Digital Twins Demystified:
Practical Applications of Process Simulation

Kyle Cheatham, EIT

Clinton Duewall, PhD

Luke Addington, PE

Bryan Research & Engineering, LLC

Bryan, TX, USA

Abstract

Recent developments in data systems and historians have allowed processing facilities to collect and store more data than ever before. While the data itself provides a record for facility operations, measurements alone do not necessarily provide the information needed to understand process performance, health, or limitations. Process simulation has been used for over 50 years to apply fundamental engineering principles to process data, providing additional engineering information and a deeper understanding of the inner workings of a facility. This paper examines how facility measurement data and process simulation can be combined to create digital twins that provide actionable engineering information and improve process insight, decision-making, and engineering workflows.

Background

Usage of the term “digital twin” is skyrocketing these days. The term has been used so broadly that to many it has become difficult to define. The most basic definition is a digital representation of a physical object or process. While it may sound new, that definition of a digital twin is something that has been around for quite some time. As far back as the 1960s and the Mercury Program, NASA used digital twins to train astronauts prior to launch, simulating the exact experience they would expect using both physical and digital representations of the spacecraft [1]. In the process industries, the use of digital twins by this definition is even older, dating back to the 1950s when the first computer simulation of a single process unit was created using punch cards [2]. In each case, you have the definition of a digital representation of a physical object.

Recently, the general understanding of the digital twin has evolved—especially in the process industries. Here, a digital twin is more often meant exclusively as a digital representation that incorporates modern data gathering systems; a marrying of the physical and virtual space with information flowing between the two [3]. Based on this definition, a simple model itself is not a digital twin. The model *must* bring in the operating data from the process it represents. This is the general definition that will be used in this work, which can be simplified as a process that [4]:

i. is in operation
ii. incorporates real-time data from the physical process
iii. updates and synchronizes continuously based on that data

It is understood that not everyone will agree with this definition, and some additional qualifiers and categories will be explored further. Regardless of other caveats, digital twins fundamentally provide

similar value of improved workflows, increased understanding of the physical system, and better decision-making. Whether these additional qualifiers are applied or not depends on the specific tasks and outcomes that are desired. Peeling back the mystery can provide clarity to engineers about what a digital twin is, how it works, and how they can apply it to their own work. Similarly, understanding how today's digital twins integrate with data systems will provide insight into the efforts required to bring the technologies online in a useful way.

Digital Twin Categories

Digital twins may be placed in two basic categories depending on the nature of the calculations or digital representation themselves.

Data Driven Digital Twins

One such category is *Data Driven,* where neural networks, statistical methods, and simple regression may be used to create functions that represent the real world. The advantage of such models is their simplicity; all that is needed is an abundance of historical data to tune the model. There is no need to understand the actual causation of the inputs to the outputs in the model. However, that lack of causality renders extrapolation, excursions from normal operations, and significant changes to the process lethal to the Twin's accuracy. To overcome this, data driven models must use other methods to artificially perceive the continuously changing relationships between each measured variable. In addition to their lack of causality, the digital twin requires that sufficient data quality and volume already exist for training or tuning.

Physics Based Digital Twins

Physics Based models, on the other hand, use underlying first principles that represent the causation and are typically preferred when available. The process simulators that midstream engineers are most familiar with are physics-based models. These simulators rely on the fundamentals of chemical engineering such as thermodynamics, kinetics, fluid mechanics, and mass and heat transfer to predict outcomes based on specific inputs. In this work, physics based digital twins will be discussed exclusively, and it assumed that the reader is generally familiar with this knowledge.

Static and Dynamic Models

Of the additional categories or qualifiers that are applied by some to the definition of digital twin, a common one is that the digital twin must be "dynamic" rather than "static" [5]. Unfortunately, these crossover terms from other industries can cause confusion for midstream engineers. It may be easily mistaken that to be considered a dynamic digital twin, the underlying model itself must be a dynamic simulation. However, this is not the case, as steady-state simulators can and have been used as either static or dynamic digital twins. The generally accepted definitions for static and dynamic digital twins are as follows:

Static: Only model inputs are updated based on real-time data.

Dynamic: Real-time data and internal model parameters are continuously updated to reflect changes to the physical process.

Both categories share the same general nature: a digital representation of a real thing that updates based on real time data. Dynamic digital twins, however, are continuously updating themselves

based on real-time data. Continuous updating here does not mean updating the inputs to the model; both static and dynamic digital twins do this. Dynamic digital twins change inherent parameters of the model itself over time to match actual changes in the physical process. In the midstream industry, an example of this would be changing heat exchanger fouling factors in the model based on actual changes in fouling of the real heat exchanger. If the digital twin is constantly calculating and adjusting a fouling factor itself, the digital twin becomes dynamic. Catalyst deactivation is another example of how the model may change dynamically to maintain an accurate representation of the physical process. Neither fouling nor catalyst deactivation are values measured by the real-time process data, but they both represent real physical changes that are occurring.

These physical changes can be critical or trivial depending on the process. The nature of the process itself and the desired objective of the digital twin will determine whether the digital twin should be static or dynamic. If the objective is Predictive Maintenance, where the digital twin provides insight into future maintenance requirements, a dynamic model may be required. If the process changes little over time and the digital twin is simply providing Key Process Indicators (KPIs) that are absent from instrumentation, then a static digital twin may suffice. It is this reason that this dynamic requirement is excluded in the basic definition of "digital twin" in this work; there are many instances where digital twins bring considerable value to operations in a static form.

