Acta Pharmaceutica Sinica B 2025;15(8):4193−4209

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com

ORIGINAL ARTICLE

AI-integrated IQPD framework of quality prediction and diagnostics in small-sample multi-unit pharmaceutical manufacturing: Advancing from experience-driven to data-driven manufacturing

Kaiyi Wang[a,b], Xinhai Chen[a,b], Nan Li[a,b], Huimin Feng[a,b], Xiaoyi Liu[a,b], Yifei Wang[c], Yanfei Wu[a,b], Yufeng Guo[a,b], Shuoshuo Xu[a,b], Lu Yao[d], Zhaohua Zhang[d], Jun Jia[d], Zhishu Tang[a,*], Zhisheng Wu[a,b,*]

[a]*School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 100029, China*
[b]*Pharmaceutical Engineering and New Drug Development of TCM of Ministry of Education, Beijing 100102, China*
[c]*Fangshan Hospital, Beijing University of Chinese Medicine, Beijing 102499, China*
[d]*Beijing Tongrentang Co., Ltd., Beijing 100062, China*

Received 26 November 2024; received in revised form 30 May 2025; accepted 1 June 2025

KEY WORDS

Smart manufacturing;
Artificial intelligence;
Intelligent quality prediction and diagnostics;
Small-sample multi-unit manufacturing;
Data-driven manufacturing;

Abstract The pharmaceutical industry faces challenges in quality digitization for complex multi-stage processes, especially in small-sample systems. Here, an intelligent quality prediction and diagnostic (IQPD) framework was developed and applied to Tong Ren Tang's Niuhuang Qingxin Pills, utilizing four years of data collected from four production units, covering the entire process from raw materials to finished products. In this framework, a novel path-enhanced double ensemble quality prediction model (PeDGAT) is proposed, which combines a graph attention network and path information to encode inter-unit long-range and sequential dependencies. Additionally, the double ensemble strategy enhances model stability in small samples. Compared to global traditional models, PeDGAT achieves state-of-the-art results, with an average improvement of 13.18% and 87.67% in prediction accuracy and stability on three indicators. Additionally, a more in-depth diagnostic model leveraging grey correlation analysis and expert knowledge reduces reliance on large

*Corresponding authors.
E-mail addresses: tzs6565@126.com (Zhishu Tang), wzs@bucm.edu.cn (Zhisheng Wu).

Peer review under the responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2025.06.001

Real-world Tong Ren Tang's Niuhuang Qingxin Pills

samples, offering a panoramic view of attribute relationships across units and improving process transparency. Finally, the IQPD framework integrates into a Human−Cyber−Physical system, enabling faster decision-making and real-time quality adjustments for Tong Ren Tang's Niuhuang Qingxin Pills, a product with annual sales exceeding 100 million CNY. This facilitates the transition from experience-driven to data-driven manufacturing.

1. Introduction

The global manufacturing industry is rapidly transitioning towards digitization and intelligent systems[1]. Key initiatives like Germany "Industry 4.0" and the United States "Industrial Internet" aim to enhance industrial competitiveness through advanced intelligent manufacturing technologies[2]. A crucial objective is intelligent quality prediction and diagnostics[3-7], which reduce resource waste by enabling real-time monitoring and process optimization[8,9]. This is especially crucial in the pharmaceutical industry, where ensuring product safety and efficacy is paramount.

Since China joined the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), digital and intelligent production processes for Chinese Materia Medica (CMM) have received significant attention[10-12]. The primary goal is to develop intelligent and accurate quality prediction and diagnostic models to enhance product consistency and ensure compliance with stringent quality standards. However, the complexity of multi-unit manufacturing processes (MMPs) and the low level of digitalization in CMM enterprises pose significant challenges[13,14]. In MMPs, each unit's output serves as the input for the next, leading to cumulative variations in product quality. Therefore, precise quality prediction is essential for the early detection and resolution of issues.

Deep learning methods have been developed for end-to-end quality prediction without complicated feature engineering[15-18]. Current research on quality prediction primarily focuses on single manufacturing processes, neglecting the dependency relationships among multiple processes[19-22]. Recent deep learning models for multi-unit quality prediction have emerged, but they typically rely on large amounts of historical data. For example, Liu et al.[23] introduced a novel quality embedding model pool (QTD) framework and a bidirectional serial-parallel LSTM (Bi-SP-LSTM) model, capable of leveraging machine dependencies to improve quality prediction. However, limited digitization in some manufacturing enterprises restricts the volume of collectible data. Small sample sizes result in overfitting and poor generalization, reducing predictive accuracy in multi-unit quality models[24-26]. Consequently, small sample size is a key challenge currently limiting the development of multi-unit quality prediction models.

In addition to prediction, effective anomaly detection is critical for maintaining quality consistency. A quality anomaly diagnosis model can enable production staff to shift from reactive measures to proactive, preventive actions, minimizing resource wastage and improving overall efficiency. However, the traditional production process and reliance on manual experience in the MMPs of CMM big honeyed pills complicate the early detection and correction of such issues[27]. Therefore, developing a quality anomaly diagnostic model for CMM big honeyed pill production is crucial for quickly identifying the root causes of anomalies and enhancing the level of manufacturing digitization.

Currently, diagnostic models for manufacturing processes primarily rely on statistical process control (SPC) and machine learning-based methods[28-30]. While these methods effectively detect quality abnormalities, they are highly dependent on large datasets. In complex products with multiple features and multi-stage processes, insufficient data often results in reduced model reliability[31]. To address these challenges, Bayesian networks[32,33] have emerged as a promising solution for quality diagnostics in MMPs. These graphical models effectively manage uncertainty and model probabilistic dependencies between units, enabling earlier detection and more precise corrective actions. However, for the production of big honeyed pills with small sample sizes across multiple units, it is even more difficult to develop accurate diagnostic models.

Here, a novel multi-unit intelligent quality prediction and diagnostic framework (IQPD) was developed to address the challenges posed by small sample sizes in MMPs. Within this framework, the quality prediction model employs an advanced path-enhanced double ensemble model named PeDGAT. This model introduced graph attention network (GAT) to capture complex dependencies in graph-structured processes. However, as GAT typically aggregates information only from first-order neighboring nodes within a single layer, Bi-LSTM was employed to encode the long-range and sequential dependencies between different process units. The resulting path information was then integrated with the adjacency matrix and incorporated into the model for enhanced predictive accuracy. At the same time, a Monte Carlo (MC) resampling-based double ensemble strategy was implemented to enhance the model's generalization for small sample datasets. Additionally, the small sample multi-unit quality anomaly diagnosis model combined grey correlation analysis with expert knowledge to enhance the accuracy and effectiveness of quality diagnosis in the IQPD framework. Last, a Human−Cyber−Physical (HCP) system based on the IQPD framework was established and integrated into the big honeyed pills production process, enabling monitoring and seamless human−machine interaction to ensure quality consistency and adaptive decision-making. This work shifts CMM production from an experience-driven to a data-driven model, enhancing efficiency, traceability, and intelligence in pharmaceutical manufacturing.

2. Materials and methods

2.1. Samples corresponding to the one-to-one process of the manufacturing process

Samples of Tong Ren Tang's Niuhuang Qingxin Pills, with an annual sales revenue exceeding 100 million CNY, were collected

from Beijing Tongren Tang Co., Ltd. (Beijing, China), one of the largest CMM manufacturers in the country[27]. These samples were sourced from four production units, including the honey refining, mixing, He-tuo, and pill making processes (Supporting Information 1.1). Raw honey was first refined into processed honey, which was then further processed into He-tuo honey. The He-tuo honey was combined with herbal powders to form He-tuo pill blocks through the He-tuo process. The final product, big honeyed pills, was produced through the pill making process. In the He-tuo process, herbal powders and He-tuo honey were thoroughly mixed to create a pill block that is soft, moderately firm, and suitably plastic. The ideal pills should exhibit appropriate adhesion without sticking to the cutting blade during extrusion, possess elasticity and cohesion for a smooth surface during rolling, and have moderate hardness for easy chewing. Hence, texture attributes are critical factors influencing the forming rate of honeyed pills.

The study included 4 batches of raw honey and 28 batches each of refined honey, He-tuo honey, herbal powders, He-tuo pill block, and big honeyed pill. In total, 59 physical and chemical quality attributes were measured across six sample types from the four production units (Supporting Information 1.2).

2.2. *Problem formulation*

This study addresses quality control challenges in pharmaceutical manufacturing involving multi-unit processes under small-sample scenarios, focusing on two critical tasks: (1) quality prediction in multi-stage production processes, and (2) anomaly traceability and diagnosis of final products. The methodological framework is validated through an industrial case study involving the production of Tong Ren Tang's Niuhuang Qingxin Pills, a multi-stage pharmaceutical manufacturing process.

The production process comprises N interconnected process units (*e.g.*, honey refining process, mixing process, He-tuo process, pill making process). Each unit i (i = 1, 2, …, N-1) generates a feature matrix $\boldsymbol{X}_i \in \mathbb{R}^{u\times F_i}$, where u is the number of samples, F_i is the feature dimension of unit i. The N-th unit represents the final product, with its quality metric $y \in \mathbb{R}^u$ as the prediction target.

Inputs:

1. Feature matrix $\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_{N-1}\}$, *i.e.* $\boldsymbol{X} \in \mathbb{R}^{u\times F}$.
2. Targets: $y \in \mathbb{R}^u$
3. Process topology: A graph construction $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, where:

1) Nodes $\mathscr{V}$: Represent process units $\{1, 2, \ldots, N-1\}$.
2) Edges $\mathscr{E}$: Define dependencies between units (*e.g.*, material flows). For example, an edge $(a, b) \in \mathscr{E}$ indicates that unit a directly precedes unit b.
3) Adjacency Matrix: An adjacency matrix $\boldsymbol{A} \in \{0,1\}^{(N-1)\times(N-1)}$, where $\boldsymbol{A}_{a,b} = 1$ if $(a, b) \in \mathscr{E}$, else 0.

Output:

Predicted quality: $\hat{y} \in \mathbb{R}^u$ for the final product.

