTM network, as a variant of recurrent neural network, can prevent the common problems of recurrent neural network: exploding gradients and vanishing gradients. Each LSTM network consists of an input gate $SR_p$, a forget gate $LE_p$, an output gate $EX_p$, and a memory cell $d_p$. Let $A_p$ be the input vector of stage $p$; $f_{p\text{-}1}$ be the hidden layer state of stage $p - 1$. Then, the forget gate can be calculated by:

$$A = \left[\frac{A_p}{f_{p-1}}\right] \tag{19}$$

$$LE_p = \varepsilon\left(\omega_{LE} \cdot \left[f_{p-1}, A_p\right] + r_{LE}\right) \tag{20}$$

The forget gate determines whether the MPMP product quality information should be preserved or discarded. Let $D_p{}^*$ be the candidate cell state used to determine the new cell state $D_p$. The input gate can be calculated by:

$$SR_p = \varepsilon\left(\omega_i \cdot \left[f_{p-1}, A_p\right] + r_i\right) \tag{21}$$

$$D_p{}^* = tan\left(\omega_d \cdot \left[f_{p-1}, A_p\right] + r_d\right) \tag{22}$$

The input gate determines how to update information. The training parameters of our model are configured as follows: the weight of LSTM network, $\theta$; bias, $r$; activation function sigmoid, $s$; point-by-point product, $\odot$. The output gate can be calculated by:

$$D_p = LE_p * D_p + SR_p * D_p{}^* \tag{23}$$

$$EX_p = \varepsilon\left(\theta_{EX} \cdot \left[f_{p-1}, A_p\right] + r_{EX}\right) \tag{24}$$

The memory cell can be expressed as:

$$d_p = LE_p \otimes d_{p-1} + SR_p \otimes tan(\theta_d \cdot A + r_d) \tag{25}$$

The hidden layer state of stage $p$ can be calculated by:

$$f_p = EX_p \otimes tan\left(d_p\right) \tag{26}$$

The activation function can be defined as:

$$\varepsilon(a) = \frac{1}{1 + o^{-a}} \tag{27}$$

$$tan(a) = \frac{o^a - o^{-a}}{o^a + o^{-a}} \tag{28}$$

The LSTM network can effectively handle product quality data with a strong sequential property and correlations. However, the network is unable to mine the information of future product quality data. This defect can be solved by combining forward LSTM with inverse LSTM into a BiLSTM. Let $\delta$ be the length of a data series; ML be the processing of a single LSTM. Then, the BiLSTM can obtain the forward and inverse information of product quality data from input vectors $f_p$:

$$\vec{f_p} = \overrightarrow{ML}\left(\vec{f}_{p-1}, a_p, d_{p-1}\right), p \in [1, \delta] \tag{29}$$

$$\overleftarrow{f_p} = \overleftarrow{ML}\left(\overleftarrow{f}_{p-1}, a_p, d_{p+1}\right), p \in [1, \delta] \tag{30}$$

The information in the two directions (Eqs. 29 and 30) is combined into the final output of BiLSTM. The output of stage $p$ can be expressed as:

$$b_p = h\left(\omega_{\vec{f}b}\vec{f_p} + W_{\overleftarrow{f}b}\overleftarrow{f_p} + r_b\right) \tag{31}$$

Let $A_{p\text{-}m}$ and $f_{p\text{-}m}$ be the state input and hidden layer output of the previous $n$ MP stages, respectively; $A_p$ be the input of production technology parameters in the current stage; $f_p$ be the output through the hidden layer; $A_{p+m}$ be the input of production technology parameters in the $m$ MP stages in future; $f_{p+m}$ be the hidden layer output. Then, we have:

$$f_p = g\left(\omega_3 A_p + \omega_4 f_{p-1}\right) \tag{32}$$

$$f_p^{'} = LE\left(\omega_1 A_p + \omega_2 f_{p-1}^{'}\right) \tag{33}$$

$$EX_p = b_p = h\left(\omega_5 f_p + \omega_6 f_p'\right) \tag{34}$$

The dynamic memory network capable of retaining time dependence can be established by:

$$f_i^p = h_i^p BiLSTM\left(d_i, f_{i-1}^p\right) + \left(1 - h_i^p\right) f_{i-1}^p \tag{35}$$

Considering the sequential property and strong correlations between production technology parameters, as well as the dynamic nonlinearity of product quality features, this paper assigns weights to different production technology parameters according to the correlations between the two variables, following the attention mechanism of quality variables, and takes the weighed parameters as the input of BiLSTM.

