Chapter 5 of:

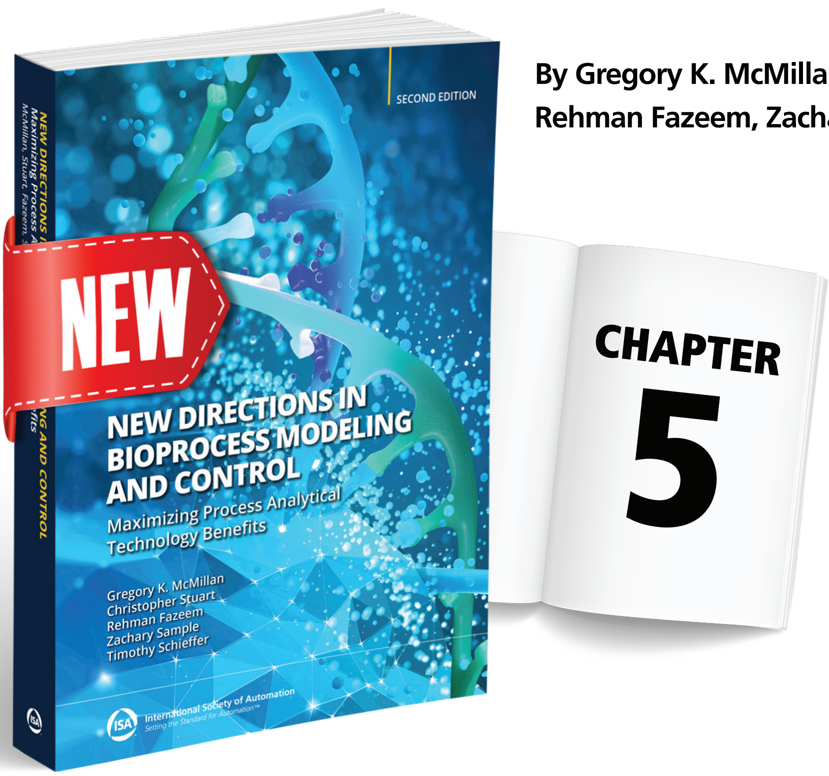
New Directions in Bioprocess Modeling and Control

Maximizing Process Analytical Technology Benefits

Gregory K. McMillan, Christopher Stuart, Rehman Fazeem, Zachary Sample, Timothy Schieffer

New Directions in Bioprocess Modeling and Control

Maximizing Process Analytical Technology Benefits



By Gregory K. McMillan, Christopher Stuart, Rehman Fazeem, Zachary Sample, Timothy Schieffer

Book Table of Contents

Buy the Complete Book

New Directions in Bioprocess Modeling and Control

Maximizing Process Analytical Technology Benefits

Second Edition

Gregory K. McMillan
Christopher Stuart
Rehman Fazeem
Zachary Sample
Timothy Schieffer

Copyright 2021. International Society of Automation. All rights reserved.

Notice

The information presented in this publication is for the general education of the reader. Because neither the author nor the publisher has any control over the use of the information by the reader, both the author and the publisher disclaim any and all liability of any kind arising out of such use. The reader is expected to exercise sound professional judgment in using any of the information presented in a particular application.

Additionally, neither the author nor the publisher has investigated or considered the effect of any patents on the ability of the reader to use any of the information in a particular application. The reader is responsible for reviewing any possible patents that may affect any particular use of the information presented.

Any references to commercial products in the work are cited as examples only. Neither the author nor the publisher endorses any referenced commercial product. Any trademarks or tradenames referenced belong to the respective owner of the mark or name. Neither the author nor the publisher makes any representation regarding the availability of any referenced commercial product at any time. The manufacturer's instructions on the use of any commercial product must be followed at all times, even if in conflict with the information in this publication.

Printed in the United States of America.
Version: 1.0

ISBN-13: 978-1-64331-104-3 (print)
ISBN-13: 978-1-64331-105-0 (ePub)
ISBN-13: 978-1-64331-106-7 (Kindle)

ISA
67 T. W. Alexander Drive
P.O. Box 12277
Research Triangle Park, NC 27709

Library of Congress Cataloging-in-Publication Data in process

Copyright 2021. International Society of Automation. All rights reserved.

Dedication

This book is dedicated to everyone working in the process industry who is seeking to advance process performance by gaining a greater understanding of first-principle process relationships and the opportunities to improve automation system dynamics and strategies. The authors are especially appreciative of practitioners who are willing to invest the time to use the digital twin and first-principle process simulations with automation dynamics to develop a deeper knowledge and cultivate innovation to get the most out of the largely untapped capability of today's measurements and control systems to achieve consistent and reliable process control improvements.

5 Digital Twin

5.1 Introduction

The requirements for management of change are stricter than ever in the process industry. Validation requirements for pharmaceuticals involving extensive and intensive documentation of design, qualification, and verification pose an even larger motivation to minimize change. The shear complexity and effort involved limits time for creativity and discourages doing anything that changes the process as defined by research and development (R&D). The Process Analytical Technology (PAT) initiative seeks to address these concerns to promote innovation and continuous improvement.

Often the process is set based on laboratory benchtop and pilot plant runs. Even here the opportunities are narrow because batch cycle times are extremely long, especially for new biologics (e.g., 10 days for mammalian cell bioreactors), and the time to market must be minimized to make the most out of patent protection for high-value-added products.

The key to addressing this reality is the digital twin. The actual configuration, including operator interface and alarm systems, is the same as used in the lab, pilot plant, or production plant. The signals can use actual input/output (I/O) channels for communication by simulated input/outputs (SIO) or a virtual input/output module (VIM). If necessary, an actual distributed control system (DCS) can be connected with a first-principle model of the process running with communications to the DCS via open platform communications (OPC). Recent breakthroughs in kinetic modeling (detailed in the new Chapter 6) enable model parameters to be readily set so that the batch profiles in the digital twin match those in the lab and plant.

Increasingly, R&D is realizing the value of the increased capability and flexibility of a small DCS to run batches automatically, greatly reducing cycle time and nonintentional variability. The transition from lab to pilot plant to production plant of the control system and digital twin is seamless and fast, greatly reducing the validation burden.

The key to the scientific method and innovation is experimentation. Presently, experimentation is largely limited to the early phases of R&D. It is increasingly recognized that the digital twin enables rapid and free experimentation with speedup factors greater than 500 so that a 10-day batch can be completed and analyzed in 30 minutes.

The digital twin previously known as virtual plant is a relatively new concept that is easily confused with existing simulation methods for process design, configuration checkout, and operating training systems. Users may not realize that most of the existing batch process simulations are off-line and noninteractive, and that most of the real-time dynamic process simulations were originally designed for continuous processes. These real-time process simulations can develop severe numerical errors or even fail under the extreme conditions of batch operations and require interfaces for communicating I/O, controlling inventory, and coordinating with the control system in speeding up, slowing down, pausing, or resuming. The control system engineer is probably most familiar with tieback simulations because these have been predominantly used for configuration checkout and operating training systems. The process response in these tiebacks is mimicked by the trial-and-error adjustment of ramp rates. Ramps are triggered by the manual inclusion of flow path logic to include the opening and closing of valves or the turning on or off of pumps.

Section 5.2 of this chapter discusses the key features that distinguish a digital twin from existing simulation systems and itemizes the important advantages of the digital twin approach. Section 5.3 provides a spectrum of uses for digital twins. Section 5.4 concludes with the basic requirements and issues concerning implementing a digital twin.

The last two sections note the role of a benchtop system with an industrial control system in getting the most out of a digital twin. This chapter shows how a digital twin creates an environment for innovation that is the heart of the PAT initiative.

Learning Objectives

- Appreciate the significance of being able to export and import to the real plant
- Recognize the differences between different types of simulations

- Know the diverse opportunities that are essential to getting the most from PAT
- Realize the importance of a benchtop and pilot plant system with an industrial control system
- Know the basic digital twin implementation steps
- Develop the ability to nonintrusively explore, develop, and demonstrate improvements
- Understand how to develop online process metrics to automatically document benefits of improvements
- Understand how to use an MPC nonintrusively for adaptation
- Be introduced to the simpler thermodynamic requirements of bioreactors
- Recognize the importance of a charge balance for pH

5.2 Key Features

The first key feature that distinguishes a digital twin from process simulators is its ability to use the configuration, historian, displays, and advanced control tool set of the real plant without translation, emulation, special interfaces, or custom modifications. The configuration database from the real plant can be exported and then imported and downloaded into a personal computer or a control system computer just as if it were an actual hardware controller. Also, files for operator graphics, process history charts, and data history from the real plant can be copied to the digital twin so the user has the entire control system of the real plant on a computer, as shown in Figure 5-2a (intent is to convey knowledge paths and not the actual process being simulated) [1].

The use of the actual configuration, database, displays, historian, and advanced control tool set without any translation, emulation, special interfaces, or custom modification enables inherent replication of the real control system.

