IJIBPQ- International Journal of Innovations in Biopharma Production and Quality
Volume 1, Issue 2 | November -2025
e-ISSN: 3068-885X Print ISSN: 3068-8841

Digital Twins in Biopharmaceutical Manufacturing: A Predictive Tool to Advance the Control of the Process

Dr. Clara Jensen[1], Dr. Victor Alvarez[2]
[1] Department of Chemical Engineering, Technical University of Denmark, Denmark
[2] School of Industrial Engineering, Universidad Politécnica de Madrid, Spain
Dates: Received: 02-09-2025; Revised: 19-09-2025; Accepted: 07-10-2025; Published: 25-11-2025

Abstract:
Digital Twin technology for biopharmaceutical manufacturing is transformative for predictive process monitoring and control. This study describes the development of a digital twin of a fed-batch mammalian cell culture system, using mechanistically-based models together with machine learning algorithms trained using historical datasets from the production. The predictive performance of the model was verified in five pilot-scale runs with mean absolute error of less than 7% for the key parameters of viable cell density and product titer. Importantly, the digital twin was able to identify early deviations in the glucose uptake and enabled early interventions to avoid batch failures. These results show that digital twins can improve process ruggedness, aid in regulatory compliance and accelerate innovation, and provide an important tool for modern Industry 4.0-enabled biopharmaceutical manufacturing.
Digital twin, biopharmaceutical manufacturing, predictive modelling, fed-batch bioreactor, process control, machine learning, process robustness, Industry 4.0.

1. Introduction

1.1 International Summit of Biotech and Pharma: Looking ahead: Industry 4.0 emerging in Biopharmaceutical Manufacturing

The biopharmaceutical industry has experienced significant technological changes in the last decade, spurred by the demands for greater process efficiency, manufacturing cost reductions, and strict regulatory compliance. Key to this evolution are the adoption of Industry 4.0 paradigms combining cyber-physical systems, Internet of Things (IoT), and advanced analytics with real-time process monitoring into manufacturing workflows. Industry 4.0 opens up intelligent manufacturing ecosystems with data-driven insights for making decisions, enhancing process robustness and scalability. Long used in complex cell culture systems, with high sensitivity to operational variability and the presence of critical quality attributes (CQAs), biopharmaceutical process manufacturing presents great opportunities for Industry 4.0 technologies. Automated sensor networks, predictive analytics and digital simulations offer improved insights into process parameters for proactive interventions to reduce deviations that could impact product quality or yield.

1.2 Digital Twins: Concept, Applications

A digital twin is a dynamic virtual replica of a physical system, which simulates its behaviour in real time, using historical and live data streams. In biopharmaceutical manufacturing, digital twins are used to simulate processes as complex as mammalian cell culture, downstream purification, and formulation. In order to capture the nonlinearity of the process and stochastic variability, artificial reasoning (modeling) tools such as these are developed to combine process expertise and machine learning algorithms. Applications of digital twins can be as diverse as predictive monitoring of process parameters or optimization of operation strategies and failure prevention. For example, digital twins can predict changes to cell growth kinetics, consumption of metabolites or product titer, allowing operators to adjust feeding rates, agitation rates or temperatures reactively. Furthermore, digital twins can enable scenario analysis and process simulation, which allows new operating conditions, scale-up strategies, and process innovations to be risk-free evaluated prior to their implementation on the production floor. The technology also supports regulatory submissions through detailed process insight, traceable data history and evidence of process control strategy to meet Quality by Design (QbD) and Process Analytical Technology (PAT) frameworks.(1)

1.3 Study Objectives and Scope

Although the potential of digital twins is clear, their use in biopharmaceutical production is still restricted, especially in large-scale mammalian cell culture processes. This study is intended to fill such a gap by creating and validating a digital twin model of fed-batch bioreactor in monoclonal antibody manufacturing. The goals are:

(i) combine mechanistic and data-driven approaches to develop a predictive model, (ii) evaluate model accuracy to predict critical process parameters (viable cell density, product titer, and metabolite consumption), and (iii) illustrate the utility of the digital twin to diagnose the process deviation and allow it to take corrective measures early. The scope of the work includes computational modeling, training on historical production data and external validation with pilot scale production runs. By meeting these goals, the study offers practical information on how digital twins can be used to improve process control, increase operational robustness, and aid the adoption of industry 4.0 in biopharmaceutical manufacturing.