Simulation-Focused Digital Twin Model Types

In practice, physics-based digital twins don't always fit perfectly into the categories seemingly designed for data driven models. Does simply allowing the simulator to calculate a fouling factor for a heat exchanger instead of predetermining its value make the model a dynamic digital twin? Is a fully connected model that fits the underlying plant through years of process fluctuations a static digital twin, if the underlying thermodynamic models aren't tuned each time the model executes? We decided that the structure of the physics-based digital twin is less important than its purpose.

Manual and Automated Digital Twins

Extending from the snapshot vs adaptive discussion is another category that can lead to confusion, and this is one of automation. While digital twins based on process simulation can be configured to be either snapshot or adaptive, they can also be configured in either an automated or manual manner. In either case, the models still satisfy the basic digital twin definition presented in the introduction; both set-ups would be a model of a physical process that updates with real-time data. However, some may argue that only models that update automatically in real-time should be considered digital twins, that synchronization must be continuous. The difference between these categories with respect to process simulation is described in further detail below. As there is significant value in both types for midstream engineers, understanding the differences between them and how they fit into a process engineer's needs is important.

Snapshot and Adaptive Digital Twins

To avoid any potential confusion moving forward, this comparison can be addressed through a different lens using similar language, while maintaining the original intent of the static and dynamic definitions: *Snapshot* and *Adaptive.* The snapshot and adaptive terms are defined as follows:

Snapshot: Model inputs are updated and specified using real-time data.

Adaptive: Model inputs are updated and specified using real-time data and characteristics of the physical object, such that the model can adapt to any changes in the physical object using first principles.

These alternative definitions still accomplish the same purpose as the static or dynamic definitions, while also clarifying the boundary between the two categories for physics-based digital twins specifically. Both definitions are now inclusive of the fact that physics-based models are subject to degrees of freedom, unlike data driven models, and each physics-based model must be specified in a way that is consistent with the system's available degrees of freedom.

Because of this, both the snapshot and adaptive models must have their degrees of freedom satisfied with real-time data, which is more specific than the static and dynamic definitions. The adaptive model must also be able to internally adapt to changing conditions per the dynamic definition, which only covers general internal parameters and does not take into account the fact that physics-based models must adhere to the system's degrees of freedom. As a result of having a clearer definition, the snapshot and adaptive terms inadvertently add an additional layer, which can be used to paint a clearer picture as to how each physics-based model type should be used:

Snapshot Purpose: determines the current state of the physical object or process

Adaptive Purpose: determines an extrapolated state of the physical object or process

For physics-based models using simulation, both the snapshot and adaptive models share the same representation of the physical process. The main difference, then, is how each of the models is specified. The snapshot model only uses real-time measurements taken at the physical facility, while the adaptive model is specified to include the equipment and hardware properties, so that it understands the limitations and performance of each physical piece of equipment.

Examples of Digital Twin Model Types

In the midstream industry, a heat exchanger can be a practical example for illustrating the application of each digital twin category. Assume that a processing facility needs to have one of its heat exchangers monitored, for suspicion of fouling. Before any monitoring practices can be used, information describing or quantifying the properties of the exchanger, and the fluids flowing through it must be available and collected for analysis. For heat exchangers specifically, it is very common for instrumentation to be installed on the inlet and outlet of both sides of the exchanger as a means of either controlling the process or measuring the fluid properties. Common fluid property measurements include, but are not limited to temperatures, pressures, and flow rates. Historically, an engineer would then take these readings and hand-calculate the exchanger's properties based on heat-transfer fundamentals, whether it be by pencil-

and-paper, or through use of software. Using process simulation, this manual calculation can be transformed into a digital twin through a one-time setup. After the initial setup, the model can calculate and monitor the current exchanger characteristics or even extrapolate to new process conditions if the current state of the plant were to change.

Snapshot Digital Twin

A snapshot digital twin in this heat exchanger example would entail a simulation that uses real-time measurements as the sole inputs, demonstrated in Figure 1. As discussed previously, degrees of freedom must be carefully considered when determining which inputs should be used in the model—regardless of whether they are manual user inputs or real-time inputs coming directly from the database. Suppose in this model, both inlet streams are fully specified. This includes their temperature, pressure, flowrate, and composition, which have been collected from the available measurements in a facility. Assuming the outlet pressure of each fluid is estimated or known, this leaves 1 degree of freedom left to satisfy the energy balance of the simulation. In this case, an outlet temperature specification would be most appropriate, as it is likely the only measurement available at the plant that can satisfy this energy balance.

Once the degrees of freedom are satisfied, the simulation calculates heat exchanger characteristics, based on the supplied inputs. For a heat exchanger, a simulation can calculate parameters such as fouling resistance, UA, and approach temperatures, among many others, which effectively characterize the exchanger's current performance based on the state of the process and current physical condition of the exchanger. This snapshot digital twin can also be used as a manual or automated model. In the manual case, an engineer can open the model, pull in the most recent data from the physical process, and run the simulation to view the results. When automated, the model provides continuous outputs of the desired performance characteristic calculations.

When the model has been in-service over a long enough period, the continuous outputs compile into a historical record. This record is beneficial for tracking exchanger performance and condition over time because it provides additional measurements that otherwise would not be known, since these metrics cannot be directly measured by physical instruments.