Most dynamic, high-fidelity process simulation software offers the ability to build a basic control strategy or sequence inside the simulation environment. However, simulation developers tend to have a process rather than a control background and focus. It is unrealistic to expect the process and batch control capability offered in simulation software to be in the same realm as the control capability of DCS software, which is the culmination of a hundred engineering years (i.e., 5 years by 20 engineers) or more of an effort by process control experts. The overall control functionality in process simulators is primitive compared to the capabilities offered in the modern DCS. The

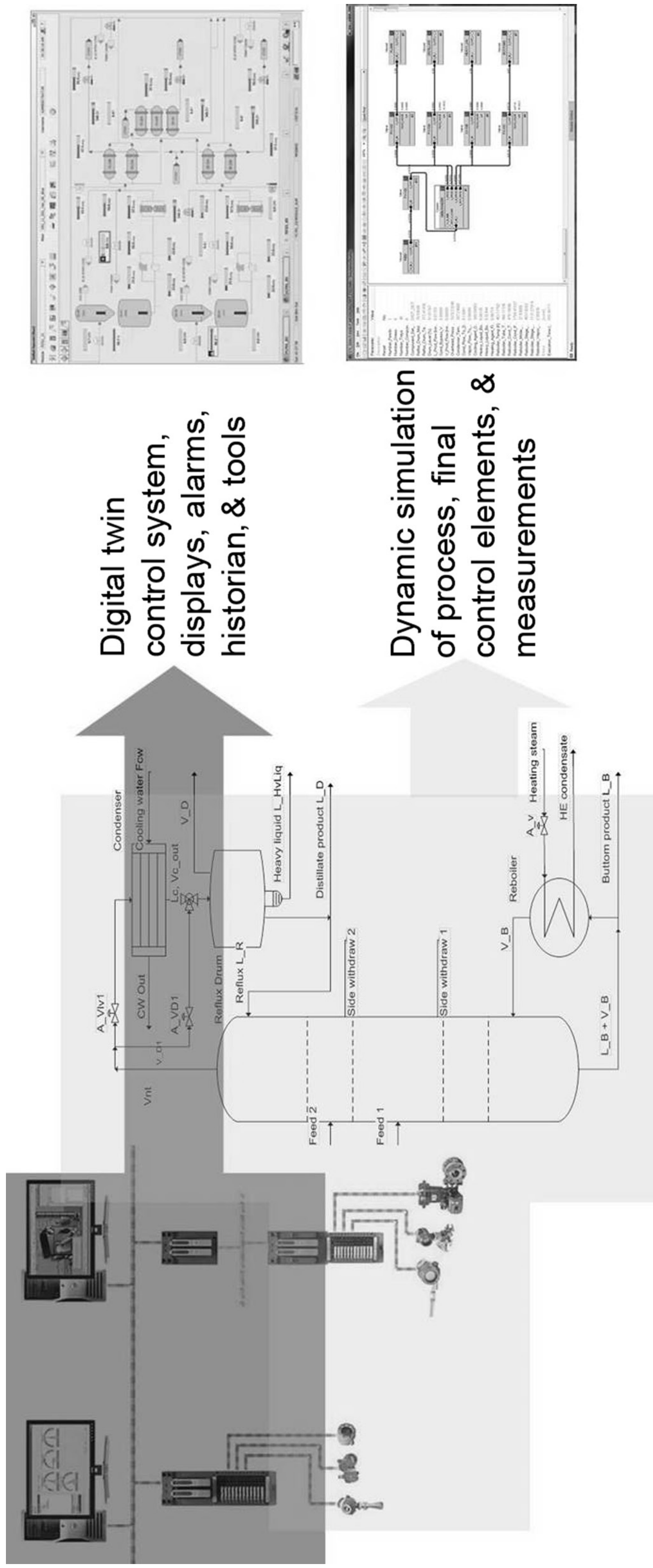


Figure 5-2a. Digital twin with actual DCS system and dynamic simulation of process and instrumentation.

DCS offers capabilities such as Sequential Function Charts and basic function blocks, the batch manager, and advanced control tools like multivariable model predictive control (MPC). These are almost nonexistent in simulators. The effort to duplicate even a simplified version of a control system in a dynamic simulation is great. At best, the user ends up with two control systems with no assurance that they will match well and no method for automatically managing changes between them.

Consequently, most simulation software now offers a standard or custom OPC interface. However, for the simulations with material balances to be able to run independently of the DCS for development and testing, the simulations must still have internal pressure and level loops set up to prevent volumes from running dry, overflowing, over pressurizing, or developing flow reversals from pressure gradient reversals, which can lead to fatal numerical errors [1]. Tables must be mapped that transfer control from these internal simulation loops to the DCS loops and initialize the proper controlled variables, set points, and manipulated variables. DCS loops that do not have a counterpart in the simulation still need to have their controller outputs initialized. The use of standard DCS blocks for split-ranged control, velocity limiting, signal characterization, and signal selection makes the proper initialization of external simulations manageable, but still difficult.

The digital twin eliminates the custom programming of control strategies into process simulations and eliminates of the significant simplification and the extra effort, verification, revision, and coordination to ensure fidelity.

The digital twin model has the material balance, component balances, energy balance, charge balance, and kinetics for a particular type of cell culture and product as detailed in Chapter 6. The user can drag blocks for process streams, mass flowmeters, pumps, fans, transmitters, control valves, and actuators into the model, and then wire them up to the bioreactor model. The process information is automatically conveyed from block to block through the "wires" that function as process and signal paths. The flow in each process path is automatically computed, as described in Chapter 6, from pressures set for streams entering or exiting the model, bioreactor pressure, and from intervening pump or fan speeds, valve positions, and pressure drops from flow resistance. It is unnecessary to duplicate level and pressure loops for inventory control and sequences for batch operation or to add interface tables for communication of I/O.

The graphical configuration of process models linked to the control modules eliminates the need to duplicate inventory and batch control or to add interface tables.

Implementing the process simulation as process models in the digital twin offers advantages beyond the elimination of the previously noted issues of duplicating loops and configuring interface tables. Foremost is the digital twin's inherent ability to run the process models at the same real-time execution time multiplier as the control modules. This ensures that the user can set a common real-time multiplier and that the simulation and control system will slow down or speed up in unison. The user can set a real-time execution multiplier of 1/30 to 30 for all modules in the digital twin. When external high-fidelity simulation packages are used, the actual real-time factor may depend on processor loading. This is because of the complexity of the calculations associated with the objectives of process or equipment design. Often during the times of greatest interest, during disturbances or failures, these simulations slow down because the integration step size has decreased for numerical stability and the control system and operator activity have increased. Even if the hook exists between the simulation and DCS real-time multipliers, the DCS is always playing catch-up.

The digital twin addresses the difficulty of ensuring that the control system is running at the same real-time multiplier as a separate process simulation during the times of greatest interest.

The speedup of the process model can be much faster than the real-time multiplier for the control system. The process model speedup is the product of speedup factors for kinetics and the material and energy balances as detailed in Chapter 6. If the kinetic speedup factor is 10 and the material and energy balance factor is 50, the process model speedup is 500. If the control system speedup is 10 that can be set via Figure 5-2b screen, the response times of the instrumentation could be reduced by a factor of 50 to keep the proportional-integral-derivative (PID) controller tuning about the same. However, the process flow ranges of final control elements and flow measurements must be increased by a factor equal to the kinetic speedup.

The process models in a digital twin can be designed to show the proper behavior for failures, start-ups, shutdowns, and batch sequences. This means that the process models can handle zero flows (closed valves and turned-off pumps or fans), empty volumes, non-equilibrium conditions, and imperfect mixing. Sophisticated external (non-digital twin) dynamic process simulations are prone to fatal numerical errors from ill-conditioned matrices used for the simultaneous solution of stream conditions by the pressure-flow solver [2]. These simulations may not only slow down but they may also shut down during extreme conditions. Consequently, the simulation of batch processes often requires specialized software that runs off-line as a single program execution. Functionally, the run conditions are set at the time of execution. These batch process simulations may appear to the user or be cited by the supplier as interactive,

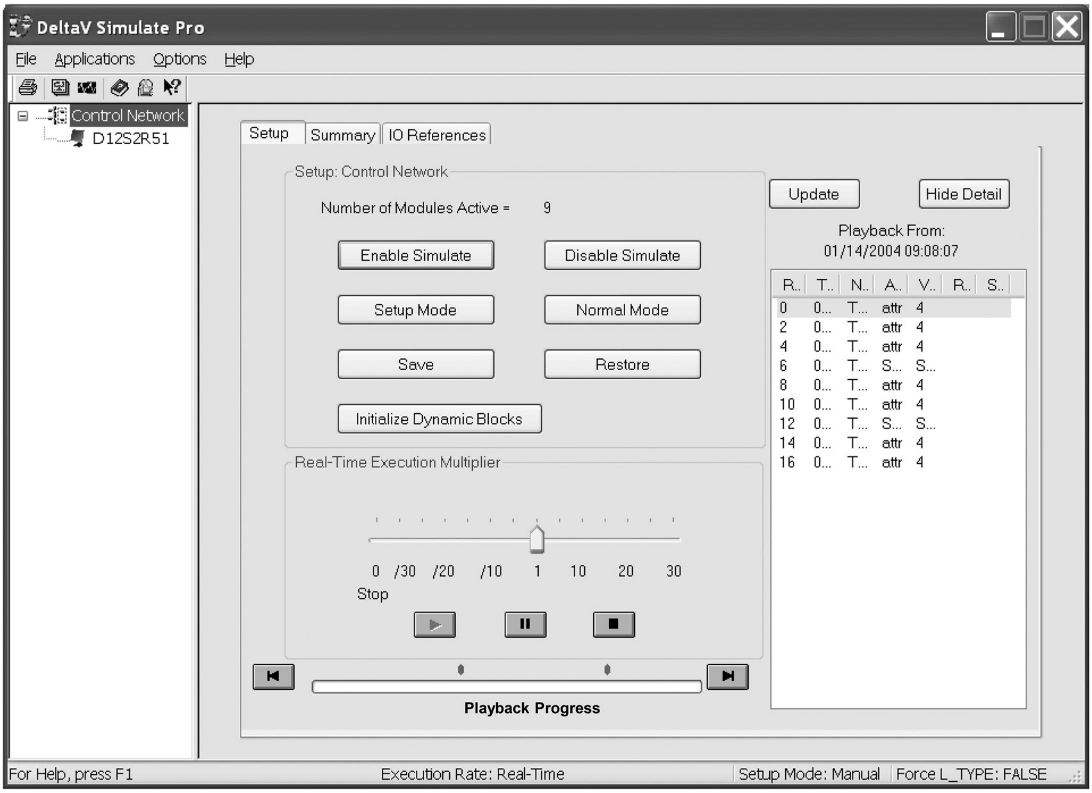


Figure 5-2b. Screen for control of digital twin speed, modes, capture, and playback.

but a change in conditions or parameters requires that the entire program and the batch process be re-executed from start to finish. The ability to include the interaction between the process and the batch and process control strategies is severely limited or impossible. Because these programs are not running in real time or a multiple of real time, it is not possible to pause, restore, or play back parts of a batch.