2.3. *PeDGAT: A path-encoded double-ensemble GAT model*

This section introduces a predictive modeling method called PeDGAT, which integrates an autoencoder, a GAT, a bidirectional LSTM-based path encoder, and a MC resampling strategy for double-ensemble training. As illustrated in Fig. 1, the primary idea is to first construct a quality embedding model[23], mapping each process unit's features into a unified feature space. Next, a static graph is established based on predefined inter-process relationships, followed by path encoding using a bidirectional LSTM. Finally, multiple GAT sub-models are trained *via* MC subsampling, and their predictions are aggregated to yield a robust final output. Pseudocode for PeDGAT is shown in Algorithm 1, which consists of the following five stages:

Stage 1 Quality embedding model

Given the raw quality data matrix $\boldsymbol{X} \in \mathbb{R}^{u\times F}$, each feature is standardized. An autoencoder (Supporting Information 1.3, Supporting Information Fig. S1) f_{AE} is applied to the feature matrix $\boldsymbol{X}$ of each process unit to encode it, projecting it into a feature space $\mathbb{R}^{d_{\mathrm{emb}}}$. The resulting embedded representation is given by Eq. (1):

$$\boldsymbol{e} = f_{\mathrm{AE}}(\boldsymbol{X}) \in \mathbb{R}^{u\times d_{\mathrm{emb}}} \tag{1}$$

where f_{AE} represents the trained encoder. The embedded features e serve as inputs to subsequent stages. This pretraining decouples feature extraction from the subsequent prediction task, thereby simplifying the overall training process and enhancing training efficiency.

Stage 2 Graph construction and data preparation

The static graph topology is constructed based on predefined process connectivity relationships among manufacturing units, where each node represents a process unit. Specifically, directed edges are established according to material flow sequences: raw honey → refined honey (node 0 → 1), refined honey → He-tuo honey (node 1 → 2), He-tuo honey → He-tuo pill block (node 2 → 4), and herbal powders → He-tuo pill block (node 3 → 4). Each node is characterized by its latent feature vector derived from the embedding matrix e. This topological configuration reveals two distinct material transformation pathways: Path 1 (node 0 → 1 → 2 → 4) characterizing the excipient honey refinement stream from auxiliary materials to He-tuo pill block formation, and Path 2 (node 3 → 4) describing the integration of herbal raw materials into the He-tuo pill block composite. All such path connections are systematically enumerated to generate a comprehensive edge list. Subsequently, this edge list was then converted into an adjacency representation, yielding an edge_index structure. For a specific path $pk = [m_1, m_2, \ldots, m_{|pk|}]$, representing the ordered sequence of nodes from m_1 to $m_{|pk|}$. The feature representation of the nodes along the path pk is expressed as Eq. (2):

$$\mathrm{B}(p_k) = \begin{bmatrix} e[m_1,:] \\ e[m_2,:] \\ \vdots \\ e[m_{|p_k|},:] \end{bmatrix} \in \mathbb{R}^{|p_k|\times d_{\mathrm{emb}}} \tag{2}$$

For each sample instance, a graph-structured data object was constructed containing three principal components: (1) dimensionally reduced node embeddings generated through the autoencoder, (2) the corresponding edge_index, and (3) the associated target quality metric y.

Stage 3 Define the GAT model with the path encoder

Motivated by the inherent limitations of GAT architectures in effectively modeling long-range and sequential node

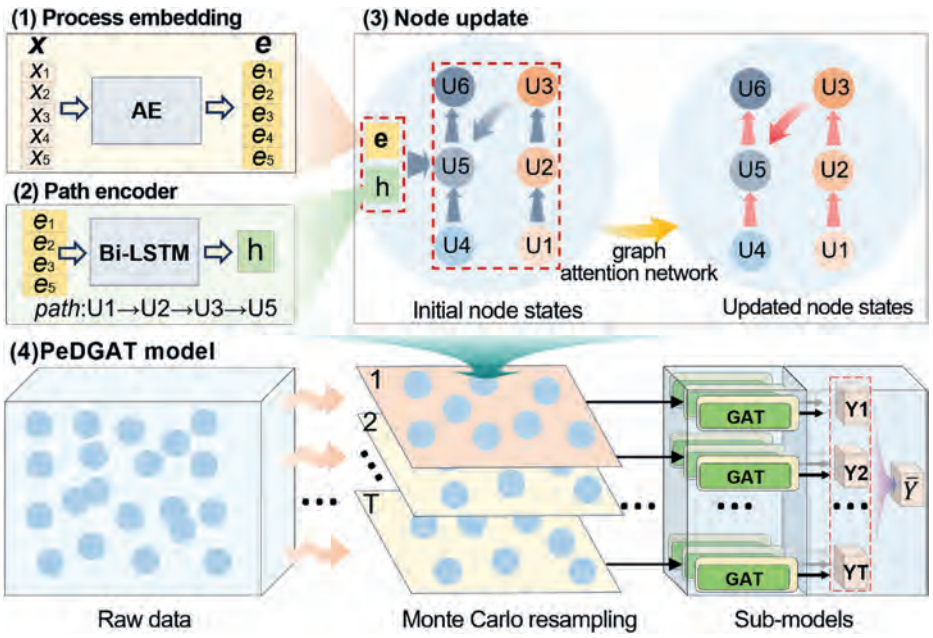

Figure 1 The principle of the PeDGAT model.

dependencies, a bidirectional LSTM (Bi-LSTM) is employed to extract its sequential dependencies features. Each path representation $B(p_k)$ is processed by the Bi-LSTM to generate hidden vectors $\tilde{h}_k$. As shown in Eq. (3):

$$\tilde{h}_k = \mathrm{Bi-LSTM}(B(p_k)) \in \mathbb{R}^{d_p} \tag{3}$$

where d_p is the hidden layer dimension of the Bi-LSTM.

The path-enhanced features $\tilde{h}_k$ are then propagated through GAT layers, where a multi-head attention mechanism bridges the graph structure. For example, for nodes such as the "He-tuo pill block" that are connected to both "He-tuo honey" and "herb powders", the GAT dynamically assigns weights to these connected nodes based on their relative importance, while simultaneously aggregating features according to the predefined topology edges. The node representation update is formalized as Eq. (4):

$$h'_a = \mathrm{GAT}(\tilde{h}_k, \mathscr{E}; \theta) \tag{4}$$

where θ represents the GAT layer parameter.

Stage 4 MC resampling model training

To ensure robustness in model performance, the MC resampling procedure is performed to generate T subsample sets for training an ensemble of GAT sub-models (Supporting Information 1.4, Supporting Information Fig. S2)[34-37]. Denote by f_t the prediction function of the t-th GAT sub-model. As shown in Eq. (5):

$$\hat{Y}_t = f_t(h'_a) \tag{5}$$

Early stopping is applied during training to mitigate overfitting.

Stage 5 Final prediction and evaluation

The final output results of T sub-models are averaged to obtain the final predicted quality value:

$$\hat{Y} = \frac{1}{T}\sum_{t=1}^{T} \hat{Y}_t \tag{6}$$

Algorithm 1: PeDGAT prediction model

Input: Raw data matrix $\mathbf{X} \in \mathbb{R}^{u \times F}$, targets y, graph structure
Output: Prediction results and evaluation metrics
//Stage 1: Quality embedding model
Initialize StandardScaler, fit_transform on X
Define autoencoder:
  Encoder: Linear (m→a) → ReLU → Linear (a→b)
  Decoder: Linear (b→c) → ReLU → Linear (c→m)
Train autoencoder:
  For epoch = 1 to 1000 do
    Forward pass: X̂ = Decoder (Encoder(X))
    Compute MSE loss: L = ‖X - X̂‖^2
    Backpropagate and update weights
  end
Extract compressed features: e = Encoder(X)
//Stage 2: Graph construction and data preparation
Define static graph topology:
  edges = [
    [0, 1], //Node 0 → Node 1
    [1, 2], //Node 1 → Node 2
    [2, 4], //Node 2 → Node 4
    [3, 4] //Node 3 → Node 4
  ]
Convert to sparse adjacency representation:
  edge_index = transpose (edges) //Shape becomes [2, num_edges]
  edge_index = make_contiguous (edge_index)
Construct graph data:
  for each sample i ∈ [1,n] do
    Create a data object with:
      x = e [i] (processed features)
      edge_index = the predefined edge_index mentioned above

(*continued*)

Algorithm 1: PeDGAT prediction model

```
            y = corresponding target
        end
10  //Stage 3: Define GAT model with path encoder
11      Path encoder: Bidirectional LSTM
            Input dimension: num_features
            Hidden dimension: h_dim
12      GAT layers:
            conv1: 8-Head attention layer
                (input_dim = 2h_dim → output_dim = 8 × 8)
            conv2: Single-head attention layer
                (input_dim = 64 → output_dim = 1)
13  //Stage 4: MC resampling model training
14  Implement MC resample:
15      Initialize the MC regressor with 100 estimators
        For each estimator do
            Train on 80% subsampled data:
                for epoch = 1 to 300 do
                    Forward pass-through GAT with path encoder
                        model
                    Compute MSE loss
                    Update weights with early stopping
                end
        end
16  //Stage 5: Final prediction and evaluation
17      Evaluate the model:
            For 20 independent runs do
                Split data into train/test samples
                Aggregate predictions from all estimators
                Calculate RMSE, MAE, R and MSE metrics
        end
```

All samples were randomly split into 80% for training and 20% for prediction. To validate the method's effectiveness, four commonly used metrics were employed: mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and correlation coefficient (R). The final experimental results are reported as the mean and standard deviation derived from 20 independent experiments, ensuring robust and reliable performance evaluation.

To evaluate the performance of the proposed model, five models were compared: ridge regression (RR), support vector regression (SVR), multi-layer perceptron (MLP), convolutional neural networks (CNN), and GAT. Besides, an early stopping strategy was implemented to avoid overfitting. The Adam optimizer was employed to optimize parameters for PeDGAT, with the training process running for 200 epochs and the number of attention heads set to 8. The optimal critical parameters for PeDGAT in predicting Chewiness, Hardness, and Adhesiveness were determined as follows: Dropout values of 0.2, 0.1, and 0.5, respectively; a Learning Rate of 0.01; and a patience setting of 80, 60, and 80 steps, respectively. The modeling environment details and parameters for the other methods are provided in Supporting Information 1.5 and Supporting Information Table S1.

All algorithms were executed using Python 3.8 and MATLAB 2022b in a computational environment powered by an AMD Ryzen 7 5800H with Radeon Graphics, a CPU clocked at 3.2 GHz, and 32 GB of memory.

2.4. Multi-unit IQPD framework for quality control in small-batch manufacturing

To address the challenges of quality control in MMPs with small samples, a multi-unit IQPD framework is proposed (Fig. 2). The framework operates in four main steps:

(1) Digital measurement of quality attributes

Firstly, the framework digitally measures the quality attributes of the manufacturing process, as shown in Fig. 2A and a1. In total, 59 physical and chemical quality attributes across 11 categories were digitally characterized.

(2) PeDGAT model

Secondly, it encodes the dependencies of the multi-unit process and establishes a double ensemble multi-unit quality prediction model—PeDGAT, illustrated in Fig. 2B. Following the method described in Section 2.3, PeDGAT is built to predict quality indices under limited sample conditions.

