# CEP
## BRINGS YOU...

Fe

Li

CO₂

# The Most Beneficial Technical ChemE Skills

# Contents

## CEP BRINGS YOU...

### The Most Beneficial Technical ChemE Skills

# PID Explained for Process Engineers: Part 1 — The Basic Control Equation

**CECIL L. SMITH, P.E.**
CECIL L. SMITH, INC.

The differential equation for PID control contains three possible modes: proportional, integral, and derivative. This article describes the control equation in language that process engineers can readily understand.

Understanding the process is the key to successful control applications in the process industries. Since process engineers do this far better than anyone else, the logical conclusion is that developing and enhancing control configurations should be the responsibility of process engineers. The counter argument is that process engineers do not have a sufficient understanding of the principles of automatic control to undertake such a task.

Many process engineers lack this understanding because the traditional method of teaching automatic control is not inherently clear. Explanations that rely on the mysterious "s" variable in LaPlace transforms do more to obscure the basic principles than to elucidate them. This article focuses on the time domain (1).

The objective of this three-part series of articles is to explain the proportional-integral-derivative (PID) control equation in language that process engineers, most being chemical engineers, can readily understand. Part 1 focuses on the basic PID equation. Next month, Part 2 examines the tuning coefficients, and in March, Part 3 explains the most common controller features and options.

## Loop representation

Figure 1 presents a piping and instrumentation (P&I) diagram for a simple temperature control loop, which consists of three components:
• measurement device — the hot air temperature transmitter
• PID controller
• final control element — the fuel control valve.

Additional components are sometimes included in the P&I diagram, but to avoid cluttering the figure, a separate control logic diagram presenting all components can be prepared.

Figure 2 is a block diagram representation of the PID loop in Figure 1. This block diagram contains three function blocks:
• *PV block* — converts the input from the measurement device to a value in engineering units (*i.e.,* °C for the hot air temperature). The result is usually referred to as a process variable (PV).
• *PID block* — provides the PID control calculations. The result of the control calculations is the output to the final
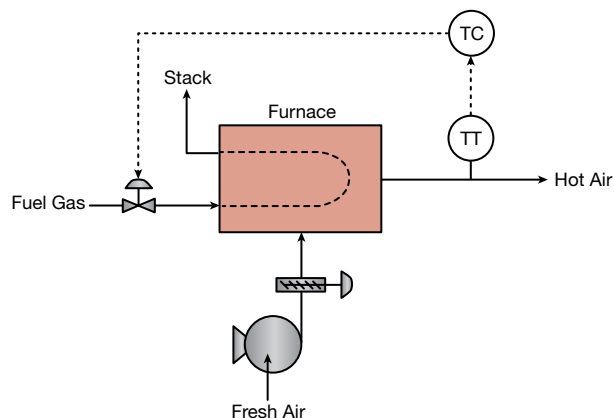


▲ **Figure 1.** In this piping and instrumentation (P&I) diagram for a simple temperature control loop, the hot air temperature is controlled by a valve on the fuel gas supply line.

control element in engineering units of % open.

• *valve block* — converts the % open input to whatever signal is required by the final control element.

In addition to the PV input, the PID block also requires an additional input — the desired value, or setpoint, for the PV input.

Control engineers often use the following terminology for the PV input and the controller output:

• *controlled variable* — the PV whose value is to be maintained at or near the setpoint

• *manipulated variable* — the variable whose value is at the discretion of the controller. For the loop in Figure 1, the manipulated variable is the fuel gas valve opening. Throughout this article, the controller output will be designated by *M*.

Process relationships such as material and energy balances are expressed in terms of flows, not valve or damper openings. For the loop in Figure 1, process-minded people sometimes refer to the fuel flow as the manipulated variable. Technically, such flows are dependent variables that are largely determined by the valve opening but are affected by upstream pressure, downstream pressure, and fluid properties.

## Valve block

In Figure 1, the symbol for the fuel gas valve suggests a valve equipped with a pneumatic diaphragm actuator. Such actuators are common in process facilities, primarily because of their failure characteristics. On a loss of either the control signal or power (supply air), the actuator can be configured to drive the valve either fully closed or fully open. This greatly simplifies the hazards analysis — each actuator can be configured to whatever is most appropriate for that loop.
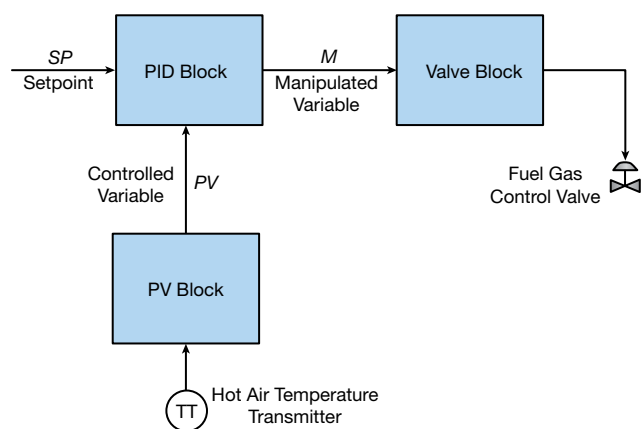
For 4–20-mA current loop installations, the actuator interprets the output signal from the controller as follows:

• *fail-closed*. The output signal determines the % open value for the final control element. For an output signal of 8 mA (25% of the 4–20-mA span), the final control element will be 25% open. Increasing the output signal increases the opening of the final control element.

• *fail-open*. The output signal determines the % closed value for the final control element. For an output signal of 8 mA (25% of the 4–20-mA span), the final control element will be 25% closed, or 75% open. Increasing the output signal decreases the opening of the final control element.

The configuration parameters for the valve block specify the fail-closed/fail-open behavior of the actuator. Based on that, the valve block performs the necessary conversion from the input value (always % open) to the output signal appropriate for the final control element.
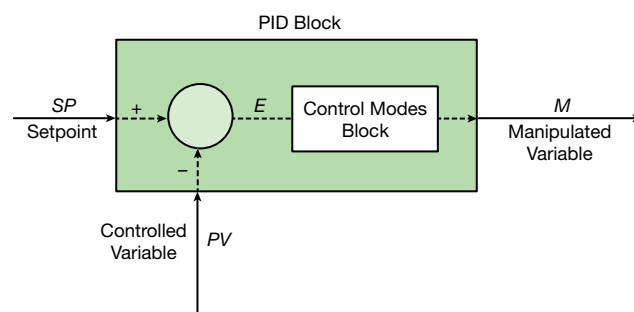
Process operators and control engineers prefer to think of a final control element in terms of % open. The use of valve blocks permits all control configurations to generate every output as % open. The valve block converts the output to % closed if necessary, that is, if the actuator is fail-open. Only the valve block needs to know the failure behavior of the final control element.

Incorporating the functions of the valve block into the PID block was very common in older models of commercial process control systems, and some continue to do so. However, this works well only for simple loops such as the one in Figure 1. When a high/low selector block (for override control) or other computation *(2)* is inserted between the PID block and the final control element, it is best to separate the valve block from the PID block.

## Controller action

Figure 3 illustrates the PID block in more detail. It has two components:

• *comparator*. The control error (*E*) is the difference between the setpoint (*SP*) and the process variable (*PV*). Fig-



▲ **Figure 2.** In this block diagram representation of the PID loop in Figure 1, the setpoint is entered as input by the operator and the process variable is determined by the hot air temperature transmitter. The PID block provides the PID control calculations.