Firstly, the key parameters that induce the abnormal product quality were identified for each MP stage. Next, a product quality prediction model was established based on BiLSTM. To improve the prediction accuracy, the product quality states of stages $p$ + 1, ..., $p$ + $m$ were predicted in real time in stage $p$. Then, the key technology parameters were further optimized.

The production technology parameters of historical MPMP and those $Z_{ij}$ of the current stage were converted into time series stage by stage. Let $Z_{mi}(p - 1)$ be production technology parameter $i$ of MP process $m$ in stage $p$ – 1, and $F(p)$ be the corresponding product quality feature. Then, we have:

$$\left[ Z_{ij} = \begin{pmatrix} Z_{11}(p-1) & \cdots & Z_{1n}(p) \\ Z_{21}(p-1) & \cdots & Z_{2n}(p) \\ \vdots & \ddots & \vdots \\ Z_{m1}(p-1) & \cdots & Z_{mn}(p) \end{pmatrix} \right] \Rightarrow \begin{bmatrix} F(p) \\ F(p+1) \\ \vdots \\ F(p+m) \end{bmatrix} \tag{36}$$

Let $\xi_{ij}(l + 1)$ be the observation of stage $l$ + 1; $\xi_{max}$ and $\xi_{min}$ be the maximum and minimum of a parameter in that stage, respectively. Then, the normalization can be expressed as:

$$Z_{ij}(l+1) = \frac{\xi_{ij}(l+1) - \xi_{max}}{\xi_{min} - \xi_{max}} \tag{37}$$

Let $F^*(l + 1)$ be the prediction for stage $l$ + 1; $\gamma_k$ be the weight of the similarity between the parameter and product quality features; $\chi_k(l + 1)$ be the corresponding attention weight. Then, we have:

$$F^*(l+1) = \sum_{k=1}^{K} \chi_k(l+1)\, \gamma_k{}^* a_{ij}(l+1) \tag{38}$$

The prediction can be expanded as:

$$F^*(l+1) = LE[F(l), F(l-1), \ldots, F(l-m+1); Z(l+1), Z(l), Z(l-n+1); \sigma(l+1)] \tag{39}$$

Eq. 39 shows that the output of product quality features in stage $l$ + 1 depends on $Z(l + 1)$, ($(Z(l)$, ..., $Z(l - n + 1))$, ($F(l)$, $F(l - 1)$, . . . , $F(l - m + 1)$) and an unpredictable error $\sigma$. Let $F_{i1}$ and $F_{i1}^*$ be the actual value and predicted value of a quality feature, respectively; $\tau||\Phi||^2$ be the regularization term of the loss function. During model training, the loss function of product quality prediction can be expressed as:

$$loss(\Phi) = -\sum_{i=1}^{m} F_{i1}{}^* log(F_{i1}) + F_{i2}{}^* log(F_{i2}) +, \ldots, F_{in}{}^* log(F_{in}) + \tau\|\Phi\|^2 \tag{40}$$