It is important to verify whether a dynamic simulation can handle the zero flows, empty volumes, and non-equilibrium conditions associated with batch operation.

Most external (non-digital twin) batch process simulation software offered to date is not truly interactive with either the user or the control system.

Process simulations developed by process engineers can predict self-regulating process gains relatively accurately because of the sophistication of the physical properties, thermodynamics, interactions, and equations of state needed for continuous process and equipment design. Process engineers tend to think steady state. Batch process design is noticeably absent in books and courses on chemical engineering.

Consequently, these engineers learn and think in terms of a steady state and equilibrium that is consistent with continuous processes.

Most dynamic real-time process simulations for chemical processes were developed by process engineers with steady-state process and equipment design in mind.

Process models designed by control engineers can be set up to do a better job than process simulations in modeling system dynamics by including the effect of transportation delays, mixing delays, mass transfer rates, kinetics, analyzer cycle times, sensor lags, noise, resolution, and deadband. The result is a closer match to the dead time, time constant, or integrating process gain and a fidelity that is more in tune with the spectrum of uses for control systems [2, 3].

The process simulation dead time is often less than the actual dead time because of the omission of transportation delays, mixing delays, mass transfer rates, kinetics, analyzer cycle time, sensor lags, noise, resolution, and deadband.

The digital twin has the ability to simultaneously stop and start the execution and restore and replay simulated conditions for all the control and the simulation models. Now emerging is the ability to replay actual plant data history files at high speeds to adapt and test the process models without having a connection to the actual plant.

Most simulations used for control system checkout and operator training become obsolete after start-up. The investment is lost. The digital twin offers a better chance of keeping the control system up to date by simply importing the most current configuration and enabling the simulation to better match process changes using the nonintrusive automatic adaptation described in Section 5.4.

External (non-digital twin) simulations used for control system checkout and operator training become obsolete after start-up because they lack the ability to be automatically updated and adapted.

Finally, in a digital twin everything is done in the same configuration environment that is used for the actual control system. The focus can be more on the application than learning the inevitable undocumented features and tips and techniques associated with any new simulation software and interface. The advantages offered by a digital twin are summarized as follows:

- Control system and graphics do not need to be duplicated, emulated, or translated.
- Special data interfaces, tables, and initialization issues are avoided.

- All batch, basic, and advanced control tools available in an actual control system can be readily tried out.
- Controls and simulation can run in unison at the same real-time multiplier.
- Controls and simulation scenarios can be saved, restored, and played back.
- Actual plant data can be played back at high speeds for testing and adaptation.
- Simulations can handle extreme conditions of batch operations and failures.
- Simulations can incorporate dynamics that are important for tuning and performance.
- Controls and simulation can stay up to date and have a longer life cycle.
- Engineers can work in the same environment and focus on the application.

5.3 Spectrum of Uses

The most familiar use of a digital twin is for testing and training. For the checkout of batch sequences and the training of operators, it is important to be able to repetitively and rapidly simulate batch phases. The ability to stop, start, save, restore, and replay scenarios and record operator actions is critical. For first-pass testing and familiarization of sequences and graphics, an automated tieback simulation may be sufficient. To test and learn about the interaction and performance of both control strategies and the process, the higher fidelity dynamics offered by process models is important. It opens the door to upgrading the process and control skills of technology, maintenance, and configuration engineers who support operations.

A process simulation with high dynamic fidelity is important for testing process and control system interaction and performance.

Not well recognized is that the dynamic models used often for training operators as part of an automation project have a much wider utility that today is more important than ever. There is a great opportunity to use the digital twin to maximize the synergy between operators, process control engineers, and the control systems. To start on this path, process control engineers must be given the time to learn and use a digital twin and set up online metrics for process capacity and efficiency. The digital twin offers flexible and fast exploring => discovering => prototyping => testing => justifying => deploying => testing => training => commissioning => maintaining => troubleshooting => auditing => continuous improvement showing the "before" and

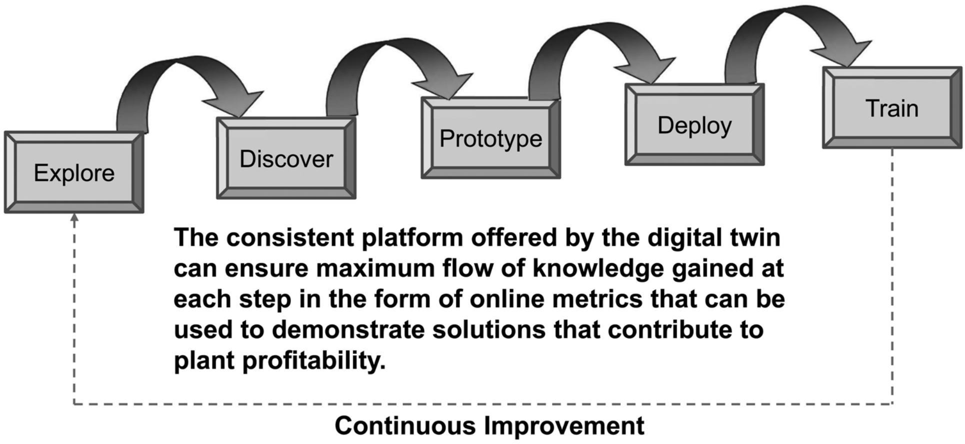


Figure 5-3a. Continuous improvement can become an inherent part of the digital twin maximizing synergy of operational, process, and control system knowledge [10].

"after" benefits of solutions from online metrics. Figure 5-3a outlines the major steps in continuous improvement maximizing innovation [10].

The capability of dynamic models for improving system performance has greatly increased, even though the use has focused mostly on training operators as an automation project nears completion. The digital twin should detail the tasks needed to address difficult situations based on the best operator practices and process knowledge and eliminate the need for special operator actions through state-based control. Advanced process control (APC) and MPC can deal with disturbances and address constraints intelligently, continually, and automatically, with great repeatability. Compare this with what operators can do in terms of constant attention, deep knowledge, and timely predictive corrections considering dead time, multivariable situations, and uncertainty in human behavior. Some operators may do well, but this is not carried over to all operators. Then, of course, an operator can have a bad day. Automation enables continuous improvement and recognition of abnormal conditions by a much more consistent operation. A better understanding by the operator of control system functionality and process performance from online metrics greatly reduces disruptions by an operator unnecessarily taking a control system out of its highest mode and/or making changes in flows. Furthermore, procedural automation can eliminate the manual operations during start-up when risk is the greatest compared to normal fed-batch or continuous operation [10].

Note that while we have singled out operators and process control engineers, the need for knowledge to attain the best performance extends to maintenance technicians, process engineers, mechanical engineers, and information technology specialists. Think of what can be realized if we were all on the same page understanding the process and operational opportunities and the value of the best measurements, valves, controllers, and software [10].

Here is a brief summary of the many possible uses of a digital twin:

- **Cause and effect relationships –** Data analytics can point to many possibilities, but the relationships identified are correlations and not necessarily cause and effect. Also, the data sets used in data analytics are limited to the range of plant operation, which by design may not show the changes possible. Process control is all about changes. We saw this from the beginning in our first course on control theory where all the variables for Laplace transforms and frequency response were deviation variables. We must see the effect of changes in process inputs on the changes in process outputs. A digital twin can make all the changes that you cannot make in a real plant. You can be creative, freeing your mind to make changes, learning the most from mistakes. We can do this all in the safe environment of the digital twin [13]. The resulting data can verify correlations, provide richer data sets for data analytics, identify the dynamic compensation needed in data analytics for the predictions by projection to latent structures also known as *partial least squares* (PLS), and provide a much leaner, faster, safer, and more focused design of experiments (DOE) for the actual plant. Root cause analysis can be greatly improved [10].
- **Interactions and resonance –** The degree of interaction and resonance depend on the dynamics, tuning, and oscillation periods. The digital twin can show all this and much more using Power Spectrum analyzers to track down the culprit if the process and automation system dead times are modeled [10].
- **Valve and sensor response –** The largest source of dead time and oscillations can be traced back in most loops to the valve and sensor response. In particular, most of the valves designed for high capacity and tight shutoff not only have a poor response time but also will cause continuous limit cycles from stiction and backlash [11]. Showing the effect of slow or noisy sensors and on-off valves posing as throttling valves can provide the justification for getting the best measurements and valves. You can show the effect of sensor and valve failures and the value of redundancy and a smooth recovery [10].
- **Process safety stewardship –** The digital twin should be used for extensive and continual testing of all the layers of protection in addition to the detailed design and implementation of the safety instrumented system (SIS). The interplay and performance of layers should be thoroughly evaluated and documented for all the important scenarios. The speed of measurements and valves is often not

thoroughly understood, and the consequences of failures should be analyzed in these scenarios [10].

- **Control system and SIS knowledge –** One of the most difficult aspect of control system and SIS design is recognizing the implications of measurement and final control element repeatability, rangeability, reliability, resolution, and response time. The digital twin enables you to explore and determine the value of the best instrumentation and best PID control and MPC, including the solutions and parameters needed. You must see how well the systems play together and help the operator in abnormal conditions. You can try all kinds of what-if scenarios to see how the system and operator perform [10].

- **Validation and regulatory support –** The validation and regulatory compliance of automation systems in pharmaceutical production requires a large expenditure of time and expertise in automation projects. The digital twin offers the verification of performance needed for confirmation and documentation [10].