▲ **Figure 3.** Within the PID function block, a comparator computes the control error, *E*. The differential equation for PID control is implemented in the control modes block, whose output is the manipulated variable, *M*.

ure 3 suggests that the control error is computed as *SP* minus *PV*. But for some loops, the control error must be computed as *PV* minus *SP*. (More on this later.)

• *control modes.* The differential equation for PID control contains three possible modes: proportional, integral (also called reset), and derivative.

When drawing a block diagram for a PID loop, control engineers often label the control modes block in Figure 3 as the controller. Indeed, this block is an important one.

Before examining the individual modes, an explanation of controller action is essential.

Consider the control loop in Figure 1. The hot air temperature is controlled by manipulating the fuel gas valve opening. Suppose the hot air temperature is increasing. How should the controller respond? The controller should decrease the fuel gas valve opening. An increase in the controlled variable (hot air temperature) requires a decrease in the manipulated variable (fuel gas valve opening or fuel flow).

An alternative approach is to control the hot air temperature by manipulating the fresh air damper opening. This configuration is illustrated in Figure 4. Suppose the hot air temperature is increasing. How should the controller respond? The controller should increase the fresh air damper opening. An increase in the controlled variable (hot air temperature) requires an increase in the manipulated variable (fresh air damper opening or fresh air flow).

Controller action is defined as reverse or direct *(3)*:

• *reverse* (also called increase-decrease) — in response to an increase in the process variable, the controller decreases its output

• *direct* (also called forward or increase-increase) — in
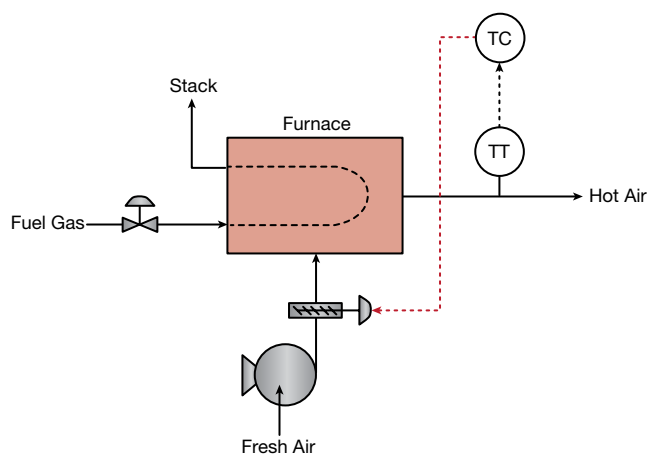
response to an increase in the process variable, the controller increases its output.

Specifically note that controller action is based on the process variable and not the control error. Why base controller action on the process variable? By defining controller action based on the process variable, the definition can also be applied to regulators (which do not contain a comparator, as the PID block in Figure 3 does).

When a valve block (or its equivalent) is used to isolate the controls from the fail-open/fail-closed nature of the final control element, the output of every PID controller is % open. The direction in which the controller must respond depends solely on the behavior of the process.

The configuration parameters for each PID block include a specification for the controller action. If the controller is reverse-acting (as in Figure 3), the control error, *E*, is computed by:

$$E = SP - PV \tag{1}$$

Increasing *PV* decreases *E*, which decreases the output of the controller.

If the controller is direct-acting, the inputs are switched, and the control error is computed by:

$$E = PV - SP \tag{2}$$

Increasing *PV* increases *E*, which increases the output of the controller.

This article uses this approach to incorporate direct or reverse action into the control equation. While most pneumatic and electronic controls used this approach, other possibilities are available with digital controls. From an application perspective, how direct or reverse action is implemented is immaterial.

## Proportional-plus-bias

In process applications, almost all PID controllers rely at least to some degree on the proportional mode of control. By proportional, it is understood that the controller output, *M*, is proportional to the control error, *E*. This suggests the following equation for the proportional mode of control:

$$M = K_C E \tag{3}$$

where $K_C$ is the controller gain (a tuning coefficient).

Consider the hot air temperature loop in Figure 1. Suppose the setpoint is 210°C and that the controller has acted so that the hot air temperature is also 210°C. Since *PV* = *SP*, the control error is zero. Substituting *E* = 0 into Eq. 3 suggests that *M* = 0, which means that the fuel gas valve is 0% open (fully closed). This makes no sense. The fresh air enters at the ambient temperature, so some fuel is required to heat the air to 210°C. The fuel valve cannot be fully closed.



▲ **Figure 4.** If the hot air temperature is controlled via the fresh air damper, an increase in the hot air temperature must result in an increase in the fresh air damper opening. In this scenario, the user must specify direct controller action.

Although the mode is referred to as proportional, the equation for the mode is best described as proportional-plus-bias:

$$M = K_C E + M_R \qquad (4)$$

where $M_R$ is the bias, specifically the controller output bias (the term bias is used in other contexts within process control). The control error $E$ is computed, then multiplied by the controller gain $K_C$, and the result added to the controller output bias $M_R$ to obtain the controller output $M$.

### Initialization and PV tracking

Where do we get a value for the controller output bias $M_R$? The PID controller has two operational modes:

• *manual* — the controller output, $M$, is specified by the process operator, control calculations are not performed, and no value is required for $M_R$

• *automatic* — the controller output, $M$, is computed based on the PID control equation; these computations require a value for $M_R$.

When a controller is switched from manual to automatic, the control equation must be properly initialized so that there is no abrupt change in the controller output, $M$. This is known as a bumpless transition.

The time of the switch is designated as time zero, and $M_0$ is the controller output just prior to the switch. For the transition to be bumpless, the first value computed in automatic mode must be $M_0$. To achieve this, the proportional-plus-bias equation is rearranged so that the initial value, $M_{R,0}$, can be computed:

$$M_{R,0} = M_0 - K_C E_0 \qquad (5)$$

where $E_0$ is the difference between the process variable and setpoint at the time the controller is switched to auto.

PV tracking is an option that affects the initialization of the controller (Table 1). With PV tracking enabled, the initial error ($E_0$) is zero, which simplifies Eq. 5. The simpler

computation, with PV tracking enabled, led to its use within pneumatic and electronic analog controls. For digital controls, the additional calculations for PV tracking disabled are trivial, so most allow PV tracking to be enabled or disabled on an individual loop basis.

When a controller is switched to auto, the operator is responsible for providing a proper value for the setpoint. How this is done depends on whether PV tracking is enabled or disabled:

• *PV tracking enabled.* The operator cannot change the setpoint while the controller is in manual. The operator must first switch the controller to auto, and then enter the value for the setpoint.

• *PV tracking disabled.* The value of the setpoint is used to compute the initial control error, $E_0$, which is then used to calculate $M_{R,0}$. The operator must first enter the value for the setpoint, and then switch the controller to auto.

Some advanced control configurations require PV tracking to be enabled or disabled. But for most loops, the decision to enable or disable is largely personal preference. The best advice is to decide which to use, and make all loops the same except where advanced control (or another specific reason) dictates otherwise.

### Offset or droop

In controllers with only proportional action, the value of the controller output bias, $M_R$, remains at its initial value, $M_{R,0}$. The consequence of a fixed value for $M_R$ is an annoying phenomenon referred to as offset, but the term droop, commonly used in process facilities, is a more accurate description.

Let's look at how droop arises in the loop in Figure 1. As Table 2 indicates, the loop is initially at an equilibrium state (lined out in control jargon) at its setpoint of 210°C. The controller output ($M$) is 82.5% open, and since $E = 0$, the controller output bias must also be 82.5% open.