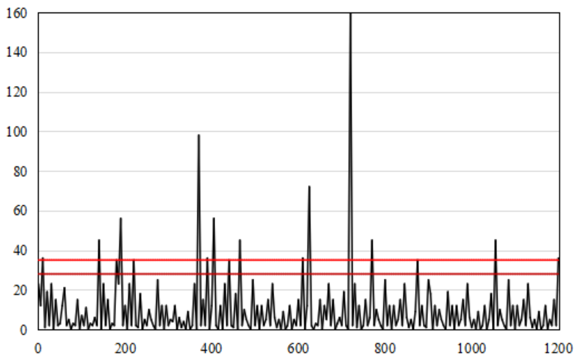
The fully-connected layer of the model was activated by Softmax function. The final output of the product quality prediction model can be given by:

$$\hat{b} = Softmax(a_i) = \frac{o^i}{\sum_{n=1}^{n} o^i} \tag{41}$$

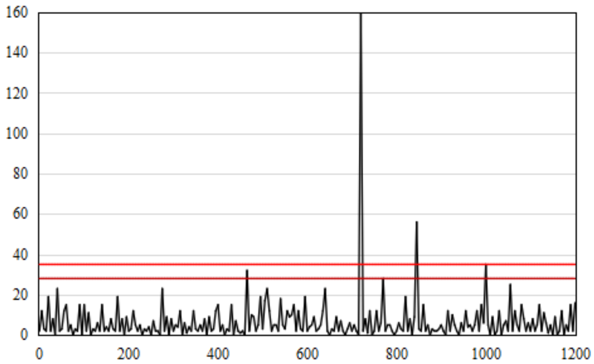
## 5. Experiments and results analysis

Fig. 6 compares the PCA results before and after the introduction of recursion. In Fig. 6a, the quality samples of 1,200 products were not recursively processed, but statically analyzed. Many PCA results surpassed the control line, indicating that the model could not effectively extract the principal component features of product quality in the latest MP stage. In Fig. 6b, the quality samples of 500 out of the 1,200 samples were recursively processed. It can be observed that the real-time performance of the analysis on samples collected by sliding time windows was effectively enhanced, as the results above the control line were reduced by more than 8 %.

In our prediction model, the MPMP production technology parameters and MPMP product quality variables being collected and stored were preprocessed and normalized, so that the product quality prediction accuracy will not be undermined by dimensional disunity. After that, the product quality variation of stage $p$ was predicted based on the samples of stage $p$ – 1. The specific data are given in Tables 1 and 2, where the product quality indices include appearance quality PQ1, performance index PQ2, and life index PQ3. Specifically, PQ1 covers cleanness PQ11, style PQ12, color PQ13, and packaging PQ14; PQ2 covers dimension parameter PQ21, motion parameter PQ22, and power parameter PQ23; PQ3 covers reliability PQ31, failure rate PQ32, and use cycle PQ33.




(a) Before




(b) After

**Fig. 6** PCA results before and after the introduction of recursion

**Table 1** Normalized values of samples in stage $p$ – 1

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $PQ11$ | 0.9625 | 0.9657 | 0.9686 | 0.9628 | 0.9639 | 0.9583 | 0.9655 | 0.9662 | 0.9675 |
| $PQ12$ | 0.1658 | 0.1534 | 0.1643 | 0.1725 | 0.1848 | 0.1538 | 0.1342 | 0.1658 | 0.1546 |
| $PQ13$ | 0.0241 | 0.0325 | 0.0521 | 0.0375 | 0.0265 | 0.0164 | 0.0335 | 0.0276 | 0.0215 |
| $PQ14$ | 0.9638 | 0.9254 | 0.9587 | 0.9352 | 0.9583 | 0.9285 | 0.9428 | 0.9647 | 0.9562 |
| $PQ21$ | 0.6521 | 0.6358 | 0.6149 | 0.6823 | 0.6951 | 0.6214 | 0.6725 | 0.6353 | 0.6438 |
| $PQ22$ | 0.3583 | 0.3869 | 0.3763 | 0.3242 | 0.3629 | 0.4233 | 0.4126 | 0.3584 | 0.3217 |
| $PQ23$ | 0.5382 | 0.6353 | 0.5876 | 0.6764 | 0.5372 | 0.6857 | 0.5728 | 0.6241 | 0.6834 |
| $PQ31$ | 0.1574 | 0.1758 | 0.1635 | 0.1634 | 0.1725 | 0.1838 | 0.1524 | 0.1326 | 0.1532 |
| $PQ32$ | 0.5271 | 0.3525 | 0.5241 | 0.3672 | 0.4258 | 0.3581 | 0.7868 | 0.5127 | 0.4136 |
| $PQ33$ | 0.3122 | 0.2174 | 0.2563 | 0.3727 | 0.1631 | 0.2836 | 0.2751 | 0.2675 | 0.3123 |