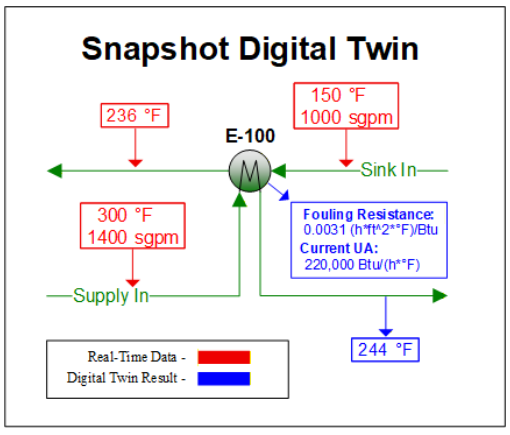


Figure 1. Snapshot Digital Twin Calculating Current Performance of a Heat Exchanger

Adaptive Digital Twin

Since physics-based models are governed by degrees of freedom, the adaptive digital twin, pictured in Figure 2, cannot be specified by the same measurement data as the snapshot digital twin, while simultaneously receiving inputs from underlying equipment performance characteristics—this is a logical inconsistency since the system would be overspecified. As a compromise, the measurements are selectively replaced by performance-based specifications. Selectivity here means that independent measurements, such as the inlet conditions, are still utilized by the adaptive digital twin. Dependent measurements, however, are replaced by physical equipment characteristics that allows the digital twin to be defined by its performance, rather than the current measured conditions. An example of this would be replacing the outlet fluid temperature used in Figure 1 by instead specifying the heat exchanger geometry and real-time fouling resistance, as seen in Figure 2.

Because the adaptive digital twin is extrapolative, it must also receive information to inform the model of where it needs to extrapolate to. Since variables tied to Programmable Logic Controllers (PLCs) or the Distributed Control System (DCS) are the only variables that can be manipulated in the physical process, the adaptive digital twin must conform to the same standards, by mirroring the control system used in the actual process. Adaptive digital twins are designed to allow only controller setpoints to be a free degree of freedom—every other model specification must be tied to either an independent measurement or an equipment performance condition.

In some instances, an adaptive digital twin may require more information other than the geometry or hardware used in the facility. When equipment deteriorates, the actual versus design performance characteristics can change significantly depending on the type of equipment and how it has deteriorated. When this effect is prevalent, the adaptive digital twin cannot function until the current condition of the equipment is known, as it cannot extrapolate while simultaneously determining the current performance. Instead, a snapshot model may need to be used in conjunction with the adaptive digital twin, so that these continuously evolving performance characteristics are first calculated in the snapshot model. As previously discussed, the performance is calculated based on the current conditions in the snapshot digital twin, and the metrics are then fed into the adaptive digital twin. Because of this quality, the adaptive digital twin is primarily used for extrapolating to new conditions using independent measurements from the physical process and equipment performance conditions from the snapshot digital twin in real time.

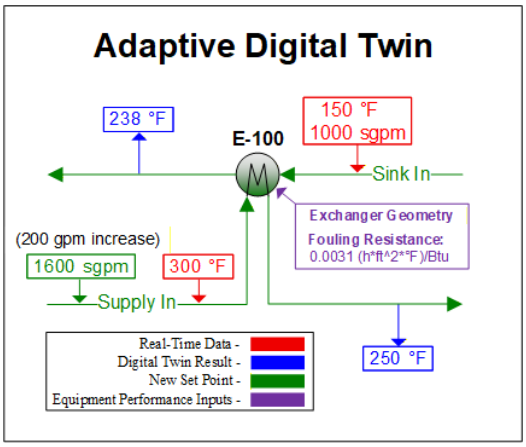


Figure 2. Adaptive Digital Twin Calculating New Process Characteristics

In the adaptive heat exchanger digital twin, the model still utilizes the same independent measurements as degrees of freedom in the simulation: inlet temperature, pressure, flow, and composition. Instead of specifying the outlet temperature of the heat sink stream, as it was in Figure 1, the heat exchanger in the adaptive digital twin is rated using the heat exchanger geometry, and the resulting temperature will be a function of heat transfer first principles. By also incorporating the fouling resistance calculated from the snapshot digital twin, the adaptive digital twin understands the current physical condition of the heat exchanger and can adjust the resulting temperatures to match what the physical exchanger would experience under the same conditions.

Once the digital twin understands the heat exchanger's characteristics; the conditions of the process can be manipulated to represent how the plant would perform under different conditions. Suppose the flow rate of the heat supply stream was increased by 200 gpm. How are the heat transfer properties of the physical exchanger effected? What is the new process state if the flowrate were to change? These are questions that a well-informed adaptive digital twin can easily answer using first principles. When scaled to represent the entire plant, the adaptive digital twin provides powerful insight into engineering and operations for informed decision-making.

Utility of Each Digital Twin Type

The four previously outlined categories of simulation-based digital twins each have their own unique use case, depending on what is desired by the digital twin. Based on the 4 categories: snapshot, adaptive, manual, and automated, these types of digital twins can be represented by their four different use cases, presented in Table 1.

Table 1: Simulation-Based Digital Twin Models

Model Type	Manual	Automated
Snapshot	Normal Simulation with Real-Time Inputs	Continuous Facility Monitoring
Adaptive	Sandbox Model	Real-Time Optimization

Normal Simulation (with Real-Time Data)

The most basic manual digital twin is a simple process model tied directly to real-time operating data. In essence, the simulator is used as it always would be with a human user operating the model. The difference, though, is that the fetch-and-input steps of having an up-to-date model are preconfigured. In other words, the flow of information between the physical and virtual space has already been established. This provides significant value over standard simulation use, as it eliminates both the time required to gather and input the data and any potential errors when sourcing the data, and is simple and straightforward. Once the mapping of the simulation variables is complete, the engineer no longer needs to know where and how to go and get the data manually. Once the model is updated with real-time data, the engineer is then free to perform whatever simulation tasks they desire. The results of the simulation are available and viewed through the GUI of the simulator itself. This non-automated set up has advantages for the engineer as it is easy to explore historical or past events and is not tied purely to real-time data. What if scenarios and re-execution of the model are also possible.