Figure 5 shows the behavior of the hot air temperature

| Table 1. PV tracking affects the initialization of a controller. When PV tracking is enabled, the initial control error, $E_0$, is 0, because the setpoint is changed to the value of the process variable. | | |
|---|---|---|
| | **PV Tracking Enabled** | **PV Tracking Disabled** |
| Actions in Manual | Setpoint is changed to the value of the process variable | Setpoint is not changed |
| Value of Control Error on Switch from Manual to Auto | $E_0 = 0$ | Direct-acting: $E_0 = PV_0 - SP_0$  Reverse-acting: $E_0 = SP_0 - PV_0$ |
| Initial Value for Controller Output Bias | $M_{R,0} = M_0$ | $M_{R,0} = M_0 - K_C E_0$ |
| Operator May Change Setpoint while in Manual | No | Yes |
| Switching from Manual to Auto | 1. Switch to auto  2. Enter value for setpoint | 1. Enter value for setpoint  2. Switch to auto |

| | Initial | Final |
|---|---|---|
| **Table 2. After a 5% increase in the fresh air damper opening, a loop with proportional-only control exhibits droop. With a fixed value for the controller output bias, $M_R$, the process variable will not line out at its setpoint.** | | |
| Setpoint | 210.0°C | 210.0°C |
| Fresh Air Damper Opening | 60.8% | 65.8% |
| Hot Air Temperature, PV | 210.0°C | 207.8°C |
| Control Error, E = SP – PV | 0.0°C | 2.2°C |
| Controller Output Bias, $M_R$ | 82.5% open | 82.5% open |
| PV Span | 300.0°C | 300.0°C |
| Controller Gain, $K_C$ | 1.0 %/% = 0.333%/°C | 1.0 %/% = 0.333%/°C |
| Proportional Term, $K_C E$ | 0.00% open | 0.73% open |
| Controller Output, M | 82.5% open | 83.2% open |

controller for a 5% increase in the fresh air damper opening (green line). An increase in the fresh air flowrate causes the hot air temperature (blue line) to droop down (from 210°C to 207.8°C, an error of 2.2°C); a decrease causes the hot air temperature to ride up.

Following the increase in the fresh air flowrate, the hot air temperature controller opens the fuel gas valve from 82.5% to 83.2%, an increase of 0.7%. As Table 2 shows, this increase is entirely due to the $K_C E$ term in the proportional-plus-bias equation (Eq. 4).

At the higher fresh air flow, a controller output of 83.5% is required to attain a hot air temperature of 210°C. But should a temperature of 210°C be attained, the control error would be zero and the controller output would be 82.5% (the



**▲ Figure 5.** A proportional-only controller exhibits droop after a 5% increase in the fresh air damper opening. Initially, the controller output (fuel gas valve opening) is 82.5% open. When the fresh air damper opening increases (green line), the temperature drops from its setpoint at 210°C to 207.8°C (blue line), and the fuel gas valve opening increases to 83.2% (orange line). For the controller to return the temperature to its setpoint, the controller output bias, $M_R$, must be adjusted.

value of the controller output bias, $M_R$).

A value of 210°C for the hot air temperature can be obtained by adjusting the controller output bias. Most commercial process controls do not provide features to readily display the value of the controller output bias nor accept changes to its value. This was not always the case. The earliest controllers were proportional-only, but provided an adjustment (often labeled manual reset) for the controller output bias, $M_R$, that permitted the process operator to remove any droop or offset.

The addition of integral action to the controller has largely eliminated the need for manual reset. Process operators are no longer familiar with manual reset, nor do most commercial control systems provide such a feature. The integral mode automatically adjusts $M_R$ so that the controlled variable lines out at its setpoint, thus eliminating droop. The term reset has stuck, though, and the integral mode is commonly referred to as the reset mode.

### Integral mode

To eliminate the droop exhibited by the hot air temperature response (Figure 5), the controller output bias, $M_R$, must be increased. This can be achieved by adjusting $M_R$ at a rate proportional to the control error, $E$:
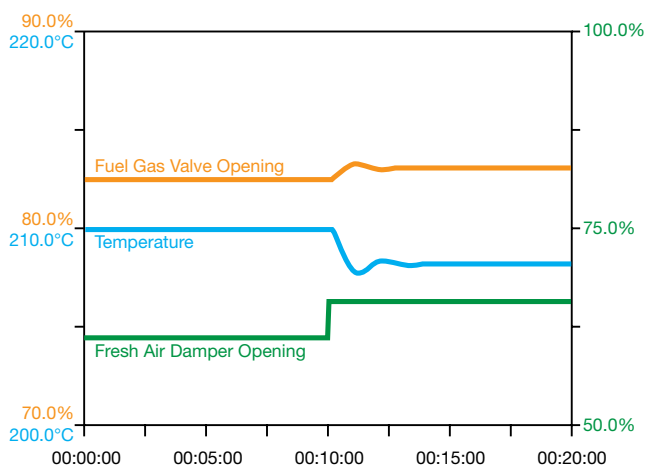
$$\frac{dM_R}{dt} = K_I E \qquad (6)$$

where $K_I$ is the integral gain (a tuning coefficient). For positive values of $E$, the controller output bias $M_R$ increases; for negative values, $M_R$ decreases.

Integrating gives the more familiar equation for the integral mode:

$$M_R = \int K_I E \, dt \qquad (7)$$

To achieve a smooth transition from manual to automatic, the initial value for $M_R$ in both equations must be the

value computed for $M_{R,0}$. In practice, the value computed for $M_{R,0}$ becomes the initial condition for the integrator.

If the integral mode is present in the control equation, the controller can only line out with the process variable equal to the setpoint. If the process variable is not equal to the setpoint, the following statements will be true:

• control error, $E$, is nonzero
• controller output bias, $M_R$, will be changing, as will the controller output, $M$
• the loop is not lined out.

Addition of the integral mode to the control equation does not assure that a loop will line out, but if it does line out, the process variable will be equal to the setpoint. A controller with integral action will not experience the droop depicted in Figure 5.

### Proportional-plus-integral (PI)

An advantage of the integral mode is that the controller can line out only at its setpoint. But if the integral mode is used alone, the controller responds very slowly. Proportional responds more rapidly, but when used alone, the loop exhibits droop. Using the two modes together creates a loop that responds more rapidly and can line out only at its setpoint.

The equations for the proportional and integral modes can be written as follows:

$$M = K_C E + M_R \tag{4}$$

$$M_R = \int K_I E \, dt \tag{7}$$

These equations are usually combined into a single equation:

$$M = K_C E + \int K_I E \, dt \tag{8}$$

The integral mode can also be expressed using the integral time or reset time, $T_I$, instead of the integral gain, $K_I$:

$$M = K_C E + \int \frac{K_C}{T_I} E \, dt \tag{9}$$

Since the controller gain, $K_C$, appears in both modes, the equation is often written as:

$$M = K_C \left( E + \frac{1}{T_I} \int E \, dt \right) \tag{10}$$

An alternative form for the PI control equation is the reset feedback form, which will be covered in Part 3 of this series.

### Derivative mode

An older term for the derivative mode was "pre-act," which suggests that the controller is anticipating the process and acting in advance. Indeed, this is what derivative control attempts to do.

For the $PV$ input to the controller, two items of information can be computed:
• current value of $PV$
• rate of change of $PV$, that is, $dPV/dt$.