- **Checkout of code –** Checking out the code and control sequences before the actual plant start-up enables the control engineer to validate that the code was written as required and designed. It allows debugging for any mistakes made during the coding and validating seen in the coordination of each sequence. The more time the control engineer spends checking out the code and exercising the sequences, the fewer issues he or she will face during the start-up and commissioning of the plant.

- **Process and equipment knowledge –** Most of our deep understanding of the implications of process and mechanical designs has come from running dynamic models. This was particularly true for the specialties of pH and compressor control. For pH control, the extraordinary sensitivity and rangeability and consequential incredible nonlinearity of a system of strong acids and bases leads to strict design requirements in terms of minimizing mixing and injection dead time and seeking better set points, and the possibility of using weak acids or weak bases and conjugate salts to help reduce the steepness and nonlinearity. Today, we have the ability to handle conjugate salts as well as a wide spectrum of weak acids and weak bases by means of an expanded charge balance equation detailed in Chapter 9 [10, 11].

- **Process equipment degradation –** The consequences of fouling of heat transfer surfaces, thermowells and electrodes, loss of catalyst activity, and plugging of filters can be studied and more optimum times for maintenance and cleaning scheduled. Better installation practices and redundancy of measurements can be justified. For pH, three electrodes with middle signal selection provide the best protection against reaction to noise, failures of any type, and fouling [10, 11].

- **Start-ups, transitions, shutdowns, and batch operation –** The design and value of procedural automation and state-based control for fed-batch and continuous processes is best achieved by trying out all conceivable scenarios in a digital twin. The digital twin is also the key for optimizing batch profiles, end points, and cycle times. If an operator claims his actions cannot be automated, there is an even a greater motivation for a digital twin to identify and test the best actions. The automated repeatable best actions reveal the other sources of changes by eliminating the variability from operator response [10, 11].
- **Optimum operating points –** Finally, you can identify and document the benefits of better set points and achieve these set points more reliably and extensively by feedforward control, override control, batch profile control, full throttle batch control, and model predictive control [10].

Before the configuration even starts in the front end of a project, the process models can be used to evaluate control strategies and advanced control tools. In the past, this was done with off-line dynamic simulations. Having ready access to an industrial tool set for both basic and advanced control and simulations that is adapted to benchtop or pilot-plant runs offers the opportunity for rapid prototyping. This can lead to control definitions that have better detail and potential performance. Benchtop or pilot-plant systems with a mini version of the industrial DCS are now available that greatly facilitate the development and scale-up of the control system [6]. Benchtop systems and pilot plants that have all the functionality of the main manufacturing systems are not yet prevalent because the development groups of these types of companies traditionally do not have the expertise (and more importantly the interest) to configure, maintain, and engineer these systems. The digital twin enables a synergy between scientists and control engineers to make the incredible capability of DCS part of the R&D process.

The digital twin can be demonstrated with the industrial batch, basic, and advanced control systems used in the benchtop or pilot-plant system. The process models can be adapted by way of a connection to a benchtop or pilot-plant system or by built-in high-speed playback of process data from experimental runs. An opportune time to take advantage of the PAT initiative is during the research and process development phase. The value of benchtop and pilot-plant systems that have an industrial DCS is significant in terms of project schedule, cost, and effectiveness because the DCS enables process control to be designed into the plant at an early stage and puts the biochemist, process technology, and configuration engineer on the same page [6]. However, the best return on investment (ROI) for PAT is realized by eventually implementing advanced process analysis and control in large-scale manufacturing processes.

In a digital twin, innovative strategies such as effective switching of the controller output can be trialed and tuned. Advanced control tools such as adaptive control, auto tuning, fuzzy logic, MPC, neural networks (NNs), principal component analysis (PCA), and PLS can be demonstrated, adjusted, and evaluated "faster than real time."

A benchtop or pilot-plant bioreactor that has an industrial control system offers significant opportunities for reducing configuration costs, improving communication, developing a digital twin, and prototyping advanced control, but the greatest ROI is realized by implementing PAT in large-scale manufacturing processes.

As a general rule of thumb, five changes are needed for each process input to develop an experimental model of a process output. For MPC, this corresponds to a minimum of five step changes in each process input, at least one of which is held long enough for the process output to reach a steady condition. For NN, PCA, and PLS it means a minimum of five batches of process input in which the respective process input differed from the normal value. Although actual plant operation is obviously the best source of data, the long batch cycle time and the desire to minimize disruptions from the introduction of perturbations severely restricts the amount of useful plant data available for the development of MPC, NN, PCA, and PLS. For example, developing a PCA with four inputs for detecting an abnormal batch would require at least 20 batches with varying inputs and at least five batches with normal inputs. If the batch cycle time is about two weeks, it would take about a year of plant production to have enough data. If you consider that you cannot deliberately optimize the spectrum of variability in the inputs, then it may take several years of production runs to have enough data.

In contrast, perturbations can be automated and introduced to a digital twin running faster than real time so that in 2 weeks there is enough data to identify models for MPC, to train NN, to develop latent variables and discriminant analysis for PCA, and to predict economic variables via PLS. The predictive ability of MPC, NN, PCA, and PLS can then be verified and evaluated by the high-speed playback of previous plant batches.

Conventional PCA assumes that all process inputs other than the ones used for the PCA are fixed. A digital twin running in real time that is synchronized with the actual plant can be used to predict the effect of variations in other process inputs by using the model-based and super model-based PCA algorithms described in Section 8.4 of Chapter 8.

The digital twin can also be used to help explore more optimal operating conditions and investigate what-if scenarios. These scenarios are important for identifying the cause of an abnormal batch. Today, PCA for batch fault detection only identifies a batch as abnormal. Logic needs to be added to diagnose the fault. Fuzzy logic rule sets have been used in conjunction with PCA to provide real-time predictive fault analysis [7]. A digital twin can help develop these rule sets off-line faster than in real time by creating scenarios and evaluating the rule sets through the high-speed playback of previous batches. An online digital twin synchronized with the real plant can be sped up and run to batch completion to predict and analyze abnormal situations based on current batch conditions. A digital twin can also be used to help create the predictive capability of the PLS “Y space” of economic variables from the PCA “X space” of process variables.

A digital twin can eliminate process testing and provides years of process data within a few weeks for advanced control, fault analysis, and performance prediction.

Additionally, a digital twin can be used to provide inferential real-time measurements of important process outputs such as nutrient, biomass, and product concentration. The built-in material balances can be used in conjunction with kinetic models to predict concentrations. These predictions are then delayed so that the values are synchronized with online or lab analysis. The concentration is shifted by a bias that is a fraction of the difference between the inferential and actual measurement, similar to what is done in the feedback correction of an NN used for property estimation [3].

A digital twin can provide faster, more reliable, and smoother measurements of concentration than online analyzers.

The rate of change of these concentrations can be used by MPC to optimize batch profiles as described in Section 4.4 of Chapter 4. A digital twin can also be run “faster than real time” to batch completion to provide an online prediction of key performance indicators (KPIs), such as batch cycle time and yield, based on current batch conditions and inferential measurements. Presently, the understanding of the value of process control in management is decreasing and the emphasis on profitability is increasing. It cannot be taken for granted that practitioners will be given the time and funding to make process control improvements. The retirement of most of the process control experts has exasperated the situation. We are at a critical point in terms of the decreasing rate of innovation and increasing tendency to copy jobs in migration projects. The best way to turn this around is by means of KPI with actual dollars correlated to process control improvements that can be revealed and tested by dynamic first-principle models in a digital twin and made available real time to operations [14].

A key insight is that the myriad improvements can be categorized as increases in process efficiency, capacity, flexibility, and safety. Increases in process efficiency show up primarily as a decrease in the ratio of the running average of raw material mass or energy used to the running average of product mass produced. Increases in process capacity show up as an increased running average of the product mass produced. Process capacity increases can be the result of higher production rates, faster start-ups, better ability to deal with abnormal operation, and greater onstream time. In all cases, the product mass must meet customer specifications. Flexibility shows up as the ability to meet different production rates or different product requirements. Safety shows up as minimizing activations of the SIS besides the obvious metric of minimizing the number of incidents, including near misses [11].

The period for metrics must be large enough to eliminate noise and inverse response and to provide the ability to make decisions based on the objective and process type. For evaluating operator and control system actions, the period is normally the batch cycle time and operator shift for batch and continuous processes, respectively. The period is a month for correlation with accounting metrics. For alerting operators as fast as possible to the consequence of actions taken (e.g., changing the controller set point or mode), the period can be reduced to be as short as six times the total loop dead time. The metrics at the end of a month, batch, and shift are historized for evaluation of plant, batch, and operator performance [11].

There is often a trade-off between process metrics. Increasing the production rate often comes at a cost of a decreasing efficiency. Similarly, changing production rates reduces both process efficiency and capacity because movement to the new process operating point takes time and the product produced in the transition may not meet specifications. Increases in yield (decrease in raw material use) can be taken as an increase in process efficiency if the raw material feed rate is accordingly decreased. There may be an accompanying decrease in the cost of recycle and waste treatment operations. Alternatively, increases in yield can be taken as an increase in process capacity by keeping the raw material feed rate constant. Prolonging a batch can improve yield and thus efficiency, but the lengthening of batch cycle time translates to less batch capacity, particularly as reaction rates or purification rates decline near end point. Time to reach a new set point can be shortened by overdriving the manipulated variable past its final resting value. For batch processes, reaching a new composition, pH, pressure, or temperature set point is often not possible without the overdrive. The process efficiency is reduced during the overdrive, but the process capacity is increased either as a reduction in batch cycle time or an increase in the continuous production rate upon reaching the set point [11].

Especially important is the translation of online metrics to the bottom-line effect on production unit profitability in the plant accounting system. This means benefits must be reported on a monthly basis and presented per accounting format and procedures. Also, obvious but often not addressed is the buy-in by the plant accounting department and plant management. This is best done by real-time accounting (RTA) [11].