Continuous Facility Monitoring

In the previous case, a human user is executing the model which, as stated before, some may argue means they are not "true" digital twins. However, any manual model can be automated to satisfy a more rigorous definition of digital twin. Automation requires an application that initiates or executes the model on some desired frequency, eliminating the need for human interaction. This is often done through a software service to ensure consistent up time. In such a service mode, the simulator executes silently without any active user interface. Instead of viewing the results of the digital twin through the simulator itself, as with the above cases, the results are fed back to the database or data historian and viewed there.

The scope of these models can vary considerably depending on the desired task and outputs. In some cases, a simulation of the entire plant is required as the total heat and material balance is important. In other cases, systems as simple as a single process stream, heat exchanger, or compressor stage may be sufficient. It is not always required to build a digital twin of an entire process or facility; digital twins of individual components can also be of value.

The size and complexity of the model will vary not only based on the specific purpose of the digital twin, but also on the availability of the operating data that will feed into the model. Where needed inputs, such as compositions, are missing in the instrumentation, the scope of the digital twin can be broadened such that the system boundaries are located where available data exists. Consider the previous illustration from Figure 1, when the snapshot digital twin monitored the performance and potential fouling of a single heat exchanger. If the composition and fluid properties of the various sides of the exchanger are dependent on distillation columns and other separation equipment around the exchanger, then measured data for the composition of these streams may not be available. Broadening the context of the digital twin to encompass this other equipment can provide sufficient information to perform the calculations desired around that single heat exchanger.

Beyond the location of the available data for inputs to the model, the scope of the simulation will depend on the desired insight that the digital twin can provide. Unit Health Monitoring is a common example of such desired insight. In many cases, the parameters needed to ensure good operating conditions and extend the life of equipment are either difficult or impossible to measure. For example, it is commonly sighted that the vapor return line for amine regenerators

should be kept at 1% CO_2 or less to avoid corrosion in the bottom of the amine regenerator [6]. While it is possible that an online gas chromatograph could be used to monitor the CO_2 concentration at this point in the plant, it would be cost prohibitive. A digital twin could be used instead to model the entire amine system and continuously calculate the concentration of CO_2 in the reboiler vapor space. The digital twin thus becomes a type of "soft sensor", supplying values in lieu of actual instrumentation at that point.

These soft sensor values can be stored in data historians to be viewed or even supplied to operator control panels or Human-Machine Interfaces (HMI) displays in real time, such as the configuration shown in Figure 3. This not only organizes the digital twin outputs in a way that is easily digestible for the viewer but also makes the information more accessible to the personnel responsible for running, maintaining, or troubleshooting the represented process. Without a digital twin, the only other way to obtain the desired soft sensor values or KPIs requires an engineer who is proficient in the simulation software to re-create the process inside of a simulator, supply data, and then interpret the outputs.

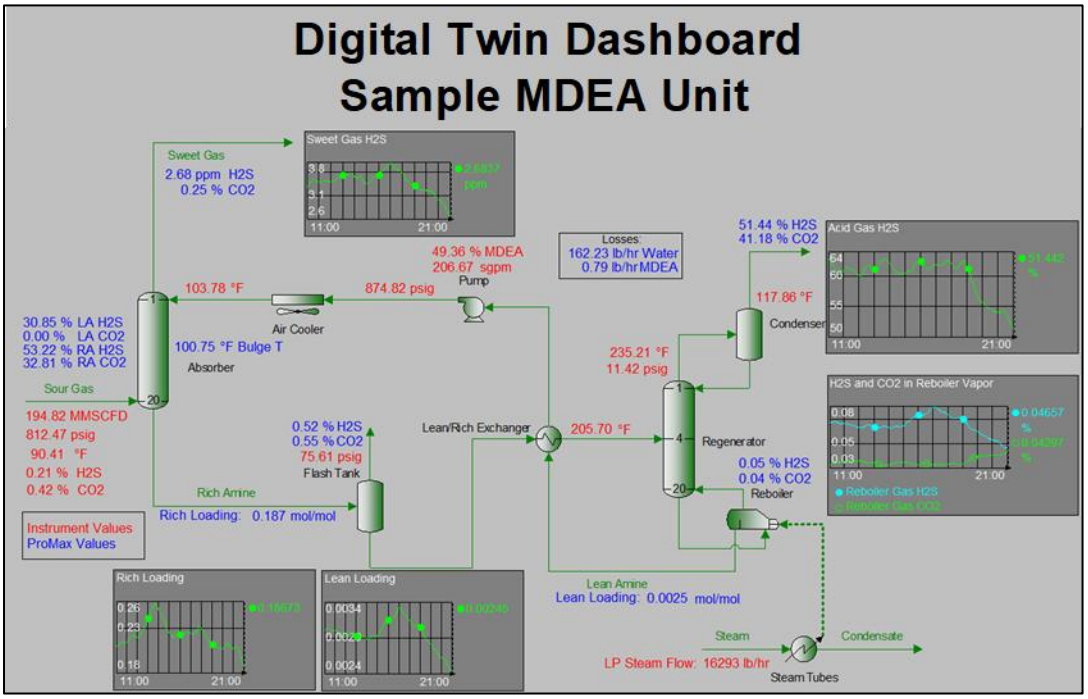


Figure 3: Amine Unit Digital Twin Dashboard

Acid gas breakout is another example of unit health monitoring where the digital twin can be applied. Acid gas breakout occurs when the outlet temperature of the rich side of the lean-rich exchanger is high enough that acid gas stripping begins in the heat exchanger itself, prior to arrival at the amine regenerator. The same digital twin that was used to calculate the CO_2 concentration in the reboiler vapor space can also be used to calculate acid gas break out in the lean-rich exchanger. Not only can the digital twin give insight into whether acid gas break-out is occurring, it can provide information as to the severity, including the composition of the vapor, potential corrosion rates, and even the velocities of the fluid through the tube side of the exchanger.