Assuming that $PV$ continues to change at the current rate for some time into the future allows a simple predictor to be formulated.

At time $T_D$ in the future, a projected value, $\hat{I}$, can be computed for $PV$ by:

$$\hat{I} = PV + T_D \frac{dPV}{dt} \tag{11}$$

This is illustrated in Figure 6. $PV$ is below the setpoint but moving toward it. The projected value $\hat{I}$ is closer to the setpoint than the current value of $PV$. Instead of basing the control calculations on the control error, $E$, computed from $PV$, the control calculations could be based on the projected control error, $\hat{E}$, computed from $\hat{I}$.

The derivative time, $T_D$, is a tuning coefficient. Its value suggests how far into the future the PV can be projected using its current rate of change. This depends on the nature of the process. One of the most common applications of the derivative mode is for temperature processes, which are slow to respond.

A sequence to perform the control calculations for a proportional-plus-derivative (PD) controller is as follows:
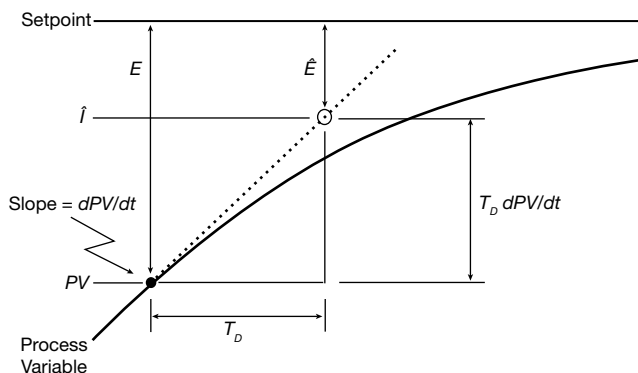
1. Compute a value for $\hat{I}$.
2. Compute the projected control error, $\hat{E}$, using either Eq. 12 or Eq. 13.

For a direct-acting controller:

$$\hat{E} = \hat{I} - SP \tag{12}$$

For a reverse-acting controller:

$$\hat{E} = SP - \hat{I} \tag{13}$$



▲ **Figure 6.** Based on the rate of change of *PV*, a projected value, $\hat{I}$, for the process variable can be computed. The control calculations are then based on the projected error, $\hat{E}$, computed from $\hat{I}$.

3. Use $\hat{E}$ instead of $E$ in the proportional-plus-bias equation:

$$M = K_C \hat{E} + M_R \tag{14}$$

To obtain the customary expression for a PD controller, the above equations are combined:

$$
\begin{aligned}
M &= K_C \hat{E} + M_R \\
&= K_C (SP - \hat{I}) + M_R \\
&= K_C \left( SP - PV - T_D \frac{dPV}{dt} \right) + M_R \\
&= K_C \left( E - T_D \frac{dPV}{dt} \right) + M_R
\end{aligned} \tag{15}
$$

Equation 15 applies to a reverse-acting controller. For a direct-acting controller, the sign on the derivative term would be positive.

### Proportional-integral-derivative

Starting with the PI control equation, a similar approach gives the equation for a PID controller:

1. Compute the projected control error, $\hat{E}$.
2. In the proportional term of the PI control equation, replace the control error, $E$, with the projected control error, $\hat{E}$.

For a reverse-acting controller:

$$\hat{E} = SP - \hat{I} = E - T_D \frac{dPV}{dt} \tag{16}$$

$$M = K_C \left( \hat{E} + \frac{1}{T_I} \int E \, dt \right) \tag{17}$$

Combining Eqs. 16 and 17 gives a single equation for PID control:

$$
\begin{aligned}
M &= K_C \left( E - T_D \frac{dPV}{dt} + \frac{1}{T_I} \int E \, dt \right) \\
&= K_C \left( E + \frac{1}{T_I} \int E \, dt - T_D \frac{dPV}{dt} \right)
\end{aligned} \tag{18}
$$

This equation applies to a reverse-acting controller. For a direct-acting controller, the sign on the derivative term would be positive.

### Derivative based on *E* or *PV*

The approach described in the previous section is based on the rate of change of the process variable and yields a PID equation with its derivative in terms of *PV*. In practice, basing derivative on *PV* is generally preferred, and some commercial control systems provide only this option. The PID equation presented in most textbooks contains a deriva-tive term based on the control error, *E*, and some commercial control systems also provide this option. Whether derivative is based on the process variable or on the control error has a minor effect on the performance of a loop.

To project the control error, *E*, one derivative time into the future and obtain the projected control error, $\hat{E}$, use the equation:

$$\hat{E} = E + T_D \frac{dE}{dt} \tag{19}$$

Replacing the control error in the proportional term of the PI control equation (Eq. 10) with the projected control error calculated by Eq. 19 gives a PID equation with deriva-tive based on *E*:

$$M = K_C \left( E + \frac{1}{T_I} \int E \, dt + T_D \frac{dE}{dt} \right) \tag{20}$$

An issue arises when derivative is based on the con-trol error. Say the operator makes an abrupt change in the setpoint, such as raising a temperature setpoint from 210°C to 215°C. An abrupt setpoint change produces an abrupt change in the control error, *E*. Mathematically, the deriva-tive of an abrupt change is the impulse function. In process control systems, the result is a rate of change with a large magnitude but a short duration — that is, a spike. Part 3 of this series examines the consequences of a spike as part of the discussion on computing rates of change.

### PID computations

Only the continuous forms of the control equations have been presented so far. Digital controls must use difference equations that are numerical approximations to the continu-ous equations.

With today's computers, the time between executions of the PID equation, known as the sampling time, $\Delta t$, is one second or less. Most processes are slow, and such sampling times do not introduce significant errors into the calcula-tions. Technically, one should write difference equations, but with such good approximations, the continuous equation is common. Some people find the difference equations more informative, so the computational procedure outlined here uses those.

An iteration consists of the computations performed on each sampling instant. Each iteration is indicated by a subscript, with *n* being the current iteration, *n*–1 being the previous iteration, and 0 being the first iteration follow-ing a switch from manual to auto. On iteration *n*, $M_n$ is the controller output, $PV_n$ is the process variable, and $M_{R,n}$ is the controller output bias. On the first iteration, $M_0$ is the controller output, $PV_0$ is the process variable, and $M_{R,0}$ is the controller output bias computed for bumpless tansfer.

The following computational procedure mirrors how

PID control actually functions:

1. Start with values for $PV_n$ and $PV_{n-1}$.

2. Compute the value of $PV'_n$, the rate of change of $PV$ on iteration $n$:

$$\frac{dPV}{dt} \cong \frac{PV_n - PV_{n-1}}{\Delta t} = PV'_n \qquad (21)$$

3. Compute the projected value, $\hat{I}_n$, for $PV$ on iteration $n$:

$$\hat{I}_n = PV_n + T_D PV'_n \qquad (22)$$

4. Compute the control error, $E_n$, and the projected control error, $\hat{E}_n$.

For a direct-acting controller:

$$E_n = PV_n - SP_n \qquad (23)$$

$$\hat{E}_n = \hat{I}_n - SP_n \qquad (24)$$

For a reverse-acting controller:

$$E_n = SP_n - PV_n \qquad (25)$$

$$\hat{E}_n = SP_n - \hat{I}_n \qquad (26)$$

5. Apply integral action to obtain a new value for the controller output bias, $M_{R,n}$:

$$M_{R,n} = M_{R,n-1} + \frac{K_C E_n \Delta t}{T_I} \qquad (27)$$

The initial condition is the value $M_{R,0}$ computed for bumpless transfer.