The KPI in the digital and actual plant should be converted to dollars of revenue and dollars of cost with a running total for the last month, batch, and shift. Online metrics on process efficiency, often expressed as raw material mass or energy used per unit mass of product, must be multiplied by the production rate at the time and the cost per unit mass or energy used to get dollars cost per unit time. Online metrics on process capacity must have a production rate multiplied by the price of product(s) sold. This is best accomplished by getting an accounting representative to participate in the development and use of the metrics [11].

An opportunity sizing details the gaps between current and potential performance that are estimated by identifying the best performance found on cost sheets and from a DOE, which is most often done in a digital twin due to increasingly greater limitations to such experimentation in an actual plant. After completion of the opportunity sizing, a 1- or 2-day opportunity assessment was led by a process engineer with input sought and given by operations, accounting and management, marketing, maintenance, field and lab analyzer specialists, instrument and electrical engineers, and process control engineers. Marketing provided the perspective and details on how the demand and value of different products is expected to change. This knowledge was crucial for effectively estimating the benefits from increases in process flexibility and capacity. Opportunities for procedure automation and plantwide ratio control making the transition from one product to another faster and more efficient were consequently identified. Agreement was sought and often reached on the percentage of each gap that could be eliminated by potential process control improvement (PCI) proposed and discussed during the meeting. A rough estimate of the cost and time required for each PCI implementation was also listed. The ones with the least cost and time requirements were noted as *quick-wins*. To take advantage of the knowledge and enthusiasm and momentum, the quick-wins were usually started immediately after the meeting or the following week. Figure 5-3b shows this PCI opportunity assessment process [11]

An improvement in one unit's operation performance may have a negative impact on another unit's operation. For example, a decrease in steam generation from a waste furnace from more efficient operation (less waste production) decreases the steam available for power generation. Rapid changes in raw materials or energy usage may

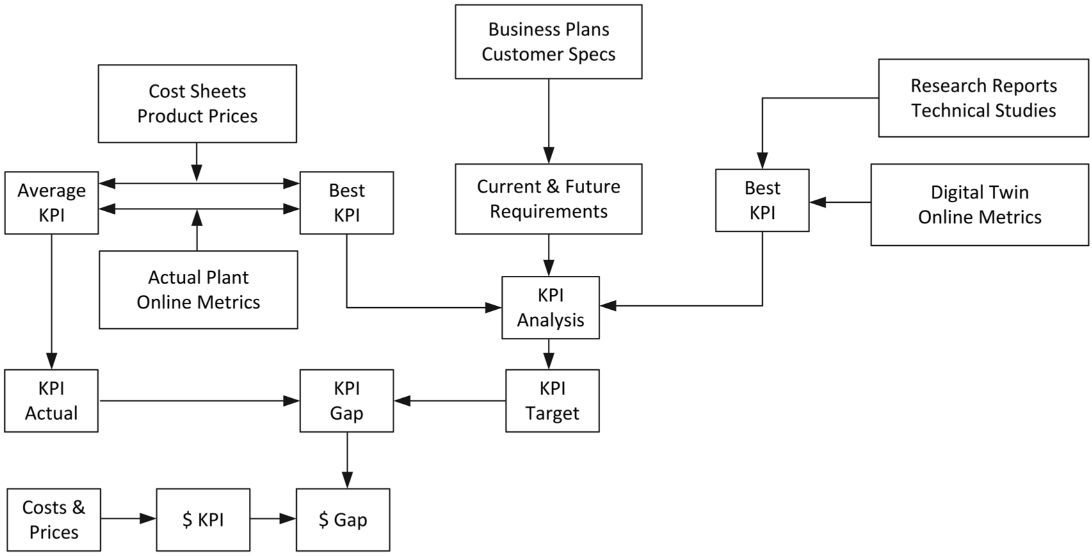


Figure 5-3b. The digital twin enables experimentation and verification to identify and achieve the best benefits from PCI [10].

upset the headers and systems that supply other systems. The online metrics should be extended to all unit operations affected to quantify the trade-offs [11].

The raw material and energy used does not have an immediate effect on product produced. Transportation and mixing delays, and time constants associated with volumes and mass transfer and heat transfer rates, prevent the start of the effect on the product and slow down the completion of the effect. Transportation delays for plug flow volumes and time constants for back mixed volumes are both inversely proportional to flow. There are dynamics and rangeability limitations associated with measurements and valves that are often not recognized and included. Also, the time to reach and settle at a new set point depends on the ability to identify a dynamic model and tune the controller to meet the best objectives [11].

The PCIs must be able to address unknowns and disturbances but also be sufficiently understood by operations to build their confidence to keep them in their highest control mode. Operators are fundamentally reluctant to relinquish manipulation of process inputs and operating points if the benefits are not seen or understood. Operators may also be initially opposed to indications and comparisons of their performance. For these reasons and many more, it is imperative that the performance metrics exhibit inappreciable noise and be consistently representative of changes in process performance. Using results in a positive way allows operators to take more ownership of process performance. The adage "if you don't measure, you can't control it" is especially important here [11].

The digital twin can speed up the benefits gained from PAT by offering users the ability to use process and end point monitoring and control, continuous improvement, and knowledge management tools in an integrated and accelerated manner. The uses of the digital twin are illustrated in the functional overview provided by Figure 1-4b in Chapter 1 and are summarized as follows:

A digital twin can accelerate the benefits of a PAT initiative.

- Testing configuration and process interactions
- Educating operators, technicians, and engineers in process and control
- Rapid prototyping of innovative control strategies and advanced controls
- Evaluating tuning settings
- Identifying MPC models
- Developing latent variables and reference trajectories for PCA
- Developing logic for fault analysis by PCA
- Predicting abnormal situations online
- Developing and testing NN and PLS models
- Making inferential real-time measurements of important concentrations
- Optimizing batch profiles
- Predicting batch KPIs, such as cycle time and yield, and PLS economic variables

Figure 5-3c shows the functional value of a digital twin, highlighting the bidirectional flows of control system and process/equipment knowledge of process control and analysis tools including online performance metrics for greater analysis, dealing with kinetics multiplicative effect in Chapter 6, and justification of improvements. The “two-way” knowledge flow is the key to improving the process/equipment and the control system besides in addition to the dynamic model and data analytics. As the fidelity of the dynamic model increases, opportunities arise for these tools to get results from the digital twin that can be used in the actual plant. The dynamic model can be run faster than real time with the tuning corrected by applying the speedup factors. New control functionality can be developed and included in the dynamic model for evaluation. If online metrics show significant improvements in control and process performance, the functionality prototyped can be added as new blocks or as improvements to existing blocks in the DCS [11].

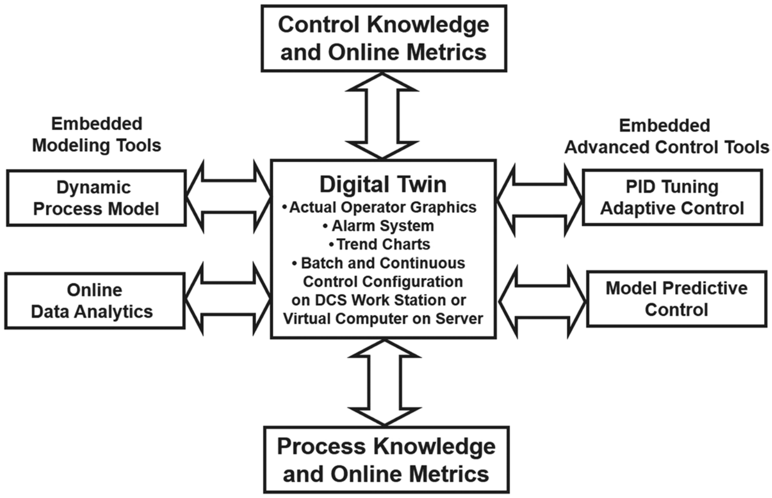


Figure 5-3c. The two-way flow of knowledge in the digital twin between tools, models, and the actual control system in the digital twin is the source of increasing synergy of knowledge between the process, control system, engineers, technicians, and data scientists [11].

5.4 Implementation

Probably the biggest obstacle to implementing a digital twin is the users' lack of knowledge of the kinetic equations that are needed to calculate the growth rates and product formation rates as a function of batch operating conditions, such as temperature, pH, DO, and substrate concentration. As detailed in Chapter 6, generalized easy-to-parametrize kinetic models have been developed that enable users to readily match historized plant batch profiles. This section discusses how an MPC can be set up to home in on the values of these parameters if they are in the ballpark. If the kinetics are unknown, NN may be able to predict kinetic parameters, terms, or rates from inferential and actual measurements. The growth rates and product formation rates are included in the net calculation of the rate of change of biomass, nutrient, and product mass in the material balance. They are then integrated to get a new accumulation or concentration of the component, as described in Chapter 6. Online calculations of the OUR and carbon dioxide production rate (CPR) as NN inputs enhance the predictive capability of the NN. The result is a hybrid first-principle and NN model [8]. In general, batch time should not be an NN input because the predictions are undesirably dependent on batch time. An NN without batch time as an input will be able to predict ethanol and cell mass [9].

Breakthroughs in kinetic models enable easily matching the batch profiles seen on trend recordings without a need to disclose proprietary R&D.

The OURs and CPRs can be used in lieu of kinetic equations, but it may be difficult to differentiate between the consumption and evolution associated with biomass growth vis-à-vis product formation rates.

If the equations are unknown, an NN may be able to predict the individual growth and production rates if there are online measurements or frequent lab measurements of biomass, product, and substrate concentration [9]. A benchtop or pilot-plant system offers the best opportunity for developing an NN. To minimize the investment, a skid of analyzers could be alternately connected to bench or plant systems to generate the test data needed for NN training and validation. This approach has proved to be successful enough for state governments to accept online NN as a permissible alternative to online analyzers for combustion emissions monitoring (CEM) systems.