(3) Quality anomaly tracing and diagnosis

When the prediction results exceed the specified thresholds, indicating potential quality anomalies, the framework employs a quality anomaly tracing and diagnosis model based on grey relational analysis and expert knowledge, as depicted in Fig. 2C and c1.

1) Construction of Bayesian network: The comprehensive correlation between abnormal nodes and their upstream nodes is calculated using grey correlation analysis to establish a Bayesian network. This network is further refined by integrating expert knowledge. The conditional probabilities within the network are computed using historical data (Supporting Information 1.6, Supporting Information Fig. S3).
2) Probabilities of quality diagnosis: Formula (25) in Supporting Information 1.6 is employed to calculate the probabilities of all upstream nodes causing anomalies. These probabilities are then ranked to identify the key causes of anomalies.

(4) Quality tracing and diagnosis software

Additionally, quality tracing and diagnosis software has been developed (Fig. 2D), user-friendly interface and offers the following key features:

1) Anomaly tracing and visualization

When predicted quality indicators exceed allowable limits, the software initiates anomaly tracing and visually highlights potential issues within the Bayesian network.

2) Root cause analysis

By integrating historical data, grey relational analysis, and expert knowledge, the platform generates ranked lists of probable causes and produces intuitive diagnostic reports.

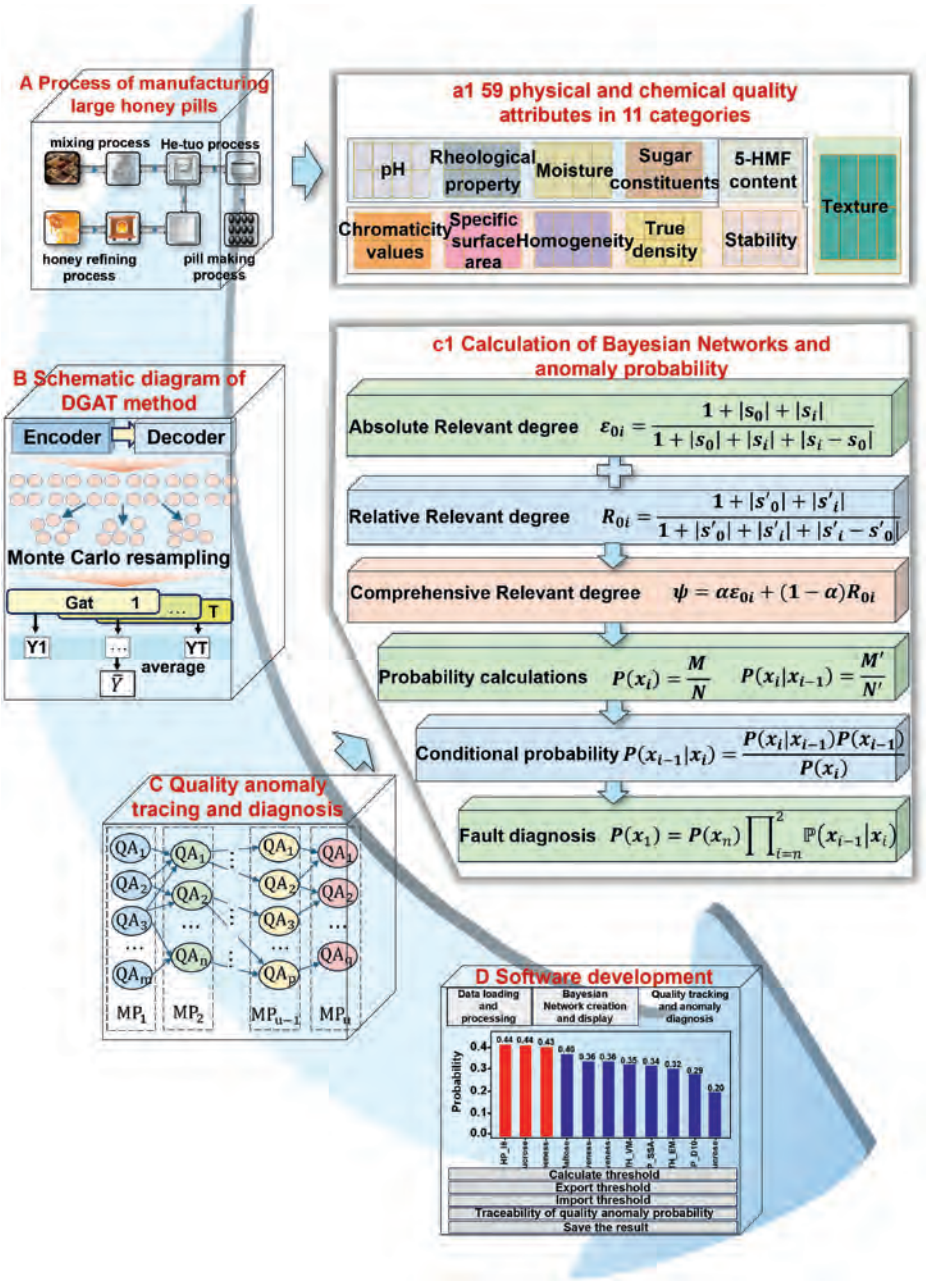

Figure 2 Multi-unit IQPD framework. (A) Manufacturing processes: (a1) Multi-source data as model inputs; (B) Schematic diagram of the PeDGAT model; (C) Quality anomaly traceability and diagnostic process: (c1) Bayesian network calculations and anomaly probability estimation; (D) Software development.

3) Decision support

The software provides actionable recommendations and best-practice guidelines to mitigate anomalies, helping practitioners rapidly resolve quality issues and make informed adjustments to the manufacturing process.

The quality tracing and diagnosis software can effectively trace and diagnose the causes of quality anomalies, providing essential guidance and visualization tools for process control.

3. Results

3.1. Overall quality diagnosis of multi-batch products via information fusion strategy

CMM product quality is critical for ensuring safe and effective clinical treatments. To evaluate quality consistency in Tong Ren Tang's Niuhuang Qingxin Pills production, the 53 physical and chemical attributes of samples from the honey refining, mixing, and He-tuo

processes (Fig. 3A) were first standardized to eliminate data scale differences (Fig. 3b1, b4, b7, and b10). Then, a multivariate statistical process control (MSPC) model based on information fusion assessed overall quality consistency across manufacturing stages.

When the MSPC strategy was implemented for refined honey, five samples were identified as outliers using Hotelling's T^2 and SPE charts (Fig. 3b2 and b3). Specifically, two samples (20010186, 20010257) were flagged by Hotelling's T^2 chart, and three samples (20010187, 20010188, 20010436) were identified by the SPE chart. For He-tuo honey, two samples (20010186, 20010188) exceeded the control limits on both charts (Fig. 3b5 and b6). In the case of herbal powders, three samples for herbs powders (20010256, 20010426, 20010439) were found to exceed the control limits (Fig. 3b8 and b9). For the He-tuo pill block, two samples (20010427, 20010438) surpassed the SPE control limits (Fig. 3b11 and b12).

These results highlight significant quality fluctuations and batch-to-batch differences during Tong Ren Tang's Niuhuang Qingxin Pills production, emphasizing the urgent need for an innovative quality prediction and diagnostic framework to identify and control potential quality issues, thereby enabling timely adjustments by production personnel.

3.2. Advanced IQPD framework for intelligent diagnosis in real-world pharmaceutical manufacturing with small sample sizes in MMPs

3.2.1. Parameter optimization of the PeDGAT model

In the PeDGAT method, a subset of samples is first selected from the original training set to establish multiple PeDGAT sub-models Fig. 4A. Sample size and epochs are key parameters. To optimize the training subset size, percentages ranging from 5% to 100% with 5% intervals are evaluated. Fig. 4b1 depicts the variation of the RMSE for chewiness with the percent of the training subset of DGAT (PeDGAT lacking path encoding). RMSE decreases sharply up to 20%, slows between 20% and 50%, and stabilizes with a minimum between 50% and 80%, before slightly increasing beyond 80%, indicating overfitting. An 80% subset is optimal for chewiness. Similarly, Fig. 4b2 shows RMSE for hardness stabilizes between 60% and 80%, and Fig. 4b3 indicates adhesiveness reaches its minimum at 80%. Thus, an 80% subset is recommended for all three indices for DGAT and PeDGAT.

Early stopping was utilized to determine the optimal number of epochs by monitoring the training loss, aiming to prevent overfitting. Fig. 4b4−b12 illustrates the loss trends for GAT, DGAT, and PeDGAT, respectively, across the three metrics. It can be observed that the fluctuations in loss trends for DGAT and PeDGAT are smaller than those for GAT, indicating that MC resampling enhances model stability. Furthermore, the significantly lower loss values for PeDGAT compared to DGAT and GAT indicate that the path encoding mechanism in PeDGAT effectively enhances model accuracy, further reducing error compared to GAT and DGAT. Using early stopping, the optimal epochs selected for the three metrics with the PeDGAT method were 167, 190, and 272, respectively.

3.2.2. Comparison of the prediction results

The effectiveness of PeDGAT in small-sample scenarios was systematically validated through 20 independent experimental trials. Under optimized parameter configurations, comprehensive performance comparisons between PeDGAT and five baseline models (RR, SVR, MLP, CNN, GAT) were conducted, focusing on three critical textural attributes of big honeyed pills: chewiness, hardness, and adhesiveness. To further investigate the contributions of key architectural components, an ablation study was implemented through two modified variants: DGAT and PeGAT (excluding MC resampling).

The superior performance of PeDGAT is systematically demonstrated in Table 1, where the lowest error metrics (MAE, RMSE, MSE) and highest R are consistently achieved across all three textural attributes. For chewiness, the PeDGAT model achieves an MAE of 0.2257, RMSE of 0.3056, MSE of 0.0934, and R of 0.9366. When compared to the suboptimal GAT baseline model (MAE = 0.2438, R = 0.8373), PeDGAT exhibited significant improvements, reducing MAE by 7.4% and enhancing the R by 11.9%. PeDGAT demonstrates similar performance advantages over traditional machine learning models. Compared to SVR and RR exhibiting MAE values of 0.2480−0.3120 and R values of 0.6452−0.7262, PeDGAT achieves MAE reductions of 9.0%−27.7% alongside R value improvements of 45.2%−29.0%. This performance advantage extends consistently to hardness and adhesiveness predictions, demonstrating that the integration of path encoding and MC resampling substantially improves predictive accuracy across multiple textural attributes.