6. Apply the proportional-plus-bias equation based on $\hat{E}$ to obtain the controller output, $M_n$:

$$M_n = K_C \hat{E}_n + M_{R,n} \qquad (28)$$

These relationships provide the basic PID control equation with derivative based on $PV$.

## Closing thoughts

This article has presented the basic PID control equation. However, stopping here would be premature.

This article uses $K_C$, $T_I$, and $T_D$ as the tuning coefficients (also called tuning parameters). All control systems provide a tuning coefficient for each mode, but not all use $K_C$, $T_I$, and $T_D$. The next article in this three-part series discusses alternative tuning coefficients for each mode. In practice, the most undervalued is the coefficient for the proportional mode. This is unfortunate; the speed of response of a loop is primarily a function of the proportional mode tuning coefficient. Part 2 examines such issues.

The flexibility of digital technology permits additional features and options to be incorporated into the PID block. The third article explains the most common of these, including windup and windup protection.  **CEP**

### LITERATURE CITED

1. **Smith, C. L.,** "Practical Process Control: Tuning and Troubleshooting," John Wiley and Sons, Hoboken, NJ (2009).

2. **Smith, C. L.,** "Advanced Process Control: Beyond Single Loop Control," John Wiley and Sons, Hoboken, NJ (2010).

3. **Instrument Society of America,** "Process Instrumentation Terminology," ISA-51.1-1979 (R1993).

**CECIL L. SMITH, PhD, P.E.,** is president of Cecil L. Smith, Inc. (2306 Avalon Place, Houston, TX 77019; Phone: (225) 268-4391; Email: cecilsmith@ cox.net). He has gained expertise in virtually every control technology used in industrial production facilities during his 50 years of experience in process control. He has presented a wide range of continuing education courses on various aspects of process control, with a primary focus on developing and enhancing control configurations. He is featured in the International Society of Automation's (ISA's) Leaders of the Pack (2003) and *Control's* Process Automation Hall of Fame (2009). He has authored five books on process control. Smith received his BS, MS, and PhD, all in chemical engineering, from Louisiana State Univ. (LSU). He is a registered P.E. in California.

# The Essentials of Continuous Evaporation

**BRYAN W. HACKETT, P.E.**
kW ENGINEERING

Evaporation is a standard thermally driven unit operation. Learn three methods to improve evaporator efficiency.

Evaporation is a unit operation that separates a liquid from solids by means of heat transfer via vaporization or boiling. The purpose of evaporation is to concentrate a solution of a nonvolatile solute (*i.e.,* solids) and a solvent (*i.e.,* liquid), which is typically water. Evaporating a portion of the solvent concentrates the solute into a more-viscous liquid product. Evaporation is regularly used in the food processing, chemical, kraft paper, and pharmaceutical industries to produce liquid concentrates.

Evaporation differs from dehydration and drying in that
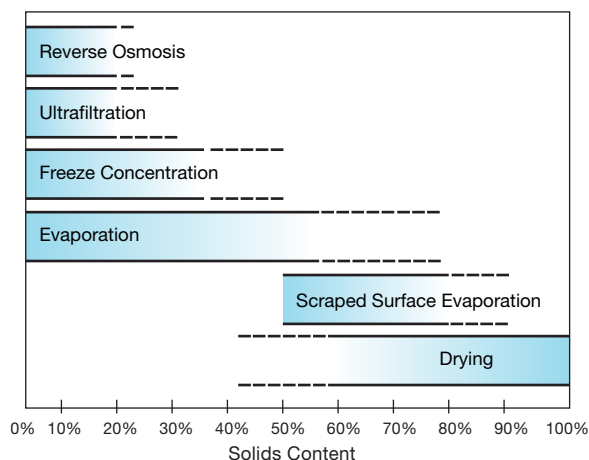


▲ **Figure 1.** Evaporation as a method of concentration can produce solids contents that range from 0% to as much as 92% with scraped surface evaporators (used in select applications).

the product of evaporation is a concentrated liquid, not a solid. Evaporation can be used as the initial step in producing a dried product if the liquid concentrate then undergoes a drying process such as spray drying. The combination of evaporation and spray drying is often used to make powdered products, such as powdered milk. This combination of processes is economically attractive because high-efficiency evaporation is significantly less costly than drying and other methods of removing water *(1)*. Evaporation also produces a higher concentration of solids than other methods of concentration (Figure 1) *(2)*.

Evaporation differs from distillation in that the concentrated solution, rather than the condensed evaporate, is typically the valuable product. A common exception is the evaporation of solutions with a high mineral content, where the vapor is condensed as the product and the concentrated brine is discarded. This process is commonly referred to as water distillation, although the process is more akin to a thermally driven liquid-solids separation operation.

Evaporation may be carried out as a batch or continuous process. This article focuses on evaporation as a continuous process, in which the feed and product streams are continuous and their concentrations remain constant.

## The basics

All evaporators are comprised of two sections: a heating section (called a steam chest) and a vapor/liquid separation section. These sections can be located within a single vessel

(body), or the heating section may be external to the vessel that houses the vapor/liquid separation section.

Evaporators may be composed of one or more effects, where an effect is defined as one or more bodies operating at the same boiling temperature. In a multiple-effect evaporator, vapor from one body heats a second body at a lower boiling temperature. The first effect is heated directly with steam, and the additional bodies are ordered based on descending boiling temperature (or pressure). Figure 2 shows a simple, single-effect evaporation system, which includes an evaporator body with an internal heating section, vapor/liquid separator, water-cooled condenser with barometric leg, and steam-ejector vacuum system.

Evaporator bodies are typically operated under vacuum to reduce the temperature of boiling (*e.g.,* 85°C). Steam ejectors or mechanical vacuum pumps are often used to create a vacuum. Depending on the level of vacuum required for the last effect, which has the lowest boiling temperature, a single pump or a series of pumps may be used.

Vacuum systems also remove noncondensible gases that originate as dissolved gases in the feed or from air leaking into the evaporator body. Most evaporation systems include either a direct or indirect water-cooled condenser to condense the vapor leaving the last evaporator effect. This increases the vacuum of the system. Evaporators that use mechanical vapor recompression (MVR) (discussed later) do not need an external condenser because the vapors generated are fully condensed within the heating section of the evaporator.