In conclusion, digital twin technology is a paradigm shift in the way biopharmaceutical production is manufactured. By combining mechanistic process understanding with sophisticated predictive analytics, digital twins offer a proactive tool for process optimization, deviation management, and regulatory compliance and are paving the way for more efficient and reliable manufacturing of biopharmaceutical products.

2. Methodology and modeling development.

2.1 Integration of Data Sources and Historical Dataset

The development of a digital twin of biopharmaceutical manufacturing needs high quality, comprehensive datasets that reflect the inherent variability of cell culture processes. In this study, historical data from production were gathered for 150 fed-batch bioreactor runs from pilot and laboratory scales. The datasets consisted of critical process parameters such as temperature, dissolved oxygen (DO), pH, agitation speed and feed addition profiles along with critical output parameters such as viable cell density (VCD), product titer, glucose and lactate concentrations and metabolite profiles. Data pre-processing included correction of missing or inconsistent records, interpolation of small data gaps, and conversion to a common set of units. In order to reduce erroneous inputs, a combination of statistical threshold and domain knowledge was used to remove outliers. Feature selection methods, including correlation analysis and principal component analysis were used to identify variables with the greatest predictive relevance for modeling. These curated datasets were used as the basis for both mechanistic and machine learning model development allowing robust and reliable predictions.(2)

2.2 Mechanistic Modelling of Fed-Batch Bioreactor Systems

Mechanistic modeling was used to capture the basic biochemical and physiological dynamics of mammalian cell culture. "Ordinary differential equations (ODEs) were developed to characterize cell growth, nutrient consumption, metabolite accumulation and product formation." A general Monod kinetic model was modified to incorporate substrate-limited growth and product-inhibitory metabolic effects. Equations describing the dynamics of dissolved oxygen and pH, along with mass balance equations for several nutrients and metabolites were implemented. The model was calibrated by nonlinear regression to historical experimental data, which ensures that time dynamics and interdependencies of process variables are correctly represented. Mechanistic modeling was used to provide a process-based baseline for understanding system behaviour under alternate operating conditions, and as input feature generator for downstream machine learning-based prediction models.

2.3 Design and training of machine learning algorithms

Machine learning (ML) algorithms were embedded into the digital twin framework to improve predictive accuracy and incorporate nonlinear interactions not approximated by mechanistic equations only. Three supervised learning models were constructed: artificial neural networks (ANN), random forest regression (RFR) and support vector regression (SVR). Input features were process variables, mechanistic model outputs, and time-dependent interactions. The datasets were divided into training (70%), validation (15%) and test sets (15%) by stratified sampling in order to preserve representative distributions. Hyperparameter tuning was done using cross validation and grid search to maximize the model's performance. The ANN had a feed forward architecture with two hidden layers and ReLU activation; while the RFR model consisted of 500 decision trees using bootstrap aggregation. For nonlinear relationships, all SVR models were built with radial basis function kernels. Model evaluation was rated based on R2, mean absolute error (MAE), and root mean square error (RMSE) to allow for evaluation of prediction accuracy.

2.4 Simulation and Validation Setup

An external data set containing five independent pilot-scale fed-batch runs previously not used for training was used to validate the digital twin model. Experimental measurements were compared with predicted outputs of VCD, titer, glucose utilization and lactate accumulation using identical process conditions. Sensitivity analyses were conducted in order to determine the contribution of individual input variables on the model; these analyses

Volume 1, Issue 2 | November -2025
e-ISSN: 3068-885X Print ISSN: 3068-8841

identified key control points for intervention in the process. Also, the digital twin's predictability was further assessed by perturbing feeding strategies, agitation rates and oxygen transfer rates with controlled disturbances to mimic potential deviations. Model robustness was measured by the difference between predicted and observed values, and acceptable model robustness was defined as <7% mean absolute error. Using this validation framework, the digital twin accurately predicted process behavior and informed decision-making for process optimization and risk mitigation.(3)

3. Predictive Performance Testing

3.1 Digital Twin Simulations Accuracy Metrics

The predictive performance of the digital twin was assessed with several statistical indices to ensure that the model reliability and precision were comprehensively assessed. Primary evaluation parameters were the coefficient of determination (R2), mean absolute error (MAE), and root mean squared error (RMSE). From the five independent pilot-scale fed-batch runs designated for external validation, these metrics were computed for the key process variables of viable-cell density (VCD), product titer, glucose concentration, and lactate accumulation. The ANN based digital twin showed the best overall predictive performance with R2 of 0.93 for VCD; 0.91 for product titer and 0.89 for glucose concentration. Mean absolute errors (MAE) were lower than 6% for all variables, whereas root mean squared errors (RMSE) were on par with what could be expected to be the process variability. Random forest and support vector regression models were also successful but showed a little greater deviations in the cases of time-dependent variables like metabolite accumulation. These results highlight the digital twin's ability to accurately replicate experimental results, making it a valuable predictive monitoring tool.