Once a digital twin is in place and the heat and material balance around a given system is closed, there are multitudes of additional calculations and valuable information that can be provided to operations with little to no additional work. Other KPIs that are desirable but are either not available or difficult to get in a timely manner can be added. Sticking with the amine system from Figure 3, additional values such as the solvent loadings can be provided. These values have already been calculated in the amine system model; it is merely adding them as additional outputs to be made available to operations.

Sandbox

The sandbox model is another form of manual digital twin that is especially useful for performing what-if cases. Like the first digital twin type, it is a pre-configured simulation that can immediately be used for traditional cause-and-effect studies, a common practice among midstream engineers. Unlike the regular simulation, however, the sandbox digital twin incorporates an additional layer by including the actual control inputs that mirror those of the plant. Since it is an adaptive digital twin, the simulation is defined based on the current equipment characteristics, rather than strictly using measured data. In instances where equipment deterioration affects the predicted performance, an automated, snapshot digital twin will more than likely be required to continuously provide the sandbox model with the most up-to-date equipment performance characteristics.

The simulator itself is generally flexible on the inputs provided sufficient degrees of freedom exists. However, in actual plant operations and control, there may be fewer actual degrees of freedom than might be expected. For example, for a given process stream, the simulator requires some definition of flow, whether it is mass, molar, volume, etc. In actual operations, it may not be possible to set a specific flowrate as the actual controller operates on a ratio of flowrates or some other process parameters. Inlet splits in cryogenic gas plants are an example of such a control, shown in Figure 4. The sandbox model takes these factors into consideration, which ultimately brings the digital twin to a closer representation of the plant as it exists in reality. It also better represents what potential moves in the plant are available to operations.

Specifically for the gas plant example, the sandbox has been useful for many midstream operators who want to manipulate setpoints to determine the set point effects on recoveries, product purities, or even compression costs. The greatest benefit of the sandbox digital twin is that it allows the user to change set points in a virtual environment. This removes the need to move the physical plant setpoints to see the outcome. Instead, the new state of the plant can be seen and understood without having to move the plant setpoints. While its primary use is to find new setpoints, the sandbox can also be used as a training aid for new operators to demonstrate and understand the effects of changing temperature, pressure, and flow controllers and their unique impacts on each individual piece of equipment within the digital twin. In the same way that the previous automated digital twin allows information to be accessible to users who may not be familiar with process simulation software, the sandbox model can be configured to receive remote setpoint inputs through the database and send back the outputs without the user needing to physically interact with the simulation software.

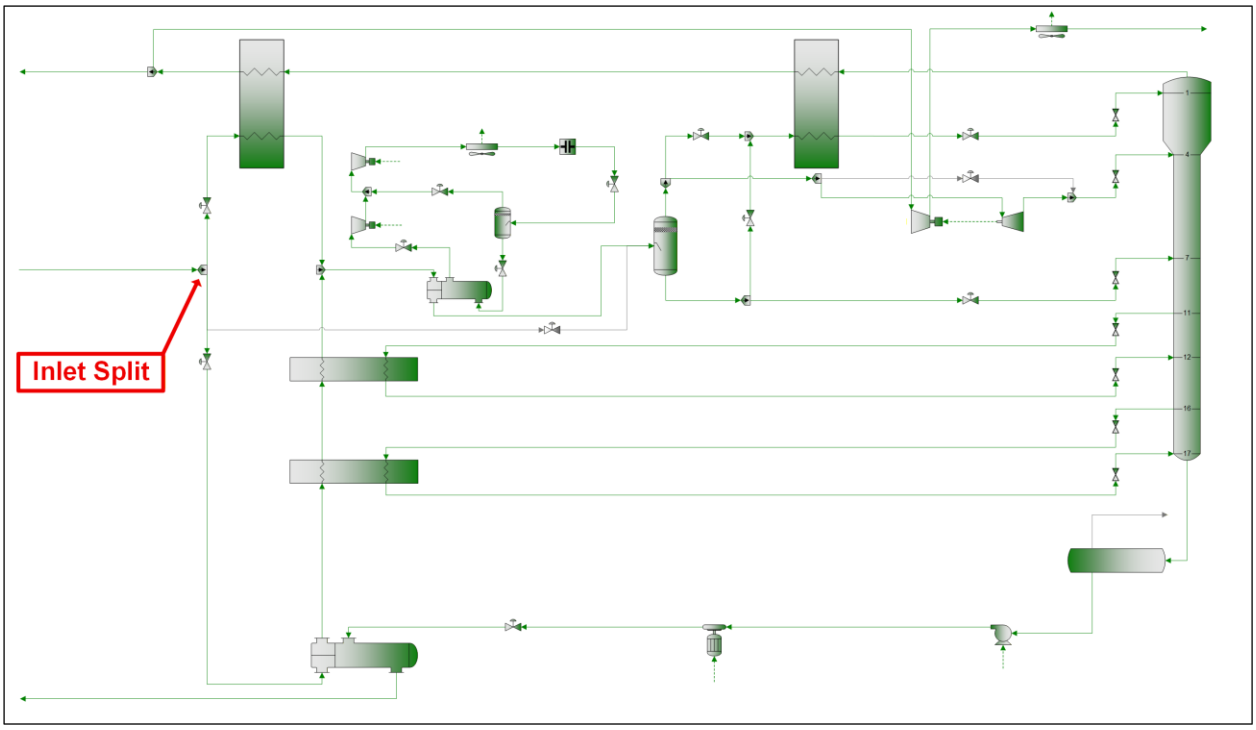


Figure 4: Sandbox Model of a Cryogenic Gas Plant Using the Gas-Subcooled Process (GSP) Technology