**Table 2** Normalized values of samples in stage $p$

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $PQ11$ | 0.9584 | 0.9625 | 0.9359 | 0.9433 | 0.9352 | 0.9375 | 0.9537 | 0.9742 | 0.9641 |
| $PQ12$ | 0.1546 | 0.1346 | 0.1738 | 0.1826 | 0.1628 | 0.1435 | 0.1628 | 0.1347 | 0.1527 |
| $PQ13$ | 0.0124 | 0.0135 | 0.0183 | 0.0175 | 0.0163 | 0.0157 | 0.0130 | 0.0175 | 0.0182 |
| $PQ14$ | 0.9258 | 0.9342 | 0.9328 | 0.9626 | 0.9731 | 0.9539 | 0.9582 | 0.9346 | 0.9504 |
| $PQ21$ | 0.6382 | 0.6426 | 0.6375 | 0.6832 | 0.6284 | 0.3505 | 0.3628 | 0.3751 | 0.3624 |
| $PQ22$ | 0.3581 | 0.3684 | 0.4213 | 0.3825 | 0.3862 | 0.4135 | 0.4253 | 0.3926 | 0.4257 |
| $PQ23$ | 0.4968 | 0.5326 | 0.4867 | 0.5315 | 0.4835 | 0.5327 | 0.4932 | 0.5147 | 0.4826 |
| $PQ31$ | 0.1538 | 0.1627 | 0.1628 | 0.1824 | 0.1736 | 0.1524 | 0.1438 | 0.1825 | 0.1926 |
| $PQ32$ | 0.2514 | 0.3012 | 0.2810 | 0.2624 | 0.2519 | 0.2746 | 0.2413 | 0.2534 | 0.2816 |
| $PQ33$ | 0.2386 | 0.2485 | 0.2104 | 0.2517 | 0.2825 | 0.2431 | 0.2042 | 0.2358 | 0.2719 |

This paper adopts two typical ensemble algorithms of machine learning, namely, AdaBoost and XGBoost, as well as our model to predict the product quality in terms of PQ1-3. The prediction results (Table 3) show that our model surpassed the other algorithms in terms of accuracy, mean accuracy, and precision.

Fig. 7 compares the AUCs of PQ1 predicted by AdaBoost, XGBoost, and our model. Our model achieved desirable prediction performance, despite the large difference of MPMP in product quality. Besides, XGBoost algorithm also demonstrated a good classification ability of various product quality features.

XGBoost and our model were separately adopted to predict PQ2 and PQ3 of 60 products with known production technology parameters. The results (Figs. 8 and 9) show that XGBoost incorrectly predicted 5 qualified products as unqualified in terms of PQ2, and incorrectly predicted 2 qualified products as unqualified in terms of PQ3. By contrast, our model successfully predicted all unqualified products. This means our model has a strong ability to predict product quality.

**Table 3** Production quality prediction results

| Product quality features | Indices | AdaBoost | *XGBoost* | Our model |
|---|---|---|---|---|
| *PQ*1 | Accuracy | 0.862 | 0.87 | 0.89 |
| | Mean accuracy | 0.85243 | 0.884672 | 0.953281 |
| | Precision | 0.842678 | 0.875689 | 0.925768 |
| *PQ*2 | Accuracy | 0.834 | 0.846 | 0.86 |
| | Mean accuracy | 0.868457 | 0.885674 | 0.896725 |
| | Precision | 0.863758 | 0.883758 | 0.895235 |
| *PQ*3 | Accuracy | 0.814 | 0.827 | 0.851 |
| | Mean accuracy | 0.823758 | 0.863074 | 0.875426 |
| | Precision | 0.825712 | 0.862875 | 0.886425 |