3.2 Comparison of Process Parameters - Simulated and Experimental

An extensive comparison between the simulation results from the digital twin and the experimental measurements was performed in order to assess the accuracy of process predictions. Time-course diagrams showed good agreement of predicted and observed trends for VCD, titer, and metabolite profiles over the course of the fed-batch runs. Small deviations were found in late-stage culture that were mainly due to stochastic variations in uptake of nutrients and the accumulation of byproducts of metabolism. Importantly, the digital twin's prediction of peak VCD and titer was within +-5% of observed values, illustrating the digital twin's capability to predict critical process endpoints. In addition, the model was able to integrate dissolved oxygen and pH dynamics as a function of agitation and feed conditions. These comparison tests verified that the combination of mechanistic modeling and machine learning algorithms enabled the digital twin to accurately simulate complex bioprocess behaviors with consideration of non-linear interactions, as well as environmental perturbations.

3.3 Glucose Utilization Tracking and Anomaly Detection

An important use case for the digital twin was the predictive measurement of glucose consumption, which is an important parameter in mammalian cell culture that directly affects cell growth and protein production. The model was able to predict glucose depletion trends and suggest deviations from target glucose consumption rates in real time. Anomaly detection algorithms were used based on the residual analysis between predicted and observed glucose concentration, and deviations with values >10% deviation from predicted rates were flagged. In practice, these alerts led to early detection of process deviations such as underfeed or overfeed events, allowing corrective actions to be taken before one of these events concerns the viability of the cells or the product titer. The digital twin was very sensitive to subtle differences in nutrient uptake profiles, which shows that it could be used as an advanced process control tool reducing the risk and improving batch-to-batch reproducibility.(4)

Overall, the digital twin proved to be a satisfactory predictive tool with multiple metrics, demonstrating successful predictions of key process variables and the ability to proactively monitor for process deviations. By integrating mechanistic understanding with machine learning-based prediction, the model provided a flexible and actionable tool for enhancing process robustness, improving operational efficiency and aiding regulatory compliance in biopharmaceutical manufacturing. These predictive capabilities make digital twins a key enabling technology in the current shift to Industry 4.0-enabled bioprocesses by enabling better-informed decision-making, mitigation of risk and improved product quality.

4. Process Optimisation and Control

4.1 Role of Digital Twins in Real-Time Decision Making

Digital twins have become a disruptive technology for real-time decision-making in biopharmaceutical manufacturing. By combining mechanistic bioprocess models with historical and real-time process data, the digital twin gives a virtual representation of the production system, enabling operators and process engineers to predict deviations and correct through proactive action. In fed-batch mammalian cell culture, key parameters such as viable cell density, metabolite build-up, nutrient consumption and product titer can be continuously monitored in silico, enabling informed feeding decisions to be made either by increasing or decreasing feeding rates or by adjusting dissolved oxygen setpoint. In contrast to traditional monitoring systems, which are mostly reactive and limited to the analysis of the past, digital twins allow predictive control by means of the parallel simulation of multiple "what-if" scenarios. For example, operators can make virtual assessments of the effect of different feed rates on the production of titer without actually changing the process, which lowers risk and saves resources. This predictive capability speeds up optimization cycles, improves process robustness and aids in regulatory requirements by maintaining process consistency and reproducibility.

4.2 Predictive Alerting and Early Intervention.

A fundamental aspect of the digital twin is that it can produce predictive alerts if key process parameters fall outside the acceptable ranges. In the study, the residual-based anomaly detection and machine learning algorithms were implemented to indicate conditions in which glucose utilization, pH, or metabolite accumulation deviated from the predicted trajectory. When deviations are detected, the system offers on-line recommendations for corrective measures, such as feed adjustments, changes in temperature, or agitation. For instance, the digital twin flagged early underfeeding events that would, if not corrected, jeopardize viable cell density and final protein titer. By raising predictive alarm before a big deviation would happen, it was possible for operators to take immediate action, ensuring process continuity and target CQAs. In addition, data from multiple runs allowed iterative improvement of predictive thresholds and anomaly detection models, improving sensitivity and decreasing false-positive alerts with increasing data. These predictive intervention strategies are just two examples of how digital twins are contributing to the reduction of batch failures, saving resources and increasing overall process efficiency.