Twenty independent trials were conducted to quantify prediction stability, with the standard deviation of each model's outputs calculated across all evaluation metrics. For example, the RMSE for hardness prediction is 0.0071 in PeDGAT, whereas MLP, CNN and GAT exhibit 8.6-, 11.8- and 19.4-fold higher (0.0610, 0.0838 and 0.1379). This difference in fluctuation amplitude persists across chewiness and adhesiveness predictions, validating the prediction stability, reliability and robustness of PeDGAT under small-sample conditions.

To systematically evaluate the contribution of individual components in PeDGAT, we conducted ablation studies comparing model performance under different architectural configurations. The experimental results reveal two critical findings: ① Removal of the path encoding module (DGAT configuration) significantly degrades prediction accuracy across all texture attributes. For example, this architectural modification increased the MAE from 0.2257 to 0.2732, representing a 17.4% performance deterioration (lower values indicate better performance) for chewiness. Concurrently, the R decreased from 0.9366 to 0.8942, reflecting reduced model explainability. This confirms the necessity of capturing long-range and sequential dependencies among manufacturing processes. ② Elimination of the MC resampling component (PeGAT configuration) substantially increased prediction instability. For example, in adhesiveness prediction, the MAE standard deviation surged from 0.0223 to 0.1336, a 6-fold increase in variance. The results show that the MC resampling strategy demonstrated significant improvements in model robustness and reliability under small-sample conditions. Overall, these results indicate that encoding dependencies between processes and employing a double ensemble approach significantly enhanced both small sample model accuracy and stability.

3.2.3. Quality anomaly tracing and diagnostic analysis for Tong Ren Tang's Niuhuang Qingxin Pills

To trace the causes of quality anomalies in the production of Tong Ren Tang's Niuhuang Qingxin Pills, a Bayesian network model was developed based on grey correlation analysis and expert knowledge, focusing on the four key manufacturing processes: honey refining, mixing, He-tuo, and pill making. The physical and chemical properties of raw materials, excipients, intermediates,

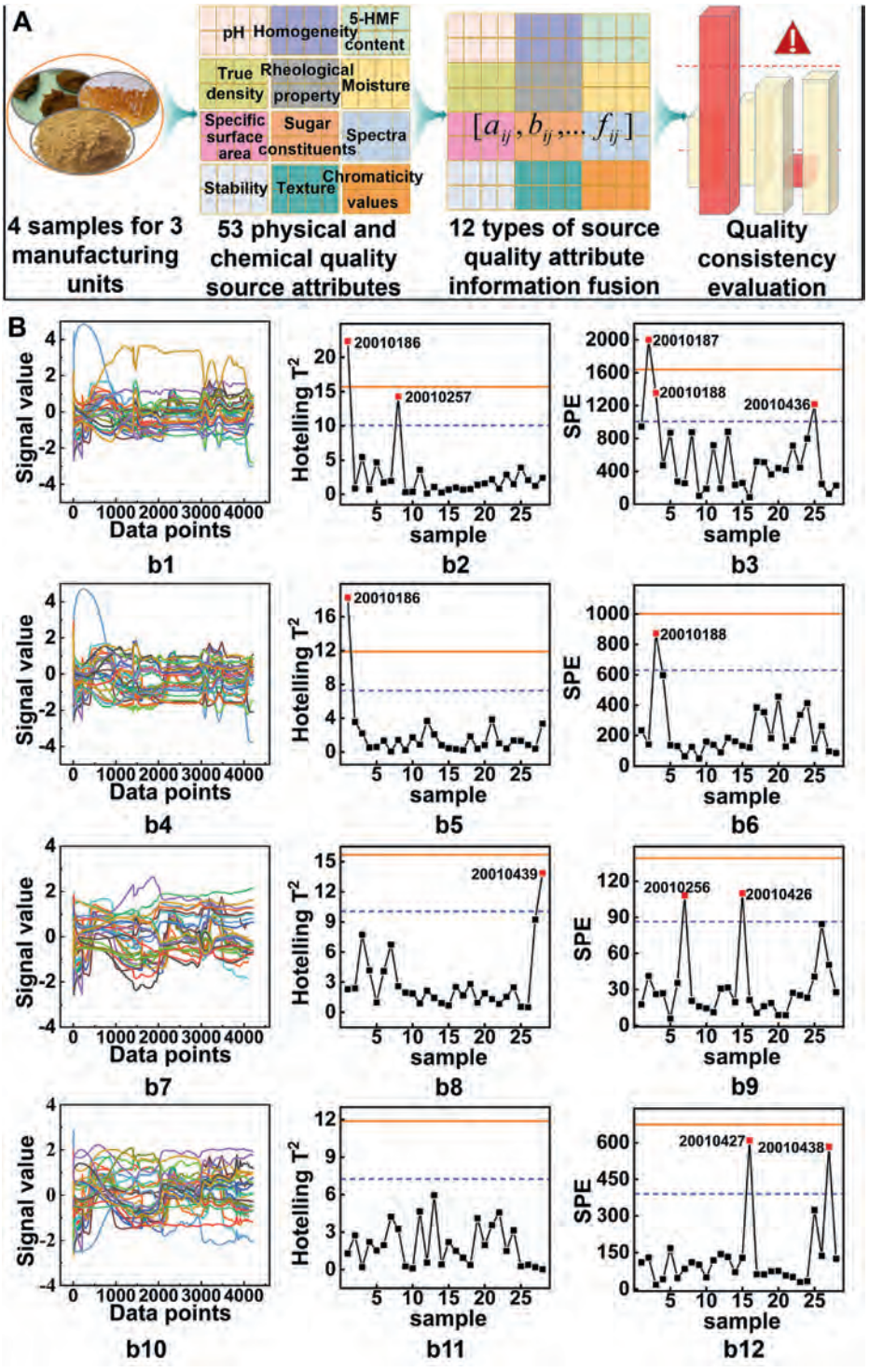

Figure 3 Standardized preprocessing and Hotelling's T^2 and SPE charts for the CMM dataset. Red squares indicate abnormal samples, blue dashed lines the 95% confidence limit, and orange solid lines the 99% confidence limit. (b1), (b4), (b7), and (b10) are standardized preprocessing charts for refined honey, He-tuo honey, mixed herbal powder, and He-tuo pill block samples, respectively. (b2), (b5), (b8), and (b11) are Hotelling's T^2 charts, while (b3), (b6), (b9), and (b12) are SPE charts for the information fusion results of the four sample types.

and final products were measured to construct a Bayesian network revealing the interactions affecting product quality, as depicted in Fig. 5.

Fig. 5A depicts the four key production processes and six critical materials, from raw honey to the final big honeyed pill. Fig. 5B−D, and F illustrate Bayesian networks for chewiness, hardness, and adhesiveness, where pink circles represent the final product's texture attributes. These are influenced by the He-tuo pill block's quality, which in turn depends on the herbal powders and He-tuo honey, ultimately tracing back to raw honey. Each network emphasizes potential key factors affecting the texture properties.

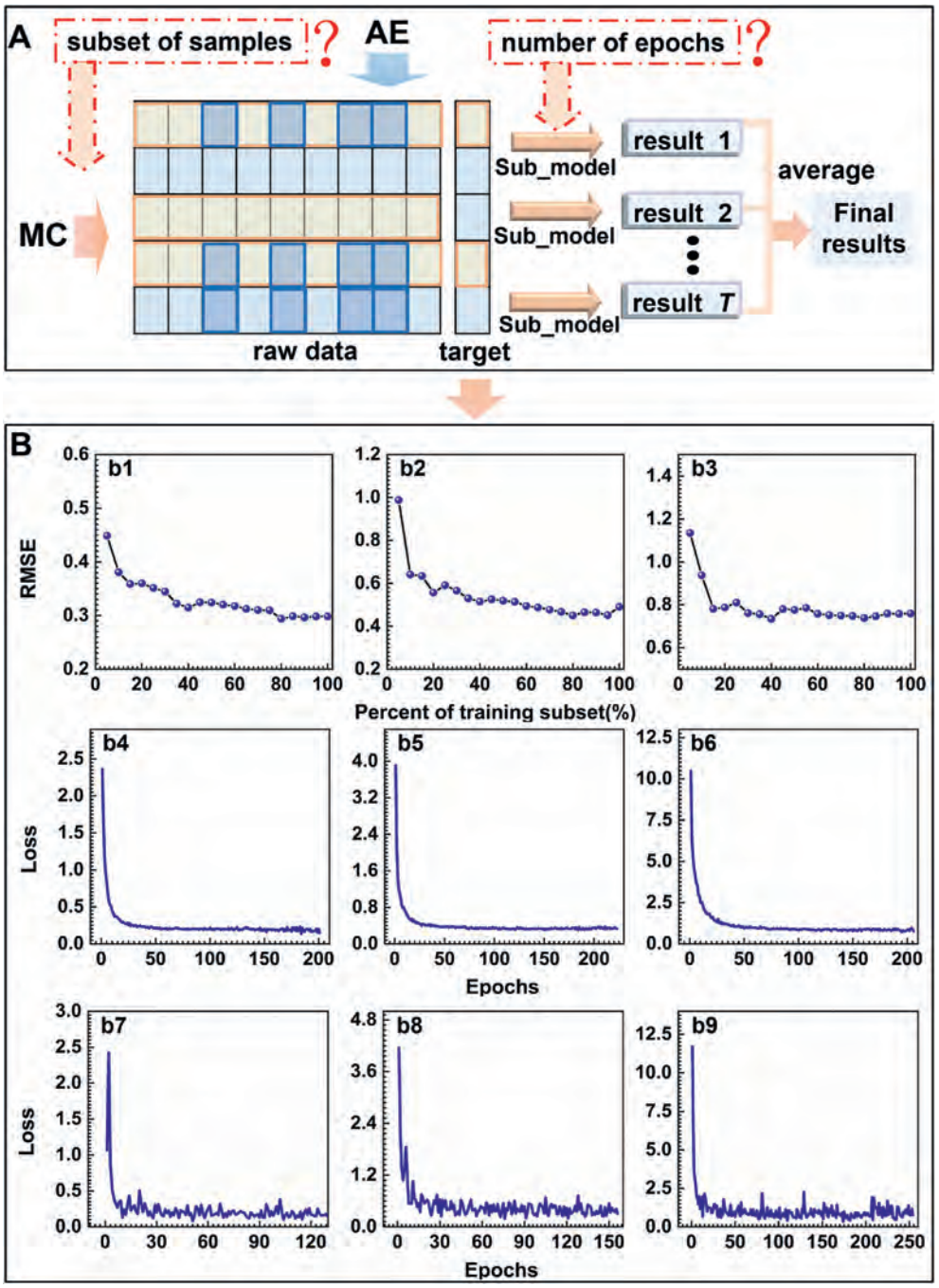

Figure 4 The parameter optimization charts for PeDGAT model. (A) The flowchart of the PeDGAT model. (b1)−(b3) the parameter optimization for the number of training subsets for the three targets: chewiness, hardness, and adhesiveness. (b4)−(b12): Epochs parameter optimization for GAT, DGAT and PeDGAT across the three targets.