Online broth or media concentration measurements and NNs can predict growth rates and product formation rates without a prior knowledge of kinetics.

Composite blocks for the bioreactor, streams, final elements, and transmitters are inserted into the digital twin process model and wired up. External input and output parameter blocks are added and connected to the final element inputs and transmitter outputs, respectively. The control system configuration is imported, and the analog input blocks in the virtual control system are switched to the simulation mode. The paths of the external inputs for the final elements are chosen via a browser to be the outputs of the analog output blocks of the virtual control system.

Similarly, the paths of the external output parameters for the transmitters are selected to be the simulation inputs of the analog input blocks of the virtual control system. Analog outputs and analog inputs of loops that are not modeled can be simulated by simple tie-backs in which the analog output is passed through signal characterizer, dead time, and filter blocks and then connected to the simulated input of the analog input block.

The concentrations, pressures, and temperatures of each inlet and outlet stream; the flow rate and pressure rise of each pump as a function of speed; the flow coefficient of each valve as a function of position; the agitator pumping rate as a function of speed; and the initial inventory, concentrations, temperature, and pressure of the bioreactor are then set. NN and PLS models that are developed from benchtop experiments or pilot-plant runs are added as necessary to enhance the kinetics described in Chapter 6. Control valve deadband and resolution are set in the *actuator* blocks, and measurement delays and lags are set in the *transmitter* blocks.

Benchtop systems offer the best opportunity for developing and validating first-principle, PLS, or NN models.

The batch sequence, basic process control system, and simulation are run in an off-line mode, and the simulation parameters are manually adjusted to match up the virtual with the real plant batch profiles of uncontrolled process outputs and manipulated process inputs. For example, consider how the model may be adjusted for better fidelity for DO control. A kinetic parameter for the oxygen limitation effect is adjusted to match up oxygen uptake rates between the virtual and actual plant. A Henry's law coefficient is then adjusted that determines the equilibrium DO (driving force for oxygen transfer) to match up pressures. Finally, an oxygen mass transfer coefficient is adjusted to match up airflows.

When representative digital twin batch profiles are obtained, the processor loading and profile alteration are checked as the real-time execution multiplier is increased, and the maximum practical speed is noted. Key model parameters are chosen as the manipulated variables, and associated process measurements are chosen as the controlled variables of the MPC for adaptation of the digital twin. An automated test sequence is then run for the MPC at the highest possible speed off-line, and the MPC models are identified and visually checked as reasonable in terms of the direction and relative magnitude of the effect [1, 4].

The digital twin is then connected to the actual plant in a read-only nonintrusive setup. The digital twin modes, set points, and batch phases come from the actual plant. The set points (targets) of the MPC's controlled variables for adaptation are externally referenced to the key respective measurements of the actual plant, as shown in the overview provided by Figure 1-4b in Chapter 1 and in the detailed signal paths outlined by Figure 5-4a. Because the relationship between model parameters and key process variables is generally nonlinear, an NN can be used to enhance the future trajectories of an MPC. Alternatively, a NN could be trained to adjust model parameters on a one-by-one basis. Because NNs are good at interpolation but not extrapolation, it is important to verify whether the inputs are within the training set range of values. The MPC offers a multivariable solution to adaptation of dynamic models [1, 4].

The use of standard tools, such as MPC, PLS, and NN, are preferred over custom and special reconciliation methods so the focus is on the application rather than the tool development and the results are more maintainable and therefore sustainable. Regardless of which technique is used, the model parameter values must be restricted to be within a practical range.

The MPC, PLS, and NN for adapting the digital twin can be developed off-line and then tested in a read-only mode.

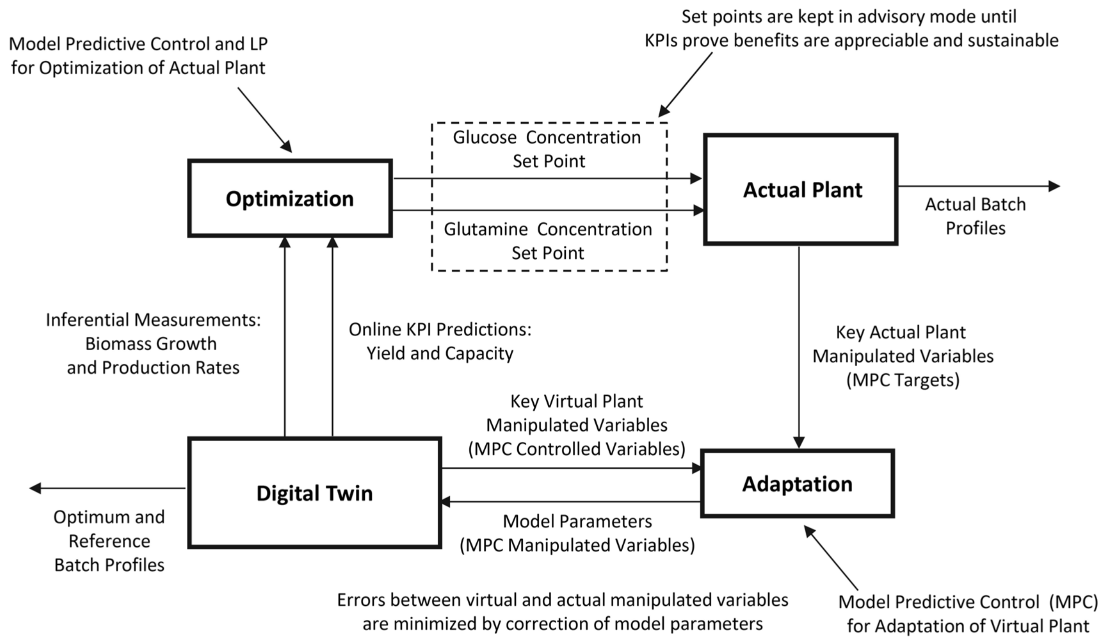


Figure 5-4a. Nonintrusive automated adaptation of model parameters to match manipulated variables with potential future optimization based on improvement in KPIs.

A mass spectrometer can be used to measure the oxygen concentration in the off-gas and a material balance can be used to measure the net difference of oxygen flow going in and coming out of the media to compute the OUR. If the DO concentration is not constant, the average rate of change of DO multiplied by the volume should be included in the calculation.

The MPC targets and controlled variables are the actual plant and digital twin manipulated variables (e.g., flows), respectively, in Figure 5-4a. That figure shows how to adapt a bioreactor model to provide more accurate inferential measurements of biomass growth and production rates that are slopes of the batch profiles for cell and product concentration. The MPC manipulated variables are model parameters (e.g., kinetic parameters) that affect these profiles [11].

The MPC can be identified and tested without the digital twin connected to the actual plant. The MPC development for adaptation is much faster than that of an actual plant MPC. This independence enables extensive adaptation and gaining associated knowledge of process relationships without any disruption to the plant [11].

The adaptation is done without affecting the actual plant because the plant's manipulated variables are being read by but not changed by the digital twin. It is critical that the digital twin have the same set points and tuning settings as the actual plant, and that the digital twin is started with controller outputs initialized to match the actual plant.

The optimized set points from MPC with inferential measurements of growth and production rate are done in an advisory mode not affecting the actual plant. Not shown in Figure 5-4a is that an MPC is run in automatic mode in another digital twin that is a duplicate of the adapted digital twin to generate and study the optimized set points. The set points are only eventually used to automatically optimize the plant if they prove more accurate, beneficial, and reliable per KPIs than an MPC with inferential measurements computed from online and at-line analyzers.

Figure 5-4b shows how an MPC was used to manipulate the Henry coefficient for oxygen to match up a change from 7700 to 7500 kPa per kmol/m^3 (intent is to convey an overview example and not the specifics of test results). This brought the controlled variable of the digital twin head pressure to its set point (target), which is the actual plant head pressure. The speed of adaptation is set by the penalty on move (PM) for the model parameter.

After adaptation, the digital twin is run off-line faster than in real time to prototype new control strategies, PCA, fault detection, and abnormal situation prediction.

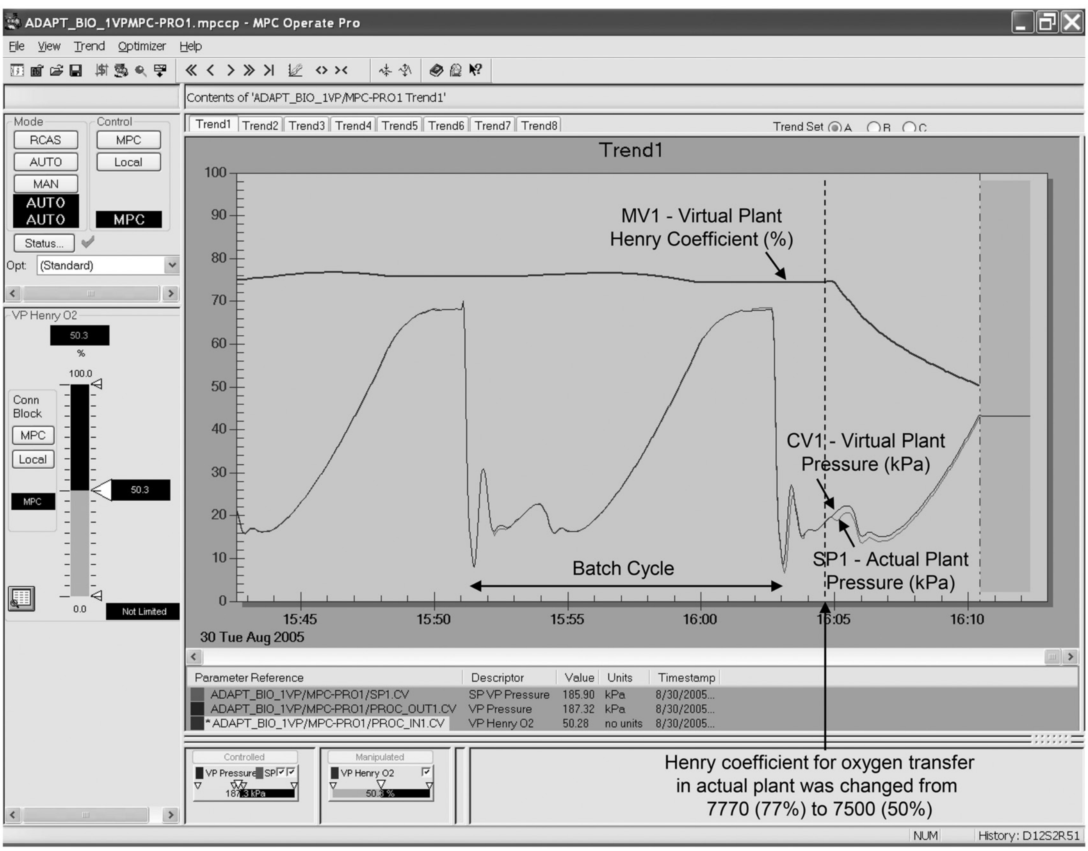


Figure 5-4b. Adaptation of Henry coefficient to match actual plant pressures.