Real-Time Optimization

The real-time optimization (RTO) digital twin is easily the most sophisticated of the four categories; however, it is also the most desired because of the potential value it can add to existing assets. The RTO shares many features with the sandbox digital twin, with the exception that the RTO is a fully automated digital twin. In the RTO, the responsibility of making simulation setpoint changes gets handed over to the software that automates the digital twin, instead of requiring the user to specify which set points to change and how far to move them. Within the automation software, an optimization algorithm is given control of the setpoints within the RTO digital twin, which moves each individual controller setpoint as a decision variable to meet or exceed either an individual or collective group of KPIs as its objective function. Common KPIs used in the optimization algorithm may include:

- **Maximize Process Throughput**
- **Minimize Operating Expenses (OPEX)**
- **Achieve or Increase Product Purity Specifications**
- **Reduce Emissions**

Regardless of what KPI may be used as the optimizer objective function, the RTO can be easily controlled using setpoint or process condition equality constraints. This ensures that the recommended setpoint changes will not violate limitations of the physical equipment, exceed product purity or other process quality requirements, or move key operating parameters outside their recommended operating ranges. It also ensures that the setpoints remain within their permitted range, so that the RTO is not extrapolating too far from the current process conditions.

Using the same Amine Unit example from earlier, the RTO can be used to manipulate controller set points to minimize the operating costs required to run the amine unit, since it is generally not a process makes money, while also ensuring the rest of the process parameters stay within acceptable ranges. Variables of interest for amine unit optimization may include:

- **Sweet Gas H_2S and CO_2 Compositions**
- **Rich and Lean Loadings**
- **Regenerator Overhead Temperature and Pressure**
- **Contactor Temperature Profile and Rich and Lean Approaches**
- **Amine Circulation Rate**

Using a similar dashboard as the continuous facility monitoring digital twin in Figure 3, the results of the optimization algorithm are best seen alongside the current plant measurements and digital twin outputs, as shown in Figure 5:

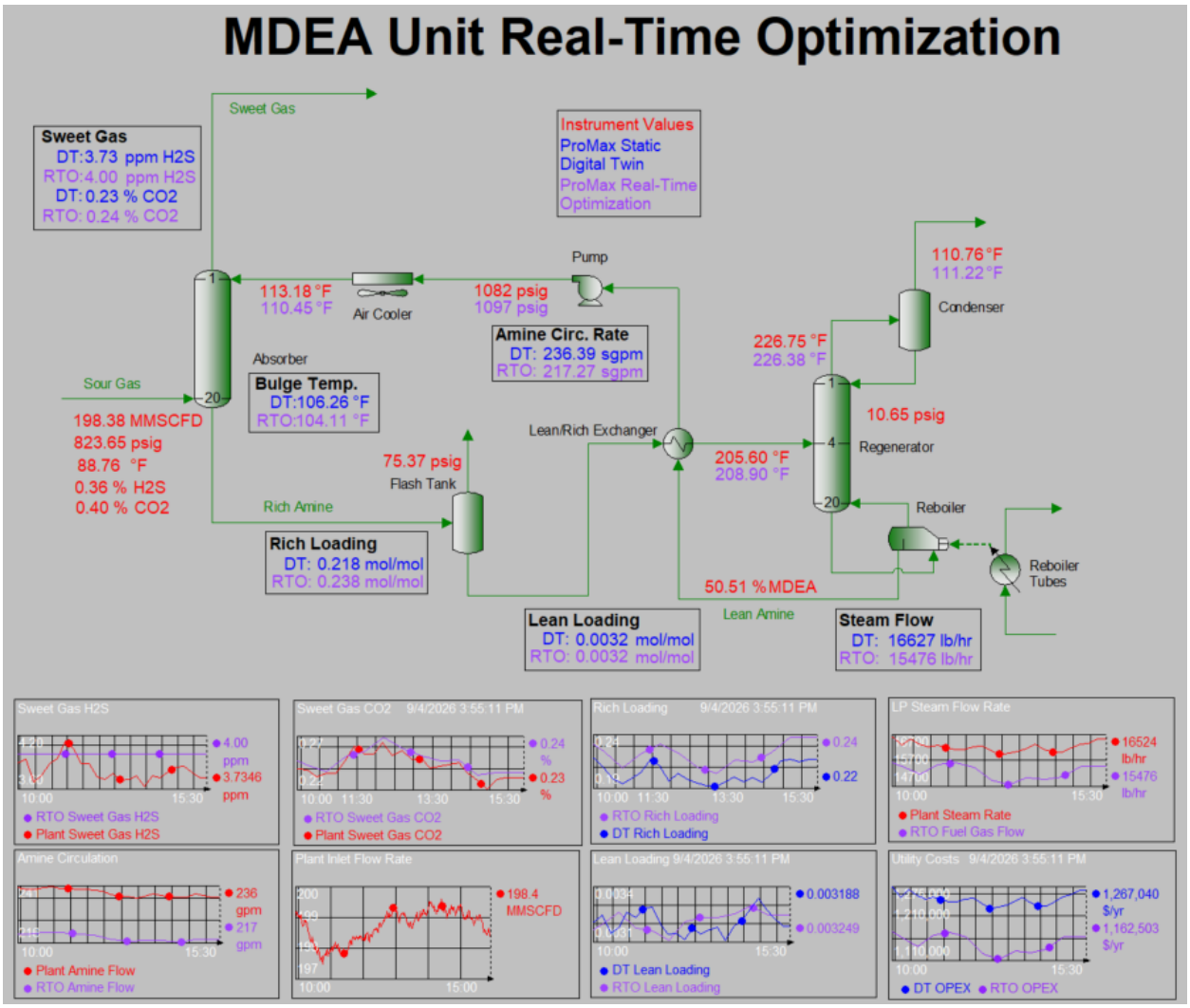


Figure 5: Amine Unit Real-Time Optimization Dashboard

Using dashboards enhances how process and digital twin information gets read and understood. It clearly depicts the current state of the process using the snapshot digital twin outputs in addition to plant measurements, while also displaying new setpoints that will result in a more desirable state of the facility, all in one place. It should be noted that the real-time optimization digital twin is not intended to be an Advanced Process Control (APC) application. The model itself is not manipulating the set points of the physical plant. This function is still performed by the operators themselves. Instead, it is a tool that can be used by both engineering and operations to enhance the general understanding and decision-making process behind the setpoint changes and facility lineups, by representing how current conditions can be changed to move the plant into a more desirable state.