In Fig. 5C, the $I\theta$ powder property of herbal powders shows the highest probability of affecting abnormal chewiness, indicating that the uniformity and consistency of the powder's particle likely play a key role in chewiness anomalies. Fig. 5E indicates that the hardness and chewiness of the He-tuo pill block are the most influential factors contributing to hardness anomalies, suggesting that process uniformity during He-tuo formation directly impacts the final product's texture. Similarly, Fig. 5G reveals that the sucrose content in He-tuo honey significantly influences adhesiveness anomalies, possibly due to its effect on the pill's stickiness and cohesion during mixing. In conclusion, to address anomalies in the big honeyed pill's texture attributes, engineers should focus on controlling temperatures during the honey refining to ensure optimal sucrose caramelization, as well as adjusting mixing times and uniformity in the He-tuo process to maintain consistent texture in the final product.

Table 1 Performance comparison of the models in three indicators.

Method	Chewiness				Hardness				Adhesiveness			
	MAE (Std[a])	RMSE (Std[a])	MSE (Std[a])	*R* (Std[a])	MAE (Std[a])	RMSE (Std[a])	MSE (Std[a])	*R* (Std[a])	MAE (Std[a])	RMSE (Std[a])	MSE (Std[a])	*R* (Std[a])
RR	0.3120 (0)	0.3435 (0)	0.1180 (0)	0.7262 (0)	0.4221 (0)	0.5191 (0)	0.2694 (0)	0.5620 (0)	0.6763 (0)	0.8102 (0)	0.6564 (0)	0.5944 (0)
SVR	0.2480 (0)	0.3288 (0)	0.1081 (0)	0.6452 (0)	0.3944 (0)	0.4897 (0)	0.2397 (0)	0.5290 (0)	0.6385 (0)	0.7844 (0)	0.6153 (0)	0.5545 (0)
MLP	0.2739 (0.0929)	0.3525 (0.0808)	0.1308 (0.0684)	0.8280 (0.0463)	0.3477 (0.0654)	0.4324 (0.0610)	0.1906 (0.0546)	0.7487 (0.1032)	0.6918 (0.0474)	0.7896 (0.0443)	0.6254 (0.0706)	0.6998 (0.1279)
CNN	0.2954 (0.0300)	0.3257 (0.0242)	0.1066 (0.0159)	0.7880 (0.0648)	0.4028 (0.0951)	0.4879 (0.0838)	0.2450 (0.0897)	0.7275 (0.0607)	0.7871 (0.0681)	0.8562 (0.0699)	0.7380 (0.1238)	0.7812 (0.0669)
GAT	0.2438 (0.0564)	0.3244 (0.0646)	0.1094 (0.0431)	0.8373 (0.0610)	0.3689 (0.1232)	0.4334 (0.1379)	0.2068 (0.1260)	0.7636 (0.0692)	0.6875 (0.1001)	0.7890 (0.0922)	0.6310 (0.1504)	0.7887 (0.1072)
DGAT (-Path encode)	0.2732 (0.0120)	0.3097 (0.0051)	0.0959 (0.0032)	0.8942 (0.0163)	0.3549 (0.0136)	0.4456 (0.0193)	0.1990 (0.0173)	0.7772 (0.0188)	0.6755 (0.0177)	0.7660 (0.0113)	0.5869 (0.0174)	0.8413 (0.0137)
PeGAT (-MC resample)	0.2528 (0.0677)	0.3106 (0.0690)	0.1012 (0.0485)	0.8464 (0.0577)	0.3098 (0.0694)	0.3444 (0.0676)	0.1232 (0.0457)	0.7739 (0.0498)	0.6375 (0.1336)	0.7170 (0.1291)	0.5307 (0.1767)	0.8068 (0.0843)
PeDGAT	0.2257 (0.0054)	0.3056 (0.0046)	0.0934 (0.0028)	0.9366 (0.0174)	0.2964 (0.0084)	0.3281 (0.0071)	0.1077 (0.0047)	0.8017 (0.0143)	0.6083 (0.0223)	0.7145 (0.0228)	0.5111 (0.0331)	0.9231 (0.0357)

Bold and underline fonts denote the best and the second-best methods, respectively.

[a]Standard deviations.

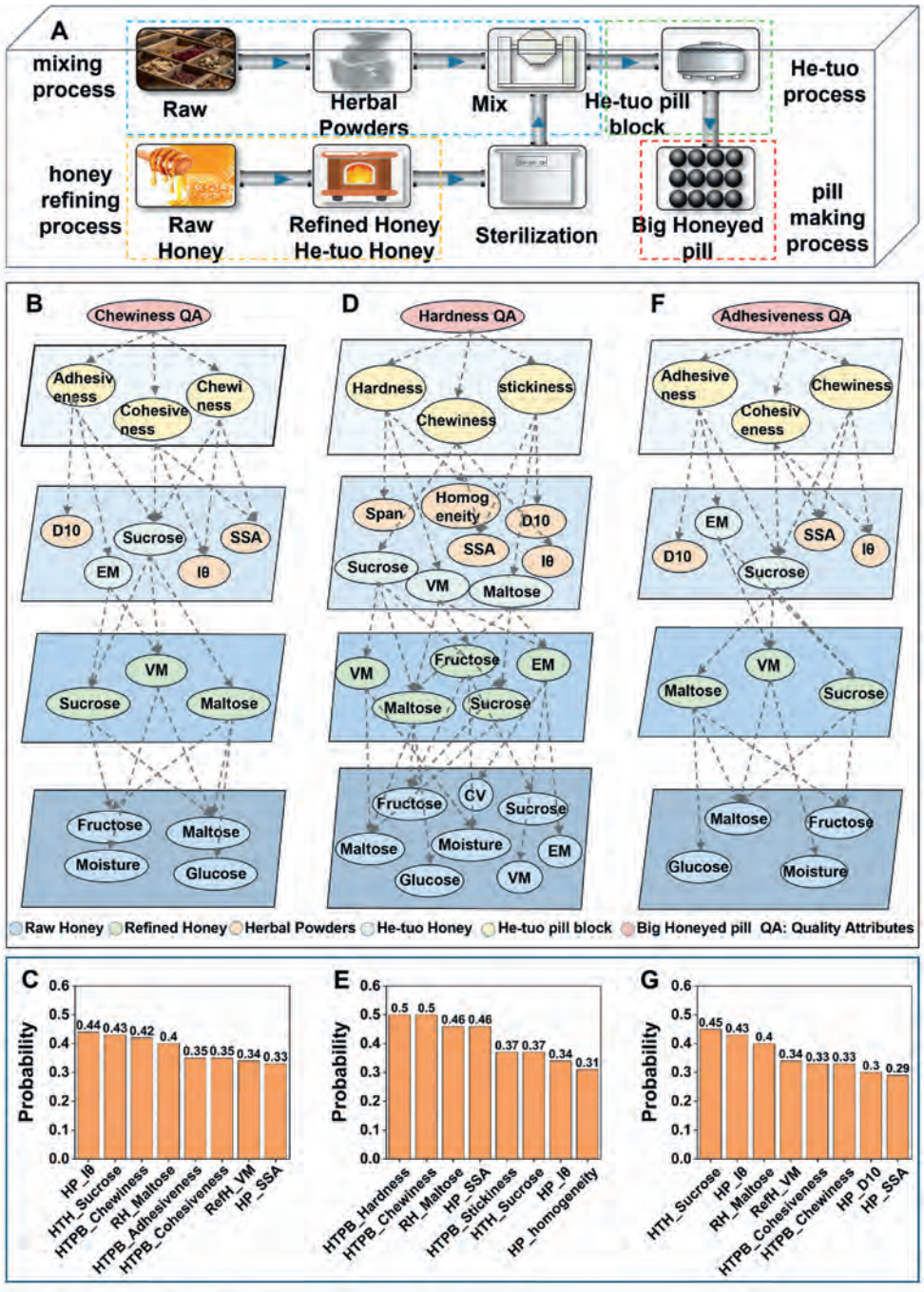

Figure 5 Tracing and diagnosis of quality abnormalities in Tong Ren Tang's Niuhuang Qingxin Pills. (A) Manufacturing process and sample source; (B), (D) and (F) Constructed Bayesian networks; (C), (E) and (G) Quality anomaly diagnosis results. Abbreviations: RH (Raw Honey), RefH (Refined Honey), HTH (He-tuo Honey), HP (Herbal Powders), HTPB (He-tuo Pill Block), and BHP (Big Honeyed Pill).

3.2.4. Development of quality anomaly tracing and diagnostic software

The software integrates data loading, Bayesian network visualization, and quality anomaly diagnosis to provide robust decision support in pharmaceutical manufacturing (Fig. 6A). As seen in Fig. 6B, users can easily import data, customize unit names, and analyze production data for different stages.

The Bayesian network module enables visualization of interconnections between units and incorporates expert knowledge to revise networks, adapting them to real-world manufacturing complexities. The network creation and display features are illustrated in Fig. 6C and D, providing a clear graphical interface for users to view and adjust the network.

The quality tracking and diagnosis module identifies potential causes of quality issues by calculating the probabilities of problematic production indicators based on predefined thresholds. These results are visually represented as bar charts, as shown in Fig. 6E, allowing for the straightforward interpretation of key risk factors in the production process.