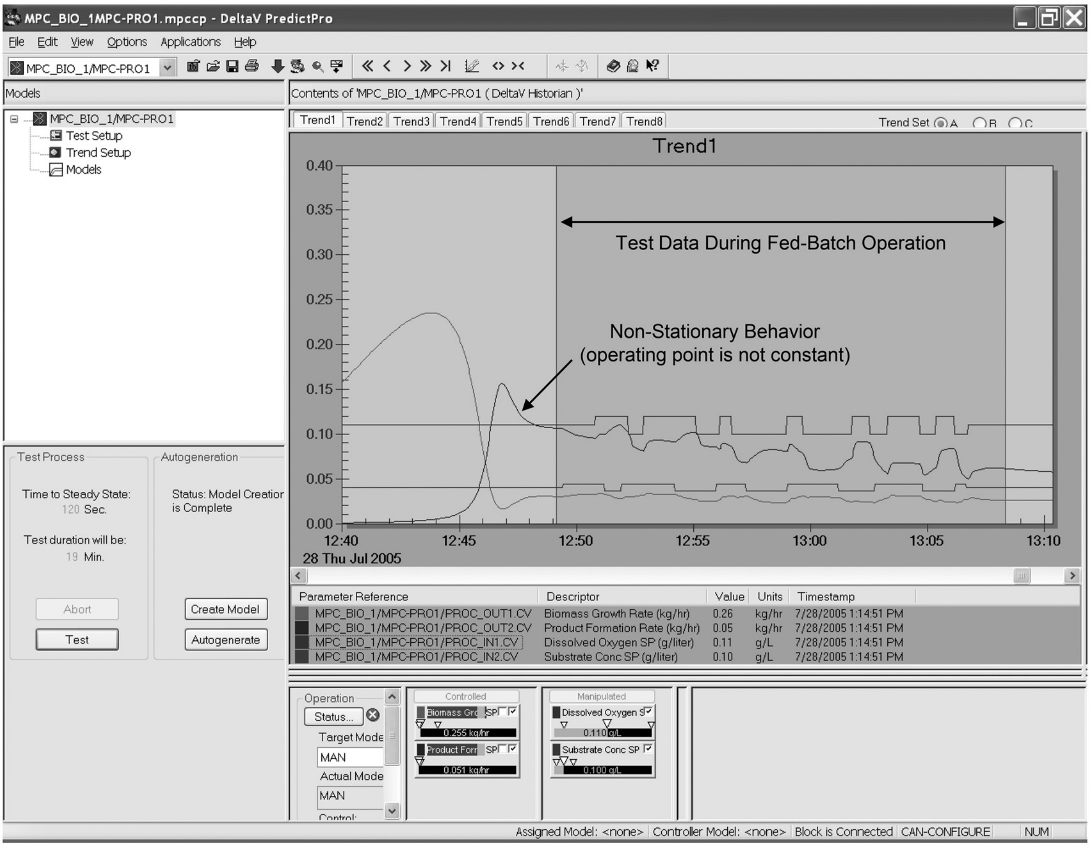


Figure 5-4c. Automated high-speed test of digital twin for model identification.

Inferential variables are computed, such as biomass growth rate and product formation rate, and the automated tests, such as those shown in Figure 5-4c (intent is to convey an overview example and not the specifics of test results), are run to identify the models of these rates versus various process inputs, such as DO and substrate concentration or feed. Even though the process operating point is moving (nonstationary) because it is a batch operation, the software is able to identify the models for optimizing growth and production rates. Inferential measurements of various concentrations are used as needed to fill in the gaps between lab measurements.

After innovative control, advanced diagnostics, and MPC applications have been prototyped and documented, they can be implemented on benchtop or pilot-plant systems. If the digital twin provides inferential measurements, it is run in real time and synchronized with these systems. After verification and validation, and accounting for scale-up factors, these applications and the digital twin can be installed for plant trials. Initially, these systems run only in the monitoring mode. When process inputs are finally manipulated, they are initially restricted to a very narrow range.

The expertise and time required for using and supporting external high-fidelity process simulation software no longer exists in the process industry, except in some large petroleum and chemical companies. The modeling of mass transfer for DO and CO_2 modeling is straightforward, and recent breakthroughs in kinetic models described in Chapter 6 make modeling of growth and product formation rates much easier than in the past.

Excluding kinetics, the thermodynamics needed for models of bioreactors are generally simpler than those needed for chemical reactors.

The charge balance is critical for computing the pH, which is important for the kinetics. Process simulations in the literature for bioreactors generally use empirical relationships for pH that do not show the effect of alternative operating conditions and upsets. Chapter 9 shows how to set up the charge balance from the component balances.

A charge balance is needed to properly simulate pH.

5.5 Conclusion

The digital twin previously known as virtual plant opens our minds to much greater possibilities by exploration and discovery without interfering with plant operations. We develop a much deeper understanding of the process and control system [5, 12]. Consequently, we can exercise much greater creativity, which translates to increasing process capacity and efficiency with benefits that are measured and documented. This leads to personal advancement and freedom for innovation. The authors of this book look forward to mentoring everyone on this path to achievement and recognition.

5.6 Digital Twin Best Practices

Here are best practices that help you use the digital twin not only for operator training but also for improving process performance and time to market.

1. Implement a digital twin that requires no changes to the DCS alarms, graphics, or logic.
2. Use first-principle process models with energy and material balances for all processes and charge balances for pH with representative physical properties and components.
3. For pH models, add a small but significant concentration of carbonic acid from exposure to atmospheric CO_2 and conjugate salts so that the titration curve

generated matches lab results, particularly in terms of the pH range and titration curve slope in the control region as detailed in Chapter 9.

4. Use an actual download of an automation system (not an imitation by emulation) that enables free exchange between the digital twin and the actual system of
   - the configuration, alarms, and APC tools (e.g., adaptive control, model predictive control, and data analytics) by exporting and importing modules and models, and
   - the historian and graphics by copying and pasting files.
5. Achieve at least medium fidelity by adjusting model parameters so that operating points sufficiently match process variables and flows on accurate process flow diagrams (PFDs).
6. Use a variable dead time block to simulate mixing and injection delays.
7. Ensure the process model can show the dynamics in start-up and transitions.
8. Have process engineers evaluate process operating conditions and response.
9. Include automation system dynamics such as the deadband and resolution limits in control valves (e.g., backlash-stiction) and in VFDs (e.g., I/O cards, speed sensor, setup), sensor transportation delays, sensing element lags (e.g., thermowell), transmitter lags (e.g., damping), wireless update rate, and analyzer cycle time and analysis time.
10. Develop, test, and document scenarios including transitions, start-up, and abnormal conditions (e.g., mechanical and automation system failures).
11. Develop and use online metrics for process efficiency and capacity.
12. Analyze and improve alarm management and the human machine interface using Center for Operator Performance and ISA standards for guidance.
13. Detail best operator responses during start-ups and to abnormal situations (e.g., equipment, valve, or sensor failures) for procedure automation (state-based control).
14. The digital twin can provide methods and practical considerations in online metrics computations, be adapted to improve fidelity by manual and automated adjustment of model parameters to match up virtual and actual flows, and be used to discover and realize process control improvement opportunities.

15. Ensure the operator training system (OTS) is continually updated and available for refreshing and enhancing operator performance.
16. Use OTS to educate technicians and engineers in maintenance and in automation, mechanical, and process technical support to improve synergy with operations. Use a multidiscipline team that includes plant operations. If the models will be deployed online, engage the operators early in the cycle. Start with your business objectives. What is the problem to be addressed, and how will you measure success in accomplishing these objectives?
17. Use a high-fidelity digital twin to develop online process metrics.
18. Synchronize process inputs with process outputs for online metrics by the individual use of dead time and filter blocks for synchronization time (e.g., total dead time and open-loop time constant).
19. Use running averages of raw materials and energy and resulting product for metrics.
20. Use different periods for running averages to show metric short- and long-term effects.
21. Improve metric signal-to-noise ratio by a filter time less than 1/5 synchronization time.
22. The shortest period for process evaluation should be greater than twice the synchronization time. Additional metrics to alert operations can have shorter periods.
23. A period equal to shift time shows the performance of operators, a period equal to the batch cycle time shows the performance of batch operations, a period equal to a month helps to predict effects on accounting cost sheets, and a period as short as six total loop dead times can alert the operator to consequences of actions taken (e.g., changing set points or modes) provided the signal-to-noise ratio is sufficient.
24. Involve accounting, plant management, and marketing in the development and use of online metrics for real-time accounting.
25. Use a digital twin for faster and more effective metrics and PCI implementation.
26. Use digital twin to reduce plant testing required for metrics and PCI.
27. Develop an MPC to adapt the digital twin without connection to the actual plant.