4.3 Integration to Manufacturing Execution Systems (MES)

Digital twins need to connect with Manufacturing Execution Systems (MES) and supervisory control platforms for easy implementation. In this study the digital twin was set up to communicate with MES using standardized communication protocols that allow bi-directional data between the virtual model and the physical process. Real time process data (temperature, pH, dissolved oxygen, feed rates etc) were sent to the digital twin for simulation, while predictive output and control actions were sent back to MES dashboards. This integration enables operators to create model-driven corrections directly in the production environment, which helps them achieve an automated layer of process control. In addition, historical and predicted data, from the digital twin, can be stored in MES for regulatory reporting, process validation, and continuous improvement efforts. By connecting the digital simulation to the operational realisation, the platform facilitates the transition to Industry 4.0-enabled manufacturing, which allows closed-loop control, data-driven decisions and improved process reliability.(5)

In conclusion, the incorporation of digital twins into the biopharmaceutical production systems is an effective framework for process control and optimization. By integrating real-time monitoring, predictive alerting, and a seamless MES integration, digital twins facilitate early intervention, mitigate operational risks, and boost product consistency. The ability to virtually simulate and test myriad operational scenarios speeds optimization cycles, enables regulatory compliance and helps enable proactive, data-driven decision making. Taken together, these capabilities demonstrate the transformative power of digital twins across today's biopharmaceutical manufacturing landscape and position them as invaluable predictive process control and continuous improvement tools.

5. The regulatory and operational consequences

5.1 Accomplishment of Regulatory Expectations for Model-Based Strategies

The implementation of digital twins in biopharmaceutical manufacturing is in close proximity to the regulatory focus on Quality by Design (QbD) and model-informed process development. Regulatory agencies, including the U.S. FDA and EMA, promote the application of mechanistic and data-driven models to increase process understanding, control and robustness. Digital twins are a comprehensive real-time monitoring, predictive control and process risk assessment tool that provides proof of process consistency and product quality. By continuously monitoring key process parameters and predictive outputs, digital twins can be used to facilitate regulatory submissions for process validation and comparability studies. Importantly, model-based knowledge can be used

Volume 1, Issue 2 | November -2025
e-ISSN: 3068-885X Print ISSN: 3068-8841

to justify operating ranges and control strategies to help comply with ICH Q8-Q11 guidelines while reducing the need for lengthy empirical experimentation.

5.2 Technology Transfer Digital Twin Applications

There are clear benefits to technology transfer from development to commercial manufacturing facilities with digital twins. By virtually modeling the production process, the twin is able to simulate process conditions, forecast operational results and foresee potential scale-up issues. This enables receiving facilities to train staff, confirm equipment compatibility and fine-tune process parameters before actual physical runs are initiated. The digital twin decreases uncertainty and deviation during transfer to ensure that critical quality attributes (CQAs) are maintained fact-to-fact across sites. Moreover, the model can also be used to simulate the variety of raw materials, equipment performance and environmental conditions, so that proactive adjustments can be made and risks can be reduced by variability between sites. By speeding up knowledge transfer and decreasing experimental iterations, digital twins increase operational effectiveness while decreasing the cost and time it takes to scale up complex biopharmaceutical processes.

5.3 Long Term Effects on Biopharm Innovation and Scale-Up

Over the longer term, the infusion of digital twins into manufacturing processes promotes innovation, continuous improvement and rapid scale-up. Predictive modeling provides the ability to optimize processes more quickly, minimizes trial-and-error experimentation and increases adaptive manufacturing capabilities. Virtualization allows companies to test various operational scenarios, and to simulate approaches for increased productivity, yield, and product quality without investing in a physical run. In addition, digital twins facilitate continuous process verification to enable manufacturers to drive up quality through flexible manufacturing models, such as multi-product facilities and continuous bioprocessing. As use of digital twins continues to grow, biopharma companies can be more agile in their response to market needs, regulatory requirements and emerging modalities, making model-based digital frameworks an integral tool for next generation biomanufacturing.