If engineering or operations wants to know more about the projected state of the plant from RTO outputs than what the dashboard displays, or if further analysis of the results of the RTO is desired to double-check or fine-tune the setpoints, they can insert the RTO recommended setpoints into a sandbox digital twin to obtain the same exact results that the RTO sees. Then, the sandbox can be used manually to adjust setpoints as needed, if desired. This is especially helpful in more complicated optimization schemes, such as the cryogenic gas plant example from Figure 4. In many instances, the optimization algorithm has provided a solution that was not considered by a human user when manually changing setpoints [7]. The RTO provides its users with a simple output that can significantly increase the effectiveness of their physical assets to even the most complicated process technologies in the midstream industry. When applied to an entire fleet of processing plants, it has the potential to save or generate millions of dollars.

Benefits of Building and Maintaining Digital Twins

Now having defined several examples of digital twins, we'd like to give specific examples of where digital twins have provided meaningful value in familiar processes. These stories share instances where twins aid decision making in ways expected and intended, and in ways we did not expect at the project outset.

Amine System Cases

A midstream company who operates a natural gas processing plant built and commissioned a continuous facility monitoring digital twin for their amine system. The twin was built to characterize loadings and approaches much like the example previously detailed. Through the digital twin they were able to characterize amine loadings, monitor approach to equilibrium, and regenerator performance.

Upon commissioning the Twin, they quickly realized their acid gas in the reboiler vapor far exceeded the typical 1 mol% CO2 rule of thumb for corrosion and had been for some time [3]. While the plant showed no signs of damage, they were able to help operations find new setpoints for the regenerator to reduce reflux and operate below their goal.

Not everyone is this lucky. A different company commissioned a digital twin of an amine CO_2 removal system after they had ruined the stainless steel packing in their generator. Upon opening the regenerator, the packing was simply gone in most places, having been dissolved by rapid corrosion. Looking at their historical operating data using a sandbox model, we determined their CO_2 in the reboiler exceeded 8 mol% at times. After repacking the regenerator, a continuous facility monitoring digital twin was put in place to help operations avoid repeating the costly

event. More information around the absorber also enabled them to reduce amine circulation by an average of 25%.

The final midstream amine example we'll discuss is a facility with two parallel absorbers that was ramping up throughput and would occasionally receive what they referred to as a "sour pill" from their upstream producers. They were running at the minimum circulation their pumps could deliver, and typically that was still more than necessary. Using an MDEA based solvent, their rich and lean loadings averaged 0.2 and less than 0.0001 mol acid gas/mol amine, respectively.

Engineering built a continuous facility monitoring, and real-time optimization model for this facility to help them characterize their equipment and provide insight on how to lower their reboiler hot oil usage. From their typical loadings and absorber temperature profile it was clear that for typical operation, they were both overcirculating and overstripping. Analyzing preliminary optimized results showed we could reduce hot oil demand by 20%, freeing up more heat medium for other units on site.

What we did not expect at the outset was we would be able to diagnose poorly calibrated amine flow meters, vindicating what appeared to be a less efficient absorber, which was receiving 10% less flow than was measured.

Gas Plant Cases

A midstream company commissioned three digital twins for the same facility: one continuous facility monitor, a sandbox, and a real-time optimization model. When building the snapshot twin we immediately noticed a discrepancy in the mass balance. Plant engineers and operators did not trust the cryo inlet flow meter since liquids accumulated on the orifice plate, causing the accompanying meter to overpredict flow.

While the feed meter itself is a useful instrument to understand plant capacity, it's even more critical given that for this facility, several of their controller setpoints were ratioed to the feed flow. This led operations to run the feed split and cold separator vapors flow control valves in manual. The Twin was able to infer both feed flow and feed composition after the heat trace failed on the gas chromatograph.

Operators were fighting excessive vibration in the natural gas liquids (NGL) pipeline pump for this facility. Engineering responded by adding a net positive suction head (NPSH) calculation to the digital twin and running it over the past year. By plotting the calculated NPSH on the pump requirement curve, it was clear that there was no margin for head loss even assuming clean strainers between the NGL surge tank and the pipeline pump. Engineering and operations were able to agree on a new setpoint that shifted the NSPH actual away from the pump's NPSH required and eliminating the cavitation induced vibration.

The same facility used the sandbox model to validate how well the adaptive digital twins mapped to changing setpoints in the facility before recording recommended optimal setpoints to either maximize ethane recovery or maximize propane recovery at a nominal ethane recovery.

Model Implementation: Responsibilities and Challenges

If you have made it to this point in this work, you likely want to know the answer to the following question: how can my company leverage digital twins? In this section, we distill our experiences implementing digital twins across industries to uncover what challenges companies face and who is needed to overcome them.

Roles and Responsibilities

What starts life as a simple engineering project almost inevitably requires the approval of multiple business units at each company. Successful digital twins require alignment and coordination between engineering, information and operational technology (IT and OT), historian administrators, and operations.