28. Use actual plant set points and controller tuning settings adjusted for model speedup in the digital twin.

29. To adapt the digital twin, adjust the model parameters to match the manipulated process inputs by the MPC and PID loops and noncontrolled process outputs.

30. To reduce the amount of identification needed in the actual plant for online process metrics synchronization, add injection and mixing delays and automation system dynamics to the digital twin for the development of metrics.

References

1. Gregory K. McMillan and Robert Cameron, *Models Unleashed: Virtual Plant and Model Predictive Control Applications* (Research Triangle Park, NC: ISA [International Society of Automation], 2004).

2. Walt Boyes, Dan Hebert, Larry O'Brien, and Gregory K. McMillan, "The Light at the End of the Tunnel Is a Train (Virtual Plant Reality)," *Control* (August 2005).

3. Gregory K. McMillan, "Process Modeling," in *A Guide to the Automation Body of Knowledge*, 3rd ed., Nicholas P. Sands and Ian Verhappen, eds. (Research Triangle Park, NC: ISA [International Society of Automation], 2018).

4. Terrence L. Blevins, Gregory K. McMillan, Willy K. Wojsznis, and Michael W. Brown, *Advanced Control Unleashed: Plant Performance Management for Optimum Benefits* (Research Triangle Park, NC: ISA [International Society of Automation], 2003).

5. Michael Mansy, Gregory K. McMillan, and Mark Sowell, "Step into the Virtual Plant," *Chemical Engineering Progress* (February 2002): 56–61.

6. *BioNet Bioreactor Control System*, Broadley-James Bulletin L0016B (2003).

7. Richard J. Fickelscherer, Douglas H. Lenz, and Daniel L. Chester, "Fuzzy Logic Clarifies Operations," *InTech* 52, no. 10 (October 2005).

8. R. G. Silva, A. J. G. Cruz, C. O. Hokka, R. L. C. Giordano, and R. C. Giordano, "A Hybrid Feedforward and Neural Network Model for the Cephalosporin C Production Process," *Brazilian Journal of Chemical Engineering* 7, no. 4-7 (December 2000).

9. K. Pramanik, "Use of Artificial Neural Networks for Prediction of Cell Mass and Ethanol Concentration in Batch Fermentation Using Saccharomyces cerevisiae Yeast," *IE (I) Journal – CH* 85 (September 2004).

10. Gregory K. McMillan and Martin E. Berutti, "Virtual Plant Virtuosity," *Control* (August 2017).

11. Gregory K. McMillan and P. Hunter Vegas, *Process/Industrial Instrumentation and Controls Handbook*, 6th ed. (New York: McGraw-Hill Education, 2019).

12. Christopher Stuart, Rehman Fazeem, and Gregory K. McMillan, "Improve pH Control," *Chemical Processing* (March 2017).

13. Gregory K. McMillan, and Christopher Stuart, "Simulation Breeds Innovation," *Control* (November 2019).

14. Gregory K. McMillan, "The Future Is Here," *Control* (November 2020).

Contents

Acknowledgments xiii

Preface xv

About the Authors xvii

Chapter 1 Opportunities 1
1.1 Introduction 1
Learning Objectives 4
1.2 Analysis of Variability 4
Process Inputs and Outputs 7
Field and Laboratory Measurements 8
Experimental Models 9
Dynamic First-Principle Models 12
Manual, Sequential-Batch, Fed-Batch, and Continuous Operation Modes 13
1.3 Transfer of Variability 16
Discrete Process Actions 17
Feedback Control 18
Selecting MSPC, PLS, and ANN Inputs 20
Setting Up Feedback Loops 21
1.4 Online Indication of Performance 24
1.5 Optimizing Performance 27
1.6 PAT 28
Origins of PAT 28
FDA Initiative 28
Business Drivers 29

PAT Tools 30
1.7 Insights 30
1.8 Best Practices 32
References 32

Chapter 2 Dynamics 35
2.1 Introduction 35
Learning Objectives 36
2.2 Performance Limits 36
Impact of Loop Dead Time 36
Impact of Controller Tuning Settings 38
Sources of Dead Time 43
Limit Cycles. 47
2.3 Self-Regulating Processes. 47
2.4 Integrating Processes 52
2.5 Wireless Devices and Analyzers 55
2.6 Best Practices 59
References 60

Chapter 3 Basic Control 61
3.1 Introduction 61
Learning Objectives 62
3.2 Key Process Measurements 62
3.3 Control Valves and VFDs 64
3.4 PID Fundamentals 67
3.5 PID Form and Structure 72
3.6 PID Tuning 73
Nomenclature for PID Gain and Reset Time Calculations. 78
3.7 PID Options 83
3.8 PID Performance 84
Nomenclature for Accumulated Error Metric 87
Nomenclature for Peak Error Metrics 91
Nomenclature for Rise Time Metrics. 94
3.9 Control Strategies. 96
3.10 Best Practices 100
RTD. 100
pH 101
Coriolis. 102
Valves and VFDs. 103
PID Controller. 105
Bioreactor Control. 107
References 107

Chapter 4 Model-Predictive Control 109
4.1 Introduction 109
Learning Objectives 110

4.2 Capabilities and Limitations 111
4.3 Multiple Manipulated Variables 119
High- and Low-Cost MV 120
Coarse and Fine MV 123
4.4 Optimization 127
4.5 Mammalian Cell Bioreactor Optimization 135
4.6 MPC Best Practices 137
References 140

Chapter 5 Digital Twin 143
5.1 Introduction 143
Learning Objectives 144
5.2 Key Features 145
5.3 Spectrum of Uses 151
5.4 Implementation 162
5.5 Conclusion 168
5.6 Digital Twin Best Practices 168
References 171

Chapter 6 First-Principle Models 173
6.1 Introduction 173
Learning Objectives 174
6.2 Modeling Challenges 174
6.3 Modeling Opportunities 175
6.4 Modeling Breakthroughs 177
6.5 General Form of Kinetics 178
6.6 Media Ordinary Differential Equations with Speedup Factors . . . 179
6.7 Concentration and Flow of Charges, Manipulated and Recycle Streams 182
6.8 Specific Growth Rate Equations with Bias and Gain Terms 183
6.9 Cell Death (Lysis) 186
6.10 Specific Product Formation Rate Equations with Bias and Gain Terms 187
6.11 Product Consumption and Degradation 188
6.12 Specific By-product Formation Rate Equations with Bias and Gain Terms 188
6.13 Utilization Rates 189
6.14 OUR and Carbon Dioxide Production Rate 190
6.15 Agitation and Sparge 190
6.16 Dissolved Gas ODEs with Speedup Factors 193
6.17 Parameters and Variables 194
6.18 Best Practices 202
References 204

Chapter 7 Analytical Technologies 207
7.1 Introduction 207

Learning Objectives . . . 208
7.2 Overview . . . 208
7.3 DO . . . 208
7.4 Dissolved Carbon Dioxide . . . 210
7.5 Turbidity . . . 211
7.6 Dielectric Spectroscopy . . . 213
7.7 Near-Infrared Spectroscopy . . . 214
7.8 Mass Spectrometers . . . 214
7.9 BioProfile FLEX2 . . . 215
7.10 Liquid Chromatographs . . . 217
7.11 Best Practices . . . 218
References . . . 218

Chapter 8 Data Analytics . . . 221
8.1 Introduction . . . 221
Learning Objectives . . . 222
8.2 PCA Background . . . 223
Short History of PCA . . . 223
Simple Example Applying PCA . . . 223
PCA Algorithm . . . 234
Different Types of PCA and Applications . . . 238
8.3 Multiway PCA . . . 238
Reference Trajectory . . . 239
PCA of a Batch Process . . . 240
Batch Data Synchronization . . . 244
8.4 Model-Based PCA . . . 246
Algorithm of MB-PCA . . . 246
Super Model-Based PCA . . . 247
8.5 Fault Detection . . . 250
Bioreactor Simulation . . . 250
PCA Application . . . 251
8.6 Data Analytics Best Practices . . . 257
References . . . 259

Chapter 9 Models to Improve Operator, Automation, and Process Performance . . . 265
9.1 Introduction . . . 265
Learning Objectives . . . 267
9.2 Overview . . . 267
9.3 Bountiful Boundaries . . . 268
9.4 Data Drives Dynamics . . . 269
9.5 Digitizing the Plant for Process Performance . . . 270
9.6 Titration Curve Modeling . . . 272
9.7 Equipment Modeling . . . 275
9.8 Sparge Modeling . . . 277
9.9 Kinetics Modeling . . . 279

9.10 Instrumentation Modeling. 284
9.11 Speedup 296
9.12 Performance Monitoring 297
9.13 Generation and Fitting of Profiles. 298
9.14 Simplifications and Practical Solutions 300
9.15 Best Practices 302
References 304

Appendix A: Automation System Performance Top 10 Concepts. 307

Appendix B: Bioprocess Biology 323

Appendix C: Enhanced PID Controller for Wireless and Analyzer Applications 335

Appendix D: Modern Myths. 355

Appendix E: Enzyme Inactivity Decreased by Controlling the pH with a Family of Bézier Curves. 357

Appendix F: First-Principle Process Relationships 369

Appendix G: Gas Pressure Dynamics 387

Appendix H: Charge Balance to Model pH. 389

Appendix I: Interactive to Noninteractive Time Constant Conversion 399

Appendix J: Jacket and Coil Temperature Control. 403

Appendix K: PID Forms and Conversion of Tuning Settings. 409

Appendix L: Liquid Mixing Dynamics 417

Appendix M: Mammalian Bioreactor Model 421

Appendix N: Debottlenecking Using Sensitivity Analysis 427

Bibliography 437

Index. 449