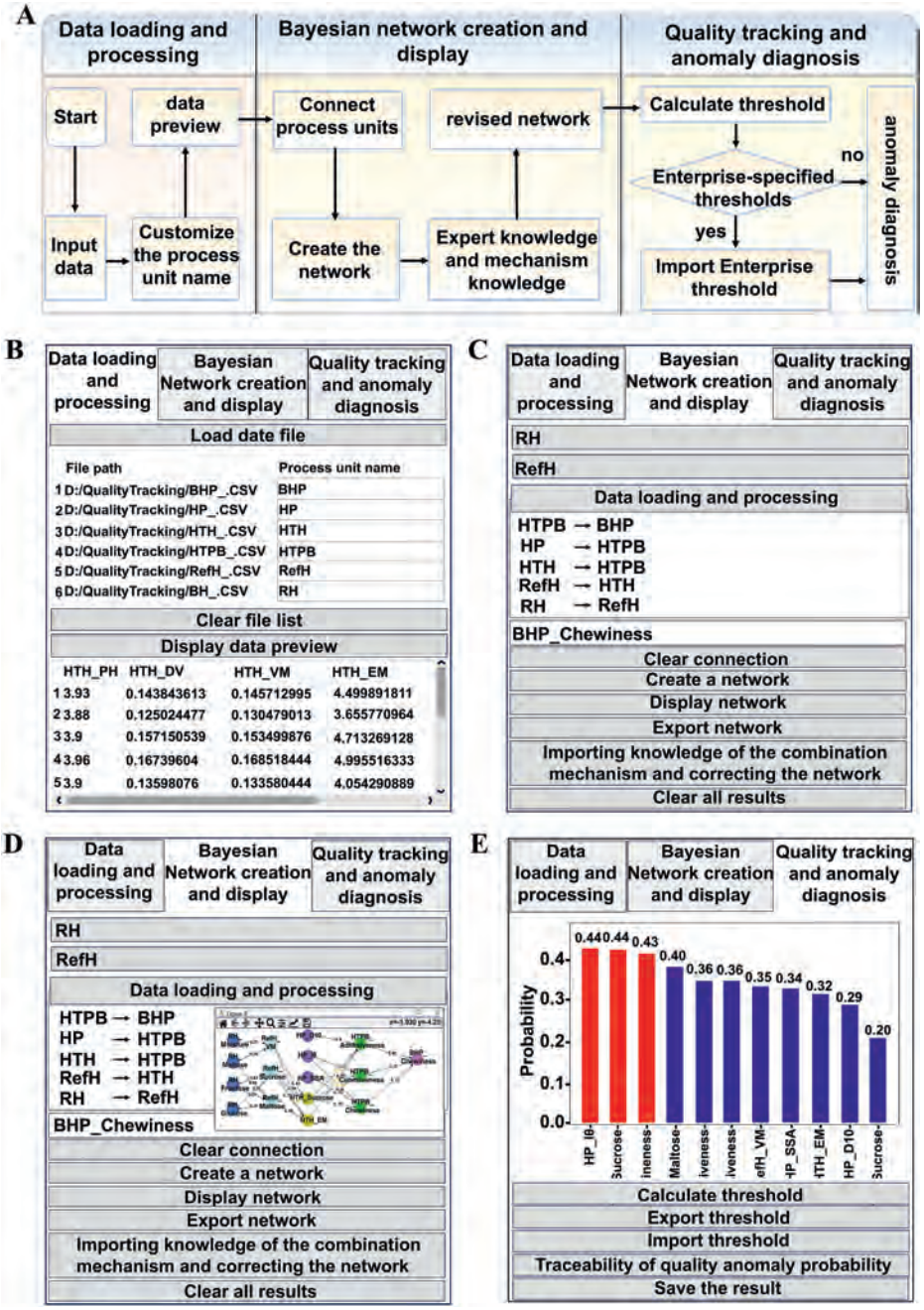

Figure 6 Quality anomaly traceability and diagnostic software. (A) Software operation workflow; (B)−(E) Software interface displays.

3.3. Digital multivariate process capability evaluation

3.3.1. Digital multivariate process capability evaluation for Tong Ren Tang's Niuhuang Qingxin Pills

To validate the diagnostic results, fuzzy mathematical theory, the analytic hierarchy process (AHP), and the criteria importance through intercriteria correlation (CRITIC) method were employed to calculate the multivariate process capability index (MPCI) of the mixing and honey refining processes, as shown in Fig. 7A and B. Fig. 7A outlines the 22 chemical attributes (*e.g.*, pH) and 37 physical attributes (*e.g.*, texture properties) used in the evaluation. Fig. 7B illustrates how weights were assigned using AHP-CRITIC. The calculation principles are detailed in Supporting Information 1.7 and Supporting Information Table S2. AHP-CRITIC provides a comprehensive weighting scheme that combines expert knowledge with statistical analysis. The comprehensive weighting scheme results are shown in Supporting Information Table S3−S6. The MPCI of the He-tuo and pill making process was calculated using principal component analysis. As illustrated in Fig. 7C, the MCp values for the mixing, honey refining, He-tuo and pill making processes are 0.71, 0.78, 0.73, and 1.02, respectively. The pill making process shows a capability index greater than 1, indicating superior performance, while the other processes have indices below 1.

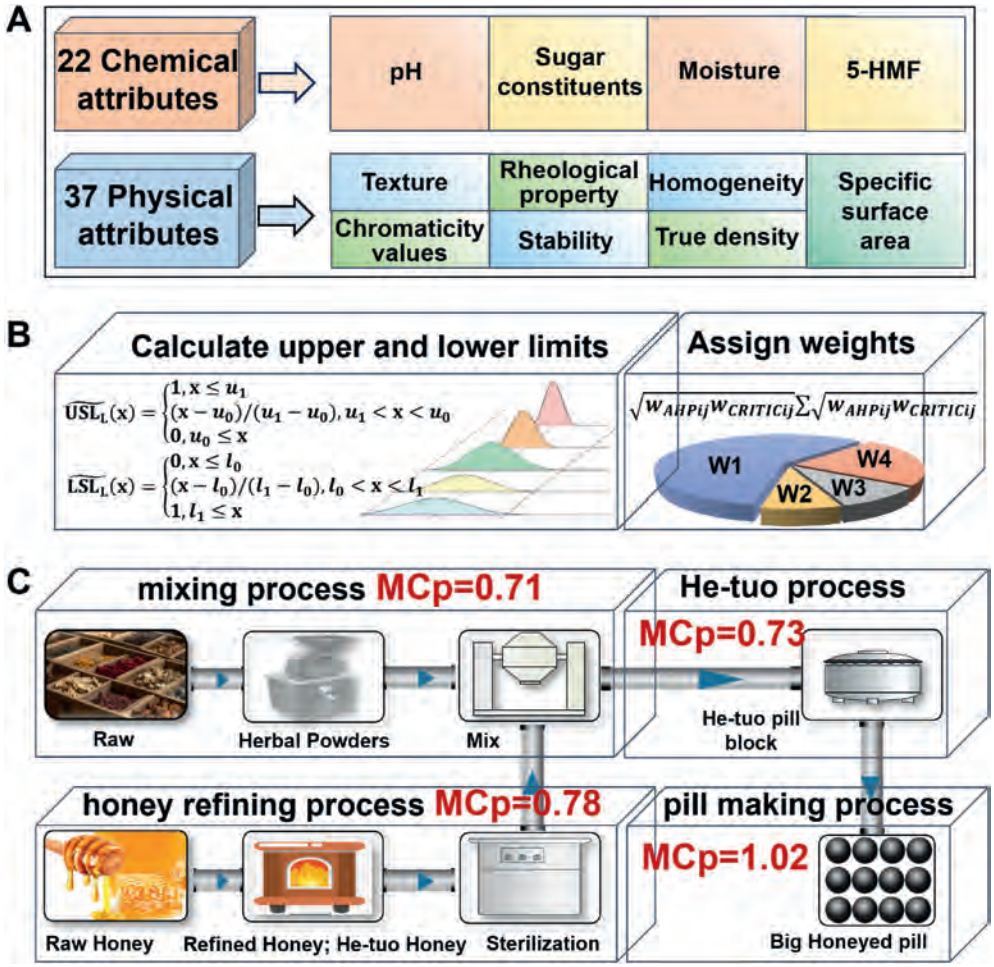

Figure 7 A novel digital multi-process capability evaluation system. (A) Acquisition of multi-source and multi-dimensional information. (B) Calculation of upper and lower specification limits for quality attributes using fuzzy set theory, and comprehensive weight determination *via* AHP-CRITIC. (C) Manufacturing process of Tong Ren Tang's Niuhuang Qingxin Pills, showcasing the MPCI for each manufacturing unit.

This highlights the need to focus on temperature control in the honey refining process and the timing of the mixing process, aligning with the previous diagnostic results.

3.3.2. Development of digital multivariate process capability evaluation software

The software was developed to evaluate MPCI at various stages of CMM manufacturing (Fig. 8A). It processes quality attribute data from each production unit (Fig. 8a1 and a2) and calculates the MPCI for each stage using the AHP-CRITIC method (Fig. 8a3), providing a comprehensive evaluation of process capability. The results are compiled into detailed reports (Fig. 8a4). The software workflow is depicted in Fig. 8B, beginning with data input and followed by normal distribution tests, weight calculations (using CRITIC and AHP), and process capability index evaluation. As shown in Fig. 8C, the user-friendly interface includes key functions such as raw data visualization, normal distribution testing, subjective and objective weight calculation, one-dimensional process capability index calculation, and comprehensive MPCI computation, with real-time status updates throughout the evaluation.

3.4. Practice innovation scenarios of the IQPD framework

To demonstrate the practical application of the proposed framework, the application scenario model has been constructed as depicted in Fig. 9. This model provides a detailed description of the implementation of the HCP system in the production process of big honeyed pills, revealing a complete manufacturing pathway from raw material processing to the final production. The HCP system is an integrated intelligent framework that combines human expertise, cyber system, and physical processes[38-41]. It is designed to enhance manufacturing efficiency, optimize decision-making, and achieve specific production objectives through the seamless interaction and coordination of its components. In this framework, the role of humans is manifested in the use of enterprise resource planning and manufacturing execution systems for information processing and decision-making. The physical system encompasses various machinery and sensors within the factory, which are directly responsible for product manufacturing and quality control. Meanwhile, the cyber system connects these two components, utilizing deep learning models to predict product