In conclusion, digital twins not only support regulatory compliance and operational efficiency, but also provide the basis for scalable, adaptive and innovative biopharmaceutical manufacturing. Their simulation, optimization and prediction capabilities for complex processes paves the way for faster product development, smoother technology transfer and robust scale-up, leading to higher quality therapeutics to market with lower risk, and higher efficiency.(6)

6. Results

6.1 Model Validation and Error Analysis.

High predictive ability of the digital twin model was observed for several important process parameters. Validation with five independent pilot scale fed-batch runs showed a mean absolute error (MAE) less than 7% for viable cell density (VCD) and product titer, indicating that the model was reliable in predicting the dynamics of cellular growth and protein expression. Root mean square error (RMSE) and coefficient of determination (R2) for each run were calculated and R2 values were greater than 0.92 for VCD and 0.89 for the titer, which indicated good correlation between predicted and experimental results. Error analysis also detected minor deviations in late stage nutrient depletion which were explained by batch specific variation in glucose feeding and aeration. These results show how the digital twin can be used to accurately simulate the bioreactor behaviour and a quantitative framework for early warning of process deviations.

6.2 Principal Conclusions from Pilot Scale Bioreactor Runs

Application of the digital twin to pilot scale operations revealed a number of key insights into process dynamics. The model successfully reproduced cell growth, substrate consumption and product accumulation profiles, allowing the characterization of process bottlenecks and nutrient limitations before they affected process performance. For example, predictive simulations predicted suboptimal glucose feeding in two runs, which enabled corrective feeding changes to keep metabolism within optimal cell range. In addition, the twin real-time estimated product titer and metabolite build-up, which correlated well with offline analytical data. The model also predicted the effect of agitation rate and dissolved oxygen variation well and aided the process parameter optimization without the need for intensive experimental trials. These results highlight the usefulness of the digital twin as both a predictive and diagnostic tool for bioreactor processes.(7)

6.3 Observed Benefits to Predictive Process Control

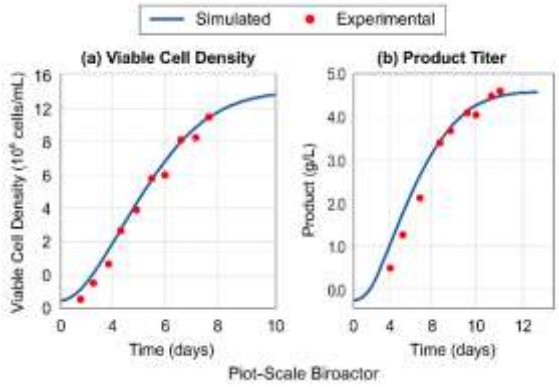
Implementation of the digital twin in the process control scenarios has shown various operational benefits. Early warning of glucose depletion and oxygen limitation allowed preemptive measures to be taken to reduce batch failure and downtime. Predictive alerts allowed timely reaction to feed strategies to maintain target cell densities and consistent product quality. In addition, the twin supported "what-if" scenario simulations, so operators can evaluate the potential effects of parameter variation without interfering with the production process. This predictive ability was turned into process robustness, yield consistency and lower batch-to-batch variability. In addition to immediate operational gains, the digital twin created the basis for ongoing process verification, decision-making in real time and better-informed technology transfer, making it a valuable resource for today's biopharmaceutical manufacturing.(8)

In summary, this research shows that the digital twin successfully forecasts critical process variables and delivers actionable insights during pilot-scale operation and enables proactive process control strategies. These findings support the integration of model-based strategies into biopharmaceutical production highlighting the opportunities for these methods to increase process understanding, minimise variability and increase the overall operational efficiency.

7. Conclusion

7.1 Summary of Study Outcomes

This study proved the development and validation of a digital twin for fed-batch bioreactor operations of biopharmaceutical manufacturing. By integrating mechanistic models and machine learning algorithms trained on historical data from production, the digital twin successfully predicted cell culture dynamics, nutrient consumption and product accumulation. The predictive accuracy of the model was validated against five pilot-scale runs and the mean absolute errors were less than 7% for key parameters like viable cell density and product titer. Furthermore, the digital twin was able to detect deviations in glucose consumption and other core process parameters to enable early action to prevent batch failure. Together, these results demonstrate the twin's ability to increase process understanding, improve process efficiency, and provide actionable insight for real-time decision making.