## Mass balance on a single-effect evaporator

Evaporators are generally rated based on their evaporation rate — the amount of water they evaporate per hour (*e.g.,* kg/hr, ton/hr). The mass balances for the single-effect evaporator in Figure 3, which include the mass flowrates for feed ($M_F$), product ($M_P$), vapor ($M_V$), steam ($M_S$), and condensate ($M_C$), are:

$$M_F = M_P + M_V \tag{1}$$
$$M_S = M_C \tag{2}$$

Drawing a control volume around the evaporator allows us to perform an energy balance of the system, in which *h* is the enthalpy of the respective streams:

$$M_F h_F + M_S h_s = M_P h_P + M_V h_V + M_C h_C \tag{3}$$

The energy balance assumes that no work is performed on the system ($W = 0$) and heat loss is negligible ($Q = 0$). In real applications, heat loss is typically estimated to be 2% of the vapor energy input to each effect (2). Latent heat from saturated steam drives the evaporation of water, and depend-
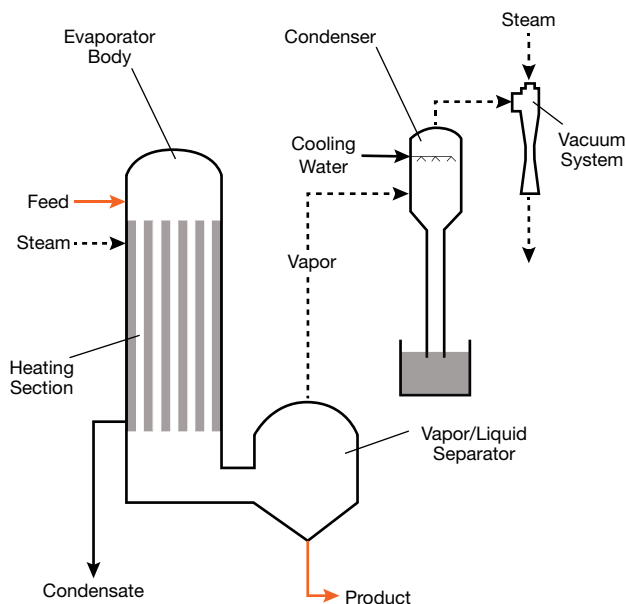
▲ **Figure 2.** The heating and vapor/liquid separation sections are separated in this diagram, but in some types of evaporators, they may be located in the same body.
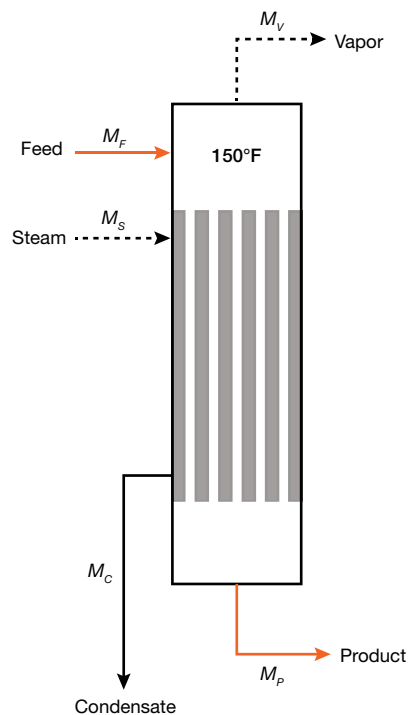
▲ **Figure 3.** An energy balance for this evaporator assumes no work is performed on the system and heat loss is zero. To get a closer approximation of the evaporator rating, however, the heat loss is estimated to be 2% of the vapor energy input to each effect.

ing on the initial feed temperature, may also be used to raise the temperature to boiling.

The product concentration, where $x_F$ and $x_P$ are the mass fractions of the respective streams, can be calculated from:

$$M_F x_F = M_P x_P \qquad (4)$$

### Steam economy

Another way to measure and compare the performance of evaporators is to consider the amount of water evaporated per amount of steam consumed. This metric is called the steam economy (SE) and is commonly expressed as the ratio of water evaporated to steam consumed.



▲ **Figure 4.** The steam economy of an evaporator can be improved by adding more effects. The first effect has the highes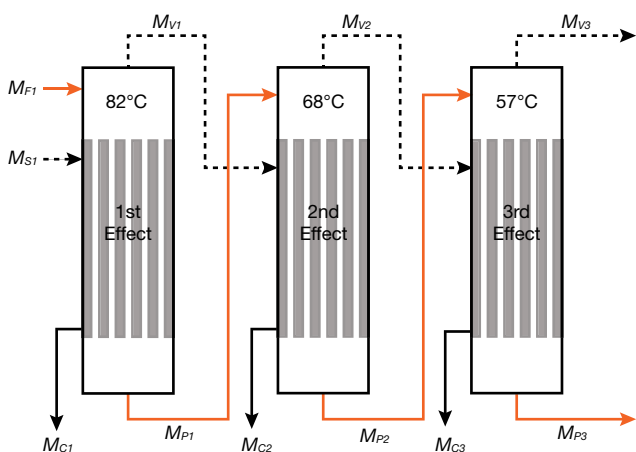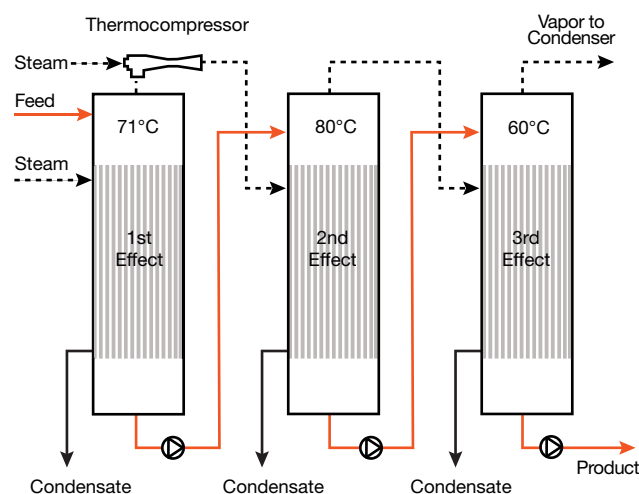t boiling temperature, while the last effect has the lowest. Each successive effect uses the vapor from the previous effect as a heat source.



▲ **Figure 5.** Thermal vapor recompression (TVR) can be applied to an evaporator system to improve the steam economy. A thermocompressor raises the saturation temperature of the vapor exiting an effect to be reused to heat the next effect.

In theory, the evaporation of 1 kg of water requires slightly more than 1 kg of steam, since the latent heat of vaporization decreases as pressure increases *(3)*. As the steam's latent heat is transferred to the feed, some of the heat raises the initial feed temperature to the boiling point and the remaining latent heat evaporates water. Thus, the steam economy will always be less than one. Depending on the initial feed temperature and system losses, the steam economy of a single-effect evaporator typically is 0.75–0.95 kg water/kg steam *(4)*.

*Improving steam economy.* Consider the example of the single-effect evaporator (Figure 3). For heat transfer to occur, the temperature of the saturated steam must be higher than the boiling temperature of the feed. If the latent heat of the water vapor leaving the evaporator can be reused, the steam economy can be increased. But, because heat transfer requires a temperature gradient, the latent heat of the water vapor can only be transferred to a lower-temperature fluid.

The two primary means of improving the steam economy of an evaporator both involve manipulating temperature via pressure change.

### Multiple-effect evaporators

If the water vapor from the first effect of a multiple-effect evaporator can be introduced into the steam chest of a second effect operating at a lower boiling point, the latent heat in the water vapor can be reused. Lowering the vapor pressure of the second effect relative to the first effect lowers the boiling point of the second effect. This arrangement of reusing vapor latent heat is called multiple-effect evaporation.

Figure 4 shows the mass flows in a three-effect evaporator. The vapor from the first effect ($M_{V1}$) becomes the heating source for the second effect ($M_{S2}$). This is possible because the second effect's boiling temperature has been lowered to 68°C (at 29 kPa), which is less than the saturated temperature of the first effect, 82°C (at 51.8 kPa). The latent heat is reused again as $M_{V2}$ becomes the heat source of the third effect, where the boiling point temperature has been lowered to 57°C (at 17.5 kPa). The steam economy of multiple-effect evaporators can be estimated by *(4)*:

$$SE = eN \qquad (5)$$

where $e$ is the steam economy of a single-effect (between 0.75 kg water/kg steam and 0.95 kg water/kg steam) and $N$ is the number of effects.