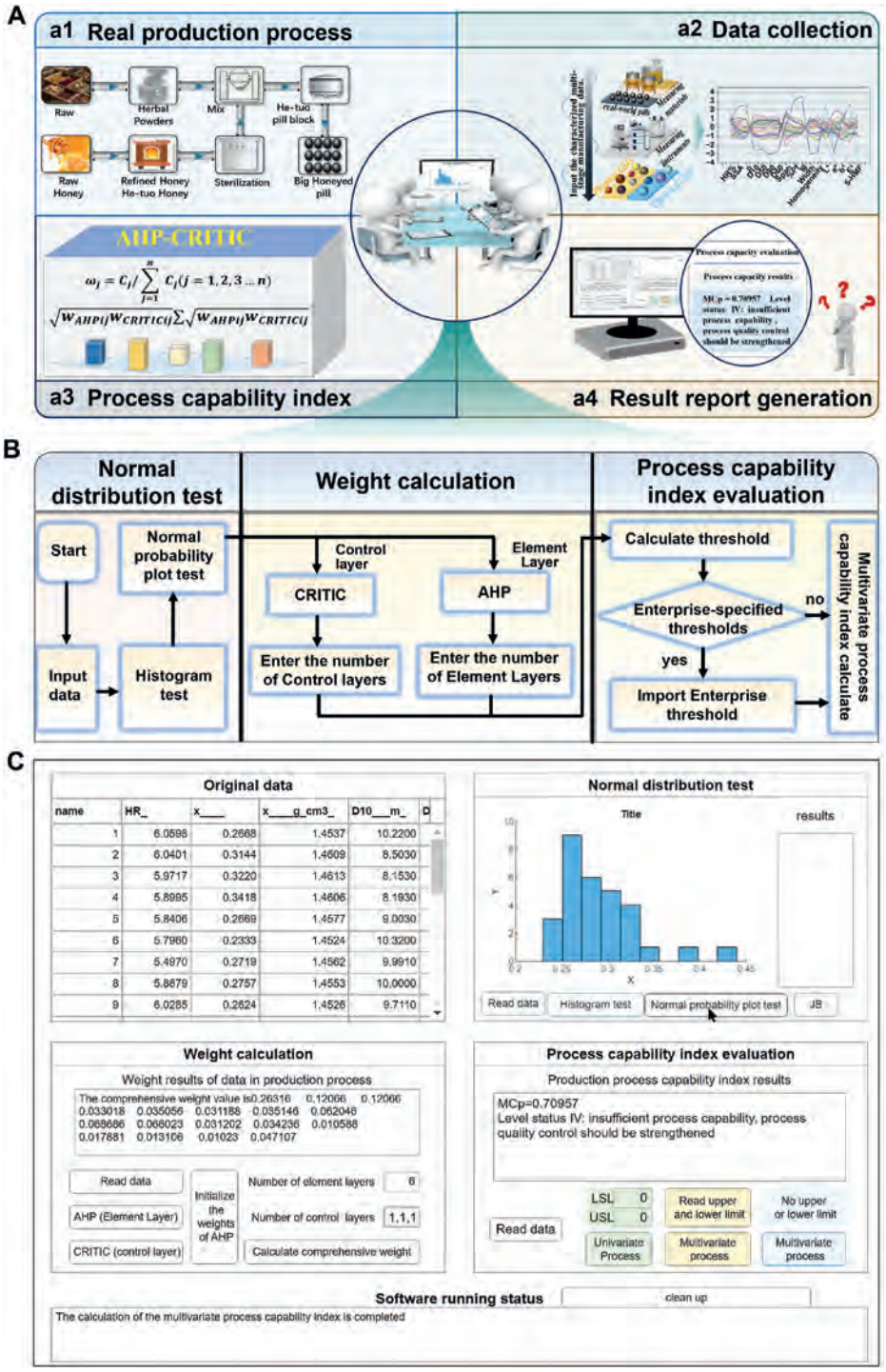

Figure 8 Digital multivariate process capability evaluation software. (A) Software development process: (a1) Real-world manufacturing process of Tong Ren Tang's Niuhuang Qingxin Pills; (a2) Collection of samples corresponding to batch numbers from actual production and measurement of their quality attributes; (a3) Calculation of the process capability index; (a4) Software development and report generation. (B) Workflow of software operation. (C) Software interface.

quality and implement anomaly diagnosis through the developed software.

Operators collect and measure samples throughout all manufacturing processes using onboard and offline measurement systems, thereby providing a dataset for training the model. Engineers input these measurement data into the developed PeDGAT to predict product quality and promptly identify any anomalies in product quality. The software is utilized to trace product quality at each critical point along the production line, enabling the diagnosis of the root causes of quality issues. This analysis guides technicians in making informed decisions to adjust process parameters, thereby optimizing

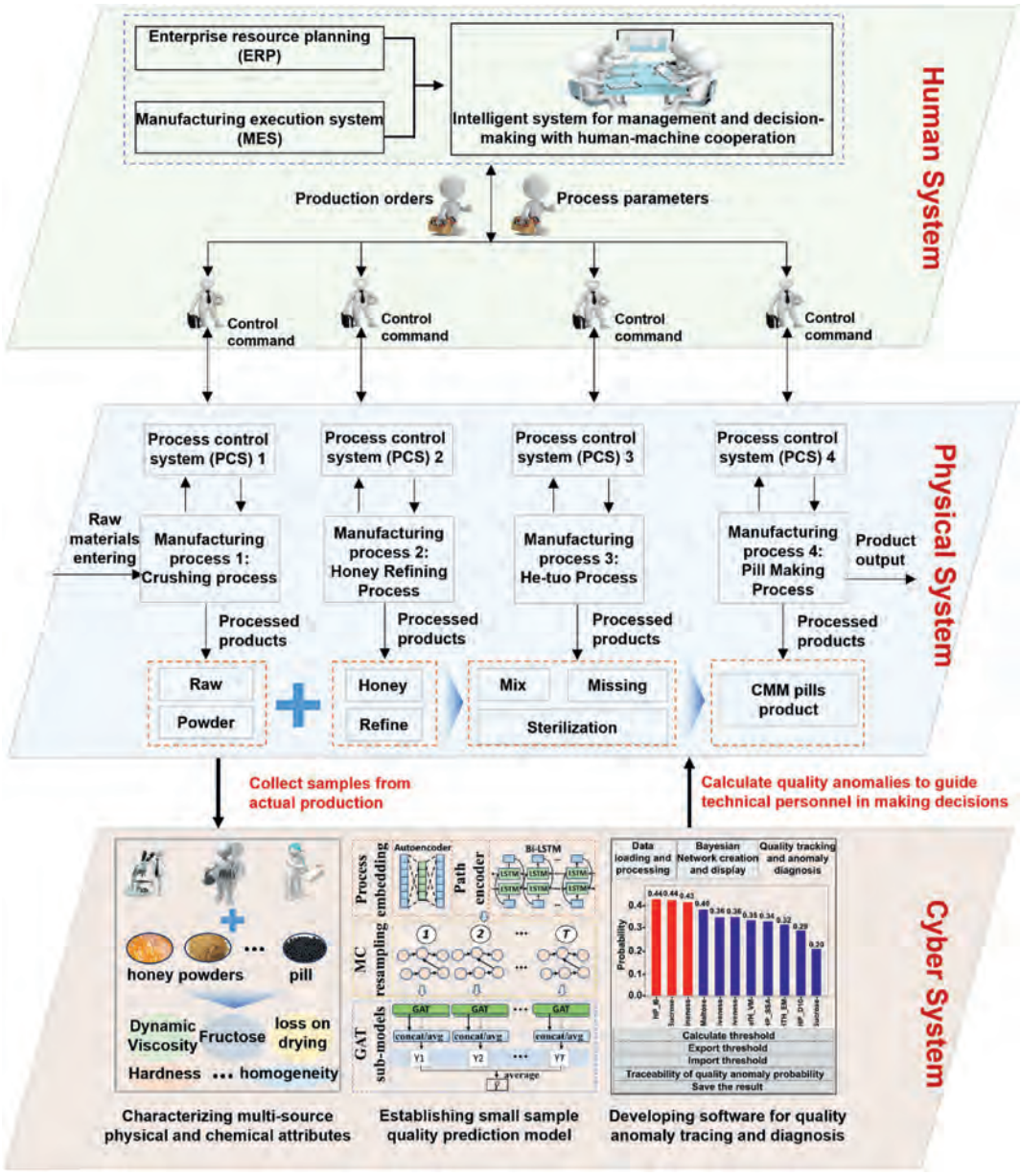

Figure 9 Practice innovation scenarios for CMM manufacturing.

the production process. Consequently, this framework establishes a solid foundation and application model for achieving an efficient and traceable intelligent manufacturing system for CMM.

4. Discussion

The core of the PeDGAT model lies the construction of a graph structure spanning multiple manufacturing stages. A Bi-LSTM was employed to encode the graph's long-range, sequential paths, while the attention mechanism of a graph attention network dynamically assigns weights to connected nodes based on their relative importance and aggregates features according to the predefined topology. This approach ultimately enables accurate quality prediction in multi-stage manufacturing processes. A double-ensemble strategy was developed to mitigate the characteristic instability and reliability challenges commonly encountered in small sample modeling environments. For effective implementation across diverse small-sample and multi-stage production processes, system-specific parameter optimization and architecture adjustments are recommended.

1. Quality embedding model

To address feature heterogeneity in multi-unit pharmaceutical production, an autoencoder-based cross-unit feature fusion approach is proposed, which projects heterogeneous features into a homogeneous dimensional space to ensure consistent node-feature inputs for graph neural networks. For adaptation to other small-sample scenarios, the autoencoder architecture and loss

functions (*e.g.*, incorporating a sparsity penalty) are adjusted according to data dimensionality and feature patterns, thereby achieving effective cross-unit feature fusion.

2. Graph construction and data preparation

The static graph topology defined by the edges should reflect the process flow of the new manufacturing scenario. Edges, node connectivity, and possible parallel steps can be reconfigured to match the real-world production stages. The data preparation step—where each sample is transformed into a data object that includes the feature embedding (x), edge_index, and the target (y)—remains unchanged, but it can be augmented with any domain-specific attributes at the node or edge level.

3. GAT model with path encoder

The path encoder (a Bidirectional LSTM) and GAT layers can be adapted to different levels of sequence complexity. If a process exhibits stronger sequential dependencies, one can increase the hidden dimension (h_dim) or consider alternative sequence models (*e.g.*, Transformers). For other application scenarios, adjusting the number of attention heads or tuning model parameters can help mitigate overfitting in small-sample settings.

4. MC resampling training

The MC resampling strategy can be fine-tuned by modifying the subsample ratio and the number of estimators to accommodate the size and variability of a new dataset. In small-sample settings, revising the early stopping criteria can further enhance model robustness.

By adjusting these components according to the characteristics of different small-sample manufacturing processes, PeDGAT can provide reliable quality predictions while retaining its core advantages of feature compression, graph-based modeling, sequence encoding, and robust model training.

5. Conclusions

In summary, this study successfully developed a multi-unit intelligent quality prediction and diagnostic (Multi-Unit IQPD) framework tailored for small sample scenarios. The framework comprises two main components: a quality prediction model and a quality anomaly diagnostic model. The quality prediction model integrates three key elements: a process embedding pool, path encoding, and a double ensemble strategy. The process embedding pool maps process features into a unified feature space, while the path encoding method captures dependencies between processes. The double ensemble strategy further enhances the model's ability to predict different metric values within a small sample context. Experimental results show that PeDGAT achieves state-of-the-art performance on small sample datasets, underscoring its feasibility for practical industrial applications. Additionally, the quality anomaly diagnostic model combines grey relational analysis with expert knowledge to improve diagnostic accuracy in small sample settings. Finally, the IQPD framework, when applied to CMM production through a HCP system, provides both method insights and software tools for real-time monitoring and process optimization. This innovative approach opens avenues for a more efficient, reliable, and traceable modern pharmaceutical industry.