Engineering: Engineers are ultimately responsible for digital twin creation, maintenance, and validation. They will need accurate drawings, equipment datasheets, and access to historical and current operating data to build a digital twin with any chance at running autonomously. Most digital twin projects start here, and it becomes the engineer's responsibility to identify who else is needed to cover what engineering cannot, either due to permissions or skills.

Technology: For good reason, most operating companies are very careful when exposing their operating data outside their network. For this reason, IT will need to build a system for the digital twin to run on inside of the corporate firewall. They will need to coordinate with the historian administrators to ensure the digital twin can communicate bidirectionally with the historian. OT and the historian administrators will need to work with engineering to ensure measurement accuracy and availability in the historian.

Operations: Engineering and operations need to agree on key performance indicators to track with the digital twin. Without operator buy-in, the twin has limited ability to assist day-to-day plant performance beyond identifying equipment degradation.

Management: Aligning these groups typically requires management buy-in and oversight. No other role at an operating company frequently has the authority to authorize work for such a wide range of responsibilities.

Common Challenges

Rarely do digital twins get commissioned without some form of setback. The most common challenge is identifying and coordinating the efforts of each group mentioned above. Frequently, a lack of buy-in from engineering, IT, operations, or management can either limit the value or eliminate the possibility of a digital twin. More tactical level problems can be organized by the group that faces them.

Engineering: Digital twin build efforts are dependent on information availability. While physics based digital twins typically require fewer measured values from the plant, they require more information about the process and equipment to function. Examples of challenges engineers must overcome include:

- Missing or outdated engineering drawings and equipment datasheets
- Insufficient instrumentation to characterize equipment performance
- Local instruments without a digital transmitter for key measurements
- Infrequent or poor-quality feed or product composition measurement

- Changes in lineups or manual valve position that aren't logged in the historian

Technology: Without the right infrastructure in place, you simply can't run a digital twin. Twins require a host system, ability to read and write data, and a mechanism to display those results to engineering and operations. A failure of any part of that system will cause the digital twin to fail. Common implementation and maintenance issues with IT and OT include:

- Poor or no data availability
- Inability to commission a server to host the digital twin
- Insufficient permissions or access to the historian
- Inability to create a space in the historian, or other database, to store digital twin outputs
- Poor data visualization
- Low system reliability

Operations: Operator buy-in is crucial to experiencing the full value a digital twin can provide. While engineers will typically look at long horizon trends of plant and equipment performance, operators are responsible for plant status at any given moment. Challenges faced by operations when implementing digital twins include:

- Critically unstable process inputs
- Damaged or fouled equipment
- Operating at control limits

Well-designed digital twins can help operations with some of these challenges once they are established. The evaluation and commissioning effort of a Twin, however, is difficult when operators are actively fighting adverse plant conditions.

Management: A manager's primary role in digital twin implementations is to coordinate the efforts between each of the other groups and provide clear business objectives to guide Twin development. Digital twin implementations can be derailed by:

- Protracted approval processes for technology infrastructure and software spending
- Ambiguous objectives for the underlying asset
- Inability to align engineering and operations goals

Conclusion

Building digital twins requires more than simply connecting a process simulation to a historian. The digital twin must be appropriately scoped and specified based on the available process data, the degrees of freedom of the underlying model, and the objectives of the application. Successful implementation also requires coordination between engineering, IT/OT, historian administrators, operations, and management. With the appropriate infrastructure, data, and organizational support, process simulation can be transformed from a tool primarily used for periodic engineering studies into a continuously available source of engineering information.

The examples presented demonstrate that the value of a digital twin extends beyond simply reproducing the current state of a facility. Snapshot digital twins can continuously characterize equipment and process performance, providing soft sensors and other calculations that are difficult or impossible to obtain from physical instrumentation alone. Adaptive digital twins can use this understanding of the physical process

to evaluate how a facility will respond to changing operating conditions. When these capabilities are combined with automation, digital twins can progress from continuous monitoring to sandbox analysis and ultimately real-time optimization, providing increasingly powerful tools for understanding and improving facility performance.

Ultimately, the value of a digital twin is determined less by how sophisticated the process technology appears and more by how effectively it applies engineering fundamentals to the needs of the physical facility. By connecting real-time process data with the established capabilities of physics-based process simulation, digital twins provide a practical means of turning large quantities of facility data into actionable engineering information. As these tools become more integrated with existing data and control systems, they can help bridge the gap between the physical facility and the engineering models used to understand it, enabling better decisions and more effective use of existing assets.

References

[1] S. Faber, "Simulation Equipment Used in the Training of Astronauts and Flight Control Crews in Project Mercury," in *Proceedings of the National Meeting on Manned Space Flight*, St. Louis, MO, 1962.

[2] A. a. W. T. Rose, "Punched card devices for distillation calculations," *Industrial and Engineering Chemistry,* pp. 42, 2494, 1950.

[3] M. Grieves, "Completing the Cycle: Using PLM Information in the Sales and Service Functions," SME Management Forum, Troy, MI, 2002.

[4] L. Wright and D. Stuart, "How to tell the difference between a model and a digital twin," *Advanced Modeling and Simulation in Engineering Sciences,* vol. 7, no. 13, 2020.

[5] A. de Souza, "Digital Twin—looking behind the buzzwords," *NAFEMS - Benchmark Magazine,* 1 April 2018.

[6] B. L. G. F. Spooner, "Why Amine Systems Fail - Gas Processing Focus," in *GPA Europe Annual Conference*, Athens, 2016.

[7] R. D. May, J. T. Baker and C. M. Duewall, "Incorporation of Producer Contracts In Gas Plant Optimization Techniques," in *GPA Convention*, 2022.