### Recompression

The second method for reusing the latent heat of vapor is to increase its condensing temperature by compressing it. This increases the vapor's condensing pressure and tempera-

ture and creates the temperature gradient necessary for latent heat reuse within the evaporator system.

*Thermal vapor recompression (TVR)* uses a steam ejector to compress the vapor leaving the evaporator vessel. In this application, the steam ejector is often called a thermocompressor; its primary purpose is to compress vapor, but its secondary benefit is creating vacuum within the evaporator vessel. As a thermocompressor, the ejector uses high-pressure steam (0.8–2.1 MPa) as the motive fluid to draw in the low-pressure vapor from the evaporator and produce a combined vapor of suitable pressure and temperature for evaporator heating.

Figure 5 illustrates one possible configuration of a three-effect evaporator with TVR. In this example, the thermocompressor raises the saturation temperature of the vapor from 71°C to 88°C so it can be reused to heat the second effect operating at 80°C.

Steam ejectors are better suited to handling large volumes of low-density vapor than mechanical vacuum pumps and are easier and cheaper to maintain. As a rule of thumb, TVR improves an evaporator's steam economy by an amount equivalent to the addition of another effect, but at a much lower cost *(5)*.

*Mechanical vapor recompression (MVR)* uses a mechanical compressor or fan to compress low-pressure vapor leaving an evaporator to a higher pressure and temperature for reuse as the evaporation heat source. The internal reuse of the vapor also eliminates the need for an external condenser, which most evaporator systems need to condense the vapor leaving the last effect. Because compression ratios of 1.2–1.4 are sufficient to produce a condensing temperature increase of 6–12°C, centrifugal fans are used for their large volumetric capacity, economy of operation, and relatively low cost. Steam turbines are typically used to drive mechanical fans at installations where the cost of steam is relatively low and low-pressure steam can be used for process heating. Otherwise, electric motors with variable-frequency drives (VFDs) are used to drive mechanical fans.

Figure 6 shows a single-effect, vertical long-tube evaporator with MVR. An MVR evaporator requires steam to bring the system up to boiling temperature during startup, but requires very little steam once operating at steady state, with only the work of the compressor adding energy to the system.
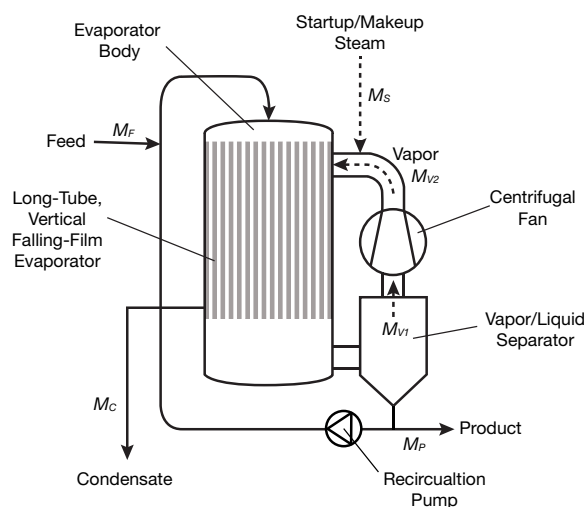
## Comparing steam economies

Table 1 compares the steam economies of multiple-effect, TVR, and MVR evaporators *(5)*. Steam economy increases steadily in multiple-effect evaporators as the number of effects increases. Adding a thermocompressor to a three-effect system can approximately double the economy, but that pales in comparison to the potential steam economy

of an evaporator with MVR — which is 5–10 times that of the three-effect evaporator.

Note that the steam economy of an evaporator is highly dependent on the initial temperature of the feed as it enters the evaporator system and the properties of the liquid being concentrated.

Feed that enters the first effect at a temperature above the boiling point first flashes some vapor before heat transfer within the evaporator even begins. Thus, the steam economy will be higher than if the initial feed temperature was below the boiling point, which would require heat transfer to the liquid before evaporation can occur.

Boiling point rise, also called boiling point elevation (BPE), is the difference between a solution's boiling point and the boiling point of pure water at the same pressure. Some solutions, especially as they become more concentrated, may require a large temperature gradient (16–27°C) to sufficiently transfer enough heat to reach boiling conditions. High temperature gradients limit the number of effects that can be applied, thereby limiting the potential steam economy increase of a system. Also, single-stage thermal



▲ **Figure 6.** Mechanical vapor recompression (MVR) improves steam economy by compressing low-pressure steam exiting the evaporator and recycling it as the heat source.

| Table 1. Comparison of evaporator steam economies. | |
|---|---|
| **Type of Evaporator** | **Steam Economy, kg water/kg steam** |
| Single-effect | 0.90–0.98 |
| Double-effect | 1.7–2.0 |
| Triple-effect | 2.4–2.8 |
| Six-effect | 4.6–4.9 |
| Triple-effect with TVR | 4–8 |
| Single-effect MVR | 10–30 |

and mechanical compressors are economically limited to compression ratios of about two, which produces a temperature increase of 5–20°C *(5)*. Thus, the potential to increase the steam economy by adding effects, TVR, and MVR may be constrained by the thermal and rheological properties of the liquid being concentrated.

### Evaporation in industry

Evaporation is used for subtlety different purposes and with different results in each industry. In the food processing industry, food products are concentrated to increase their shelf life, reduce volume and weight, reduce storage costs, reduce transportation costs, and increase value.

Concentration of fresh foods, such as fruit juices, is an economic way to make use of perishable crops during peak harvest periods *(6)*. Fruit juice concentrates and tomato paste are two examples of seasonal foods that evaporation preserves for year-round consumption. Lower water activity, which is a general indication of microbiological stability, preserves concentrated foods.

In the chemical and kraft paper industries, evaporation is often used to recover valuable, nonvolatile chemicals. For example, in the kraft pulping process, water is evaporated from spent cooking liquid (*i.e.,* black liquor), which concentrates the solids. The concentrated inorganic process chemicals can then be refined for reuse within the process.

In the pharmaceutical industry, evaporation is often used to create a concentrated solution, which can then be dried to create a powdered product.

### Evaporators in the food industry

The economics of a process dictate the evaporator type and configuration best suited to a particular application. The three most common types of evaporators used in food concentration are long-tube vertical evaporators, forced-circulation evaporators, and plate-type evaporators.

Falling-film evaporators are constructed around a vertical bundle of large-diameter long tubes, with a liquid distributor on top and a vapor/liquid separator section below. Liquid is fed into the top of the heat exchanger tubes and flows as a film down the inner walls. Heat transferred from the condensing steam through the tube walls creates boiling conditions, and the vapor generated escapes into the inner void. As the film flows downward and water evaporates, the film becomes thinner and the liquid becomes more viscous. Since the film flows by gravity down the long tube, residence times in the heat exchanger section are short (0.5–2 min) *(3)*, which is beneficial to heat-sensitive foods, such as fruit juices and milk.

Forced-circulation evaporators use a centrifugal pump to recirculate the liquid through the heat exchanger (*i.e.,* steam chest), where the liquid is heated by steam. As liquid is pumped upward through the heat exchanger, the static head pressure suppresses vapor formation until the heated liquid enters the vapor/liquid separator, where a portion flashes to vapor. The centrifugal pump allows the residence time in the heat exchanger to be controlled based on the heat-sensitivity of the liquid being concentrated. This feature is important when concentrating pulpy, viscous liquids, such as tomato juice and mango juice.