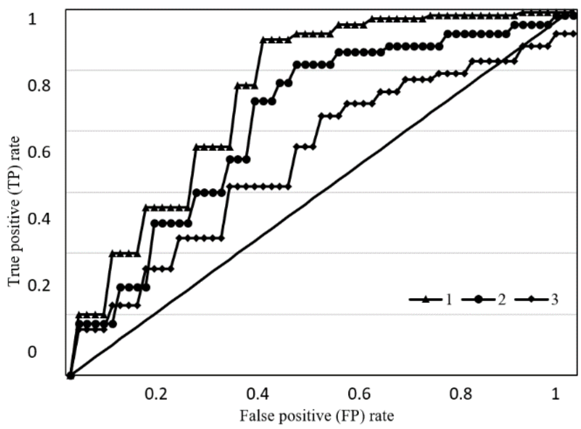



**Fig. 7** Area under the curve (AUC) of PQ1

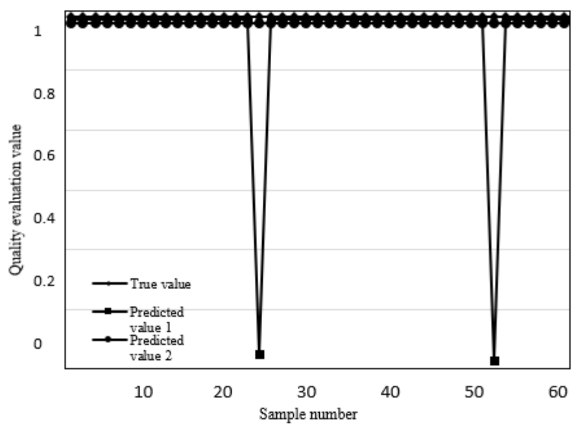



**Fig. 8** Predicted PQ2 vs. true PQ2

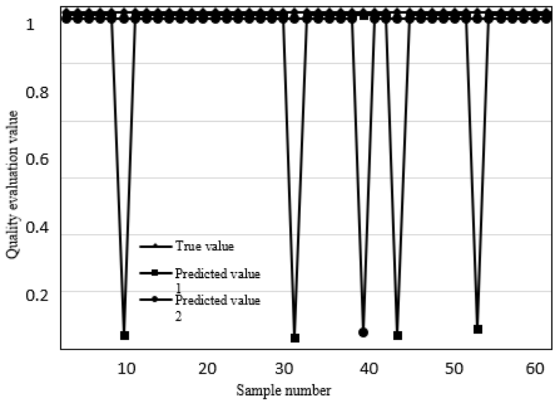



**Fig. 9** Predicted PQ3 vs. true PQ3

## 6. Conclusion

This paper analyzes the MPMP quality prediction and control based on big data analysis. Firstly, the MPMP product quality prediction strategy and flow were detailed, followed by the extraction of MPMP product quality features. Then, the MPMP product quality features were described in multiple dimensions. Next, the attention mechanism was introduced to the prediction process, and a product quality prediction model was established based on BiLSTM. Through experiments, the PCA results before and after the introduction of recursion were compared, revealing the effectiveness of our feature extraction algorithm. After that, our model and several state-of-the-arts were separately adopted to predict product quality in terms of PQ1-3. The results show that our model achieved better accuracy, mean accuracy, and precision of product quality than other algorithms.

This paper innovatively introduces the attention mechanism to establish a quality prediction model. But many details are worthy of further study. For example, the future research could try to optimize the product quality classification rules, and apply reinforcement learning to adaptively optimize various process parameters and the manufacturing quality features collected by sensors. In addition, it is an urgent need to develop a real-time system that can automatically perceive abnormal parameters, and decide on how to optimize these parameters, in the light of the manufacturing state. The product quality prediction model can be embedded into the system to manufacture high-quality products.

## References

[1] Liukkonen, M., Tsai, T.-N. (2016). Toward decentralized intelligence in manufacturing: Recent trends in automatic identification of things, *International Journal of Advanced Manufacturing Technology*, Vol. 87, No. 9, 2509-2531, doi: 10.1007/s00170-016-8628-y.