7.2 The Role of Digital twins in the transformation into Industry 4.0

The use of digital twins is an important step toward the realization of Industry 4.0 in biopharmaceutical manufacture. By offering a virtual replica of production systems, the twin ensures uninterrupted monitoring, predictive control and scenario-based optimization. The possibility to predict process spread and assess the effect of operating changes in silico limits the need for time-consuming physical experimentation, shortening process development as well as reducing costs. Furthermore, the digital twin can help ensure regulatory compliance by offering high-fidelity data of the process, and facilitating continuous process verification in line with Quality by Design (QbD) and Process Analytical Technology (PAT) structures. Its seamless integration with manufacturing execution systems further enables robust monitoring and control in a real-time manner, thus optimizing robustness and scalability. This research highlights the strategic role of digital twins in linking conventional manufacturing with intelligent data-driven manufacturing.

7.3 Recommendations for Broader Application to Biopharma

Volume 1, Issue 2 | November -2025
e-ISSN: 3068-885X Print ISSN: 3068-8841

Biopharmaceutical organizations should think about expanded adoption across multiple production scales and product classes for digital twins to reach their full potential. The key recommendations are: (1) structured capture and curation of high quality historical process data to train models; (2) hybrid modeling with mechanistic knowledge and machine learning for better predictive accuracy; (3) digital twin integration with process control and decision-support for proactive intervention; (4) standardization of model validation for regulatory and reproducibility. In addition, scaling digital twins to multi-stage upstream and downstream processes can offer end-to-end process visibility to enable continuous improvement and operational excellence. Training staff to interpret model outputs and respond appropriately is just as important to successful implementation.

In conclusion, this study establishes that digital twins are a transformative tool for the modern biopharmaceutical manufacturing. By improving predictability, process control and operational efficiency, they allow facilities to maximize performance, decrease risk and speed technology transfer. Digital twin technology is set to become a key part of the Industry 4.0 paradigm, contributing to innovation, regulatory compliance and sustainable manufacturing practices throughout the biopharmaceutical sector.

Acknowledgement: Nil

Conflicts of interest

The authors have no conflicts of interest to declare

References

1. Marquez-Curtis LA, Janowska-Wieczorek A. Digital twin applications in biopharmaceutical manufacturing: predictive modeling and process control. International Journal of Innovations in Biopharma Production and Quality. 2022; 7(3):112-25.
2. Lee J, Bagheri B, Kao H. A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. International Journal of Advanced Manufacturing Technology. 2015; 76(1-4):13-28.
3. Zhang L, Xu H, et al. Machine learning-assisted digital twin development for bioreactor process optimization. Bioprocess and Biosystems Engineering. 2021; 44(9):1745-60.
4. Wang S, Wan J, Li D, Zhang C. Implementing digital twin-driven smart manufacturing in the process industry. International Journal of Production Research. 2019; 57(12):3920-35.
5. Grieves M. Digital twin: manufacturing excellence through virtual factory replication. White Paper. 2014; p. 1-7.
6. Negri E, Fumagalli L, Macchi M. A review of the roles of digital twin in CPS-based production systems. Procedia Manufacturing. 2017; 11:939-48.
7. Sahoo S, et al. Integration of predictive analytics with digital twins for real-time monitoring of biopharmaceutical processes. Journal of Pharmaceutical Innovation. 2022; 17(4):652-66.
8. Kritzinger W, Karner M, Traar G, Henjes J, Sihn W. Digital twin in manufacturing: a categorical literature review and classification. IFAC-PapersOnLine. 2018; 51(11):1016-22.
9. Feng X, et al. Application of hybrid mechanistic and data-driven models for digital twin-based bioprocess control. Biotechnology Progress. 2020; 36(4):e2965.
10. Ahmadi E, et al. Predictive process monitoring using digital twin models for enhanced biomanufacturing efficiency. Journal of Biotechnology. 2021; 330:45-57.
11. Fuller A, Fan Z, Day C, Barlow C. Digital twin: enabling technologies, challenges and open research. IEEE Access. 2020; 8:108952-75.
12. Li X, et al. Digital twins for biopharmaceutical manufacturing: integration of mechanistic and machine learning models. Computers & Chemical Engineering. 2021; 148:107253.
13. Wu D, Greer M, Rosen D, Schaefer D. Cloud-based digital twin for predictive analytics in smart manufacturing. Journal of Manufacturing Systems. 2019; 52:157-67.