While the proposed model demonstrates promising capabilities, several improvements warrant further investigation. First, the generalization capability of the proposed model under limited data conditions remains constrained. To overcome this limitation, continuous integration of newly accumulated production data is planned, progressively enhancing predictive robustness as datasets expand. Second, the current model utilizes a static graph structure informed by prior process knowledge. Although this design ensures interpretability and aligns well with domain expertise, dynamic adjustments or unexpected deviations frequently encountered in practical manufacturing scenarios may not be fully represented. Thus, the incorporation of adaptive or dynamic graph neural network architectures will be explored in future research to enhance model responsiveness to evolving production conditions. Thirdly, although methodological viability has been rigorously demonstrated through validation on a single pharmaceutical product, broader generalization across diverse pharmaceutical products and manufacturing environments has yet to be assessed. Consequently, collaborations with additional pharmaceutical partners are actively being pursued to further evaluate and improve cross-domain adaptability. Lastly, the current approach emphasizes independent prediction of individual quality indicators. To better align with comprehensive quality management requirements, multi-task learning methodologies and joint-indicator optimization strategies will be investigated in future studies, enabling a more holistic representation of overall production quality.

Acknowledgments

We would like to express our sincere gratitude to Jingqi Zeng and Mingli Zhu (from Beijing University of Chinese Medicine, Beijing, China) for kindly providing the raw data that was essential for this study. This study was graciously supported by the National Key R & D Program of China (2023YFC3504505), National Medical Research and Development Platform for Industry Education Integration Innovation of China—Traditional Chinese Medicine Smart Manufacturing Engineering (90010062820031, China), Undergraduate and Postgraduate Research Project of Beijing University of Chinese Medicine (BXX23018, China).

Author contributions

Kaiyi Wang: Conceptualization, Data curation, Methodology, Software, Validation, Visualization, Writing − original draft, Writing − review & editing. Xinhai Chen: Data curation, Writing − original draft, Funding acquisition. Nan Li: Funding acquisition, Supervision. Huimin Feng: Writing − original draft, Writing − review & editing. Xiaoyi Liu: Writing − original draft. Yifei Wang: Supervision. Yanfei Wu: Writing − original draft. Yufeng Guo: Writing − review & editing. Shuoshuo Xu: Writing − review & editing. Lu Yao: Resources. Zhaohua Zhang: Resources. Jun Jia: Resources. Zhishu Tang: Supervision. Zhisheng Wu: Conceptualization, Funding acquisition, Supervision.

Conflicts of interest

The authors declare no conflicts of interest.

Appendix A. Supporting information

Supporting information to this article can be found online at https://doi.org/10.1016/j.apsb.2025.06.001.

References

1. Tao F, Qi QL. Make more digital twins. *Nature* 2019;**573**:490−1.
2. Yli-Ojanperä M, Sierla S, Papakonstantinou N, Vyatkin V. Adapting an agile manufacturing concept to the reference architecture model industry 4.0: a survey and case study. *J Ind Inf Integr* 2019;**15**: 147−60.
3. Xu LD, Xu EL, Li L. Industry 4.0: state of the art and future trends. *Int J Prod Res* 2018;**56**:2941−62.
4. Diez-Olivan A, Del Ser J, Galar D, Sierra B. Data fusion and machine learning for industrial prognosis: trends and perspectives towards industry 4.0. *Inform Fusion* 2019;**50**:92−111.
5. Iqbal R, Maniak T, Doctor F, Karyotis C. Fault detection and isolation in industrial processes using deep learning approaches. *IEEE Trans Industr Inform* 2019;**15**:3077−84.
6. Yuan Y, Ma G, Cheng C, Zhou B, Zhao H, Zhang HT, et al. A general end-to-end diagnosis framework for manufacturing systems. *Natl Sci Rev* 2020;**7**:418−29.
7. Yang YH, Jia Y, Yu HY, Liu X. Spatiotemporal graph attention network for soft sensor modeling of suspended magnetization roasting process. *IEEE Trans Instrum Meas* 2024;**73**:1−15.
8. Kusiak A. Smart manufacturing must embrace big data. *Nature* 2017; **544**:23−5.
9. Yang T, Yi XL, Lu SW, Johansson KH, Chai TY. Intelligent manufacturing for the process industry driven by industrial artificial intelligence. *Engineering* 2021;**7**:1224−30.
10. Xie PS, Leung AY. Understanding the traditional aspect of Chinese medicine in order to achieve meaningful quality control of Chinese materia medica. *J Chromatogr A* 2009;**1216**:1933−40.
11. Cheung F. TCM: made in China. *Nature* 2011;**480**:S82−3.
12. Zhong K, Han M, Han B. Data-driven based fault prognosis for industrial systems: a concise overview. *JAS* 2019;**7**:330−45.
13. Normile D. The new face of traditional Chinese medicine. *Science* 2003;**299**:188−90.
14. Liu FX, Dai Y. Product quality prediction method in small sample data environment. *AEI* 2023;**56**:101975.
15. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**: 436−44.
16. Yuan XF, Wang YL, Yang CH, Ge ZQ, Song ZH, Gui WH. Weighted linear dynamic system for feature representation and soft sensor application in nonlinear dynamic industrial processes. *IEEE Trans Ind Electron* 2017;**65**:1508−17.
17. Yao L, Ge Z. Cooperative deep dynamic feature extraction and variable time-delay estimation for industrial quality prediction. *IEEE Trans Ind Inf* 2020;**17**:3782−92.
18. Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci* 2021; **2**:420.
19. Chih HY, Fan YC, Peng WC, Kuo HY. *Product quality prediction with convolutional encoder-decoder architecture and transfer learning*. *Proceedings of the 29th ACM CIKM*. 2020. p. 195−204.
20. Kim J, Yang Z, Ko H, Cho H, Lu Y. Deep learning-based data registration of melt-pool-monitoring images for laser powder bed fusion additive manufacturing. *J Manuf Syst* 2023;**68**:117−29.
21. Jiang RB, Smith J, Yi YT, Sun T, Simonds BJ, Rollett AD. Deep learning approaches for instantaneous laser absorptance prediction in additive manufacturing. *npj Comput Mater* 2024;**10**:6.
22. Wang XF, Yan JH. Deep learning based multi-source heterogeneous information fusion framework for online monitoring of surface quality in milling process. *Eng Appl Artif Intell* 2024;**133**:108043.
23. Liu ZY, Zhang DH, Jia WQ, Lin XK, Liu H. An adversarial bidirectional serial−parallel LSTM-based QTD framework for product quality prediction. *J Intell Manuf* 2020;**31**:1511−29.
24. Yuan XF, Ou C, Wang YL, Yang CH, Gui WH. A layer-wise data augmentation strategy for deep learning networks and its soft sensor application in an industrial hydrocracking process. *IEEE Transact Neural Networks Learn Syst* 2019;**32**:3296−305.
25. Bilali AE, Taleb A, Bahlaoui MA, Brouziyne Y. An integrated approach based on Gaussian noises-based data augmentation method and AdaBoost model to predict faecal coliforms in rivers with small dataset. *J Hydrol* 2021;**599**:126510.
26. Wang KY, Bian XH, Tan XY, Wang HT, Li YK. A new ensemble modeling method for multivariate calibration of near infrared spectra. *Anal Methods* 2021;**13**:1374−80.
27. Li QQ, Zeng JQ, Ma LJ, Zhu JY, Zhang FY, Wei YN, et al. Successive challenge for multi-batch overall pharmaceutical manufacturing control from two-step digital process to OQC strategy by integrated intelligent algorithm. *J Ind Inf Integr* 2023;**33**:100454.
28. Chu KY, Liu R, Duan GJ. A gray correlation based Bayesian network model for fault source diagnosis of multistage process—small sample manufacturing system. *Adv. Eng. Inform* 2023;**56**:101918.
29. Jiang GQ, Wang J, Wang LJ, Xie P, Li YW, Li XL. An interpretable convolutional neural network with multi-wavelet kernel fusion for intelligent fault diagnosis. *J Manuf Syst* 2023;**70**:18−30.
30. Zhang XY, Ma L, Peng KX, Zhang CF. A novel quality-related distributed fault diagnosis framework for large-scale sequential manufacturing processes. *IEEE Trans Ind Inf* 2024;**20**:4397−407.
31. Tran KP. *Control charts and machine learning for anomaly detection in manufacturing*. Springer; 2022.
32. Puga JL, Krzywinski M, Altman N. Points of significance: Bayesian networks. *Nat Methods* 2015;**12**:799−800.
33. Puerta JM, Aledo JA, Gámez JA, Laborda JD. Efficient and accurate structural fusion of Bayesian networks. *Inform Fusion* 2021;**66**: 155−69.
34. Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. *Graph attention networks*. *International conference on learning representations*. ICLR); 2017. p. 1−12.
35. Huang J, Su JY, Chang Q. Graph neural network and multi-agent reinforcement learning for machine-process-system integrated control to optimize production yield. *J Manuf Syst* 2022;**64**:81−93.
36. Chen J, Zheng L, Hu YZ, Wang W, Zhang HX, Hu XP. Traffic flow matrix-based graph neural network with attention mechanism for traffic flow prediction. *Inform Fusion* 2024;**104**:102146.
37. Zhang WQ, Zhao F, Li Y, Du C, Feng XB, Mei XS. A novel collaborative agent reinforcement learning framework based on an attention mechanism and disjunctive graph embedding for flexible job shop scheduling problem. *J Manuf Syst* 2024;**74**:329−45.
38. Zhou J, Zhou YH, Wang BC, Zang JY. Human−cyber−physical systems (HCPSs) in the context of new-generation intelligent manufacturing. *Engineering* 2019;**5**:624−36.
39. Liu Q, Liu M, Zhou HL, Yan F, Ma YY, Shen WM. Intelligent manufacturing system with human-cyber-physical fusion and collaboration for process fine control. *J Manuf Syst* 2022;**64**:149−69.
40. Wang BC, Zheng P, Yin Y, Shih A, Wang LH. Toward human-centric smart manufacturing: a human-cyber-physical systems (HCPS) perspective. *J Manuf Syst* 2022;**63**:471−90.
41. Wang SL, Yang JH, Yang B, Li D, Kang L. An intelligent quality control method for manufacturing processes based on a human−cyber−physical knowledge graph. *Engineering* 2024;**41**: 242−60.