[2] Nazifa, T.H., Ramachandran, K.K. (2019). Information sharing in supply chain management: A case study between the cooperative partners in manufacturing industry, *Journal of System and Management Sciences*, Vol. 9, No. 1, 19-47, doi: 10.33168/JSMS.2019.0102.

[3] Li, B.-H., Hou, B.-C., Yu, W.-T., Lu, X.-B., Yang, C.-W. (2017). Applications of artificial intelligence in intelligent manufacturing: A review, *Frontiers of Information Technology & Electronic Engineering*, Vol. 18, No. 1, 86-96, doi: 10.1631/FITEE.1601885.

[4] Istokovic, D., Perinic, M., Vlatkovic, M., Brezocnik, M. (2020). Minimizing total production cost in a hybrid flow shop: A simulation-optimization approach, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 559-570, doi: 10.2507/IJSIMM19-4-525.

[5] Shafiq, S.I., Sanin, C., Szczerbicki, E., Toro, C. (2017). Towards an experience based collective computational intelligence for manufacturing, *Future Generation Computer Systems*, Vol. 66, 89-99, doi: 10.1016/j.future.2016.04.022.

[6] Wang, C., Yang, B., Wang, H.Q. (2020). Multi-objective master production schedule for balanced production of manufacturers, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 678-688, doi: 10.2507/IJSIMM19-4-CO17.

[7] Liang, Q. (2020). Production logistics management of industrial enterprises based on wavelet neural network, *Journal Européen des Systèmes Automatisés*, Vol. 53, No. 4, 581-588, doi: 10.18280/jesa.530418.
[8] Stasiak-Betlejewska, R. (2015). Construction product quality improvement with applying production problems analysis, *Manufacturing Technology*, Vol. 15, No. 5, 756-761, doi: 10.21062/ujep/x.2015/a/1213-2489/MT/15/5/756.
[9] Zhao, Z., Yuan, Q. (2022). Integrated multi-objective optimization of predictive maintenance and production scheduling: perspective from lead time constraints, *Journal of Intelligent Management Decision*, Vol. 1, No. 1, 67-77, doi: 10.56578/jimd010108.
[10] Ben-Ammar, O., Bettayeb, B., Dolgui, A. (2020). Integrated production planning and quality control for linear production systems under uncertainties of cycle time and finished product quality, *International Journal of Production Research*, Vol. 58, No. 4, 1144-1160, doi: 10.1080/00207543.2019.1613580.
[11] Amar, H., Ghodbane, H., Amir, M., Zidane, M.A., Hamouda, C., Rouane, A. (2020). Microstrip sensor for product quality monitoring, *Journal of Computational Electronics*, Vol. 19, 1329-1336, doi: 10.1007/s10825-020-01517-2.
[12] Schiermeyer, A. (2020). Optimizing product quality in molecular farming, *Current Opinion in Biotechnology*, Vol. 61, 15-20, doi: 10.1016/j.copbio.2019.08.012.
[13] Poth, A., Riel, A. (2020). Quality requirements elicitation by ideation of product quality risks with design thinking, In: *Proceedings of 2020 IEEE 28th International Requirements Engineering Conference (RE),* Zurich, Switzerland, 238-249, doi: 10.1109/RE48521.2020.00034.
[14] Hou, Y., Cao, Z.J., Yang, S.L. (2019). Cloud intelligent logistics service selection based on combinatorial optimization algorithm, *Journal Européen des Systèmes Automatisés*, Vol. 52, No. 1, 73-78, doi: 10.18280/jesa.520110.
[15] Xu, M., Cai, H., Liang, S. (2015). Big data and industrial ecology, *Journal of Industrial Ecology*, Vol. 19, No. 2, 205-210, doi: 10.1111/jiec.12241.
[16] Chen, D. (2020). Multiple linear regression of multi-class images in devices of internet of things, *Traitement du Signal*, Vol. 37, No. 6, 965-973, doi: 10.18280/ts.370609.
[17] Wang, Y., Jan, M.N., Chu, S., Zhu, Y. (2020). Use of big data tools and industrial internet of things: An overview, *Scientific Programming*, Vol. 2020, Article ID 8810634, doi: 10.1155/2020/8810634.
[18] Song, Y.-J., Lee, J.-K. (2020). A blockchain and internet of things based architecture design for energy transaction, *Journal of System and Management Sciences*, Vol. 10, No. 2, 122-140.
[19] Yin, B., Wei, X., Wang, J., Xiong, N., Gu, K. (2019). An industrial dynamic skyline based similarity joins for multi-dimensional big data applications, *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 4, 2520-2532, doi: 10.1109/TII.2019.2933534.
[20] Gupta, A.K., Chakraborty, C., Gupta, B. (2019). Monitoring of epileptical patients using cloud-enabled health-IoT system, *Traitement du Signal*, Vol. 36, No. 5, 425-431, doi: 10.18280/ts.360507.
[21] Lu, J. (2020). Industrial pollution governance efficiency and big data environmental controlling measures: A case study on Jiangsu province, China, *Nature Environment and Pollution Technology*, Vol. 19, No. 4, 1743-1748, doi: 10.46488/NEPT.2020.v19i04.046.
[22] Zhang, X., Ge, Z. (2020). Local parameter optimization of LSSVM for industrial soft sensing with big data and cloud implementation, *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 5, 2917-2928, doi: 10.1109/TII.2019.2900479.
[23] Rosin, F., Forget, P., Lamouri, S., Pellerin, R. (2021). Impact of Industry 4.0 on decision-making in an operational context, *Advances in Production Engineering & Management*, Vol. 16, No. 4, 500-514, doi: 10.14743/apem2021.4.416.
[24] Nowakowska, M., Pajecki, M. (2021). Latent class analysis for identification of occupational accident casualty profiles in the selected Polish manufacturing sector, *Advances in Production Engineering & Management,* Vol. 16, No. 4, 485-499, doi: 10.14743/apem2021.4.415.
[25] Ren, L., Meng, Z., Wang, X., Zhang, L., Yang, L.T. (2020). A data-driven approach of product quality prediction for complex production systems, *IEEE Transactions on Industrial Informatics*, Vol. 17, No. 9, 6457-6465, doi: 10.1109/TII.2020.3001054.
[26] Chu, F., Peng, C., Jia, R., Chen, T., Lu, N. (2021). Online prediction method of batch process product quality based on multi-scale kernel JYMKPLS transfer model, *CIESC Journal*, Vol. 72, No. 4, 2178-2189, doi: 10.11949/0438-1157.20200995.
[27] Lughofer, E., Zavoianu, A.-C., Pollak, R., Pratama, M., Meyer-Heye, P., Zörrer, H., Eitzinger, C., Radauer, T. (2019). Autonomous supervision and optimization of product quality in a multi-stage manufacturing process based on self-adaptive prediction models, *Journal of Process Control*, Vol. 76, 27-45, doi: 10.1016/j.jprocont.2019.02.005.
[28] Hao, L., Bian, L., Gebraeel, N., Shi, J. (2016). Residual life prediction of multistage manufacturing processes with interaction between tool wear and product quality degradation, *IEEE Transactions on Automation Science and Engineering*, Vol. 14, No. 2, 1211-1224, doi: 10.1109/TASE.2015.2513208.
[29] Zheng, R., Pan, F. (2017). Multi-phase support vector regression soft sensor for online product quality prediction in glutamate fermentation process, *American Journal of Biochemistry and Biotechnology*, Vol. 13, No. 2, 90-98, doi: 10.3844/ajbbsp.2017.90.98.
[30] Melhem, M., Ananou, B., Ouladsine, M., Pinaton, J. (2016). Regularized regression models to predict the product quality in multistep manufacturing, In*: Proceedings of 2016 5th International Conference on Systems and Control (ICSC),* Marrakesh, Morocco, 31-36, doi: 10.1109/ICoSC.2016.7507067.