Plate-type evaporators operate similarly to long-tube vertical evaporators, but with the film flowing on plates rather than tubes. Their advantages over tubular-type evaporators include compact size that allows installation within buildings and their ability to be disassembled for easier cleaning.

### Wrapping up

The primary means of increasing the steam economy of an evaporator is to reuse the latent heat of the water vapor. A multiple-effect evaporator uses the water vapor from one effect as the heating medium for the next effect, which operates at a lower boiling point. The latent heat in water vapor can also be reused by thermally or mechanically compressing the vapor to a higher pressure and temperature. These three methods of improving efficiency have their limits, however, which are specific to the physical properties of the liquid feed being evaporated.  **CEP**

**LITERATURE CITED**

1. **Smith, P. G.,** "Evaporation and Drying," Chapter 12 in "Introduction to Food Engineering," Kluwer Academic/Plenum Publishers, New York, NY (2003).

2. **Morison, K. R., and R. W. Hartel,** "Evaporation and Freeze Concentration," Chapter 8 in "Handbook of Food Engineering," 2nd ed., Heldman, D. R., and D. B. Lund, eds., CRC Press, Boca Raton, FL (2007).

3. **Saravacos, G. D., and A. E. Kostaropoulos, eds.,** "Handbook of Food Processing Equipment," Kluwer Academic/Plenum Publishers, New York, NY (2002).

4. **Chen, C. S., and E. Hernandez,** "Design and Performance Evaluation of Evaporation," Chapter 6 in "Handbook of Food Engineering Practice," Valentas, K. J., *et al.,* CRC Press, New York, NY (1997).

5. **Minton, P. E.,** "Handbook of Evaporation Technology," Noyes Publications, West Wood, NJ (1986).

6. **Roa, M. A., and A. A. Vitali,** "Fruit Juice Concentration and Preservation," Chapter 7 in "Handbook of Food Preservation," Rahman, M. S., ed., Marcel Dekker, New York, NY (1999).

**BRYAN W. HACKETT, P.E.,** is a project manager and senior energy engineer at kW Engineering (Oakland, CA; Phone: (510) 834-6420; Email: bhackett@kw-engineering.com), where he leads the industrial services team to provide energy auditing, retro-commissioning, technical support services, and project management to food processing and industrial facilities. He has more than 20 years of experience in identifying, analyzing, and implementing energy efficiency projects. He has a BSc in mechanical engineering from San Francisco State Univ. and is a licensed professional engineer in California.

# An Oil Refinery Walk-Through

**Tim Olsen**

Emerson Process Management

Refineries may appear complicated at first glance. Breaking them down into a series of units makes them easier to understand. This article describes the basic building blocks of a fuels refinery, from crude oil inlet to fuels distribution.

Refineries that process crude oil into transportation fuels can be overwhelming to the outside observer — with their multitude of tanks and vessels, numerous pipes of varying sizes, fired heaters, pumps and compressors, instrumentation and control systems, and more. The primary purpose of all of this is to economically convert raw crude oil into more-useful fuels and petrochemicals.

This article gives an overview of each unit within a refinery and explains how intermediates are transformed into transportation fuels. It focuses only on refineries that produce transportation fuels, such as gasoline, diesel, and jet fuel, and does not discuss integrated petrochemical complexes.

The simple flow diagram in Figure 1 shows how crude oil is converted into higher-value fuel products. First, the raw crude is washed in a desalter and heated. Next, it enters the atmospheric crude fractionator, followed by the vacuum fractionator. These first units in a refinery involve no chemical reactions or catalysts; rather, based on each component's boiling point, they separate and distribute the range of components for further downstream processing and conversion. Downstream units include such equipment as pumps and compressors, heat exchangers, reactors, and distillation columns. Some of these downstream units contain a specific catalyst to convert one product to a different product with more desirable qualities. Streams are then separated into intermediate products and off-gases. Finally, certain intermediate products from these units are blended into final products as per required specifications, such as octane and Reid vapor pressure (RVP) for gasoline.

Crude oils vary significantly with respect to their API gravity (a measure of relative density), component mixture, and levels of metals, sulfur (sweet crudes contain less than 0.42% sulfur, sour crudes more than that), and various salts. Crude assays are descriptions of different crude oils with a distillation range to indicate the quantity of components in each boiling point range. This information is used to determine relative flows out of the atmospheric crude fractionator.

Each refinery has been designed to process crude oil of a particular composition, whether that is the composition of a single crude oil or the composition of a blend of crude oils. If an individual crude oil with the desired feed properties for a refinery is not available for processing, multiple crudes can be blended to obtain a feedstock that better matches the refinery's capabilities.

As the crude oil is separated and distributed to downstream processing units, more scrutiny is given to four main types of components — paraffins, olefins, naphthenes, and aromatics (PONA). The crude's PONA content is important in setting the objectives of the catalysts used in downstream units in terms of functionality and desired reaction conversion.

In the U.S., crude oil is measured on a volumetric basis at a standard temperature of 60°F, typically in units of barrels (1 bbl = 42 gal) per day (bpd). Outside the U.S., crude oil tends to be measured on a mass basis in units of metric tons per day. In general, a small refinery produces 100,000 bpd or less, a mid-size refinery is in the range of 100,000–250,000 bpd, and large refineries (typically integrated with a petrochemical processing facility) produce 250,000 bpd or more.

## Crude desalter

After it is received at the refinery, crude oil is stored in large tanks in a tank farm. The first step of the refining process is mixing the crude oil with water in the line upstream of the desalter vessel to dissolve the salts contained in the crude oil. The salts and sediment are captured in the water phase (now referred to as brine), which is then separated from the oil. The desalter typically contains electric grids that assist with the oil-water separation. This allows for a smaller vessel design; without the grids, the required residence time in the vessel would be much higher and a larger vessel would be required. Salts need to be removed from the crude oil to mitigate vessel and piping fouling and corrosion, as well as poisoning of downstream catalysts.

Before the desalted oil enters the atmospheric fractionator, it is preheated by several heat exchangers and a fired heater. Up to 50% of the required heat may come from heat exchange with the side-cut and product draws exiting the atmospheric crude fractionator; the remaining heat comes from the fired heater. This preheating raises the crude oil's temperature to 650–700°F. Above this temperature, thermal cracking would more readily occur, producing carbon or coke deposits on the piping and equipment that would require a shutdown for cleaning.

## Atmospheric crude fractionator

The primary objective of the atmospheric crude fractionator (Figure 2) is to separate the desalted, 650–700°F crude oil into fractions, or cuts, based on the boiling point ranges of the components (known as cut points).

Lighter fractions come off the top of the column and progressively heavier components with higher boiling points are pulled off as side draws further down the tower. Each side-cut draw has a target initial boiling point (IBP) and end boiling point (EBP) to match the capabilities and specifications of downstream units. Typical cut points for the atmospheric fractionator and the vacuum fractionator that follows it are:
- light straight-run (LSR) naphtha, 90–190°F
- heavy straight-run (HSR) naphtha, 190–330°F
- kerosene, *i.e.,* jet fuel, 330–480°F
- light atmospheric gas oil (LAGO), 480–610°F
- heavy atmospheric gas oil (HAGO), 610–800°F
- vacuum gas oil (VGO), 800–1,050°F
- vacuum-reduced crude (VRC), above 1,050°F.

The target cut-point ranges typically do not change, but flowrates must change if a different feedstock crude oil is used.

Each side draw feeds a stripping column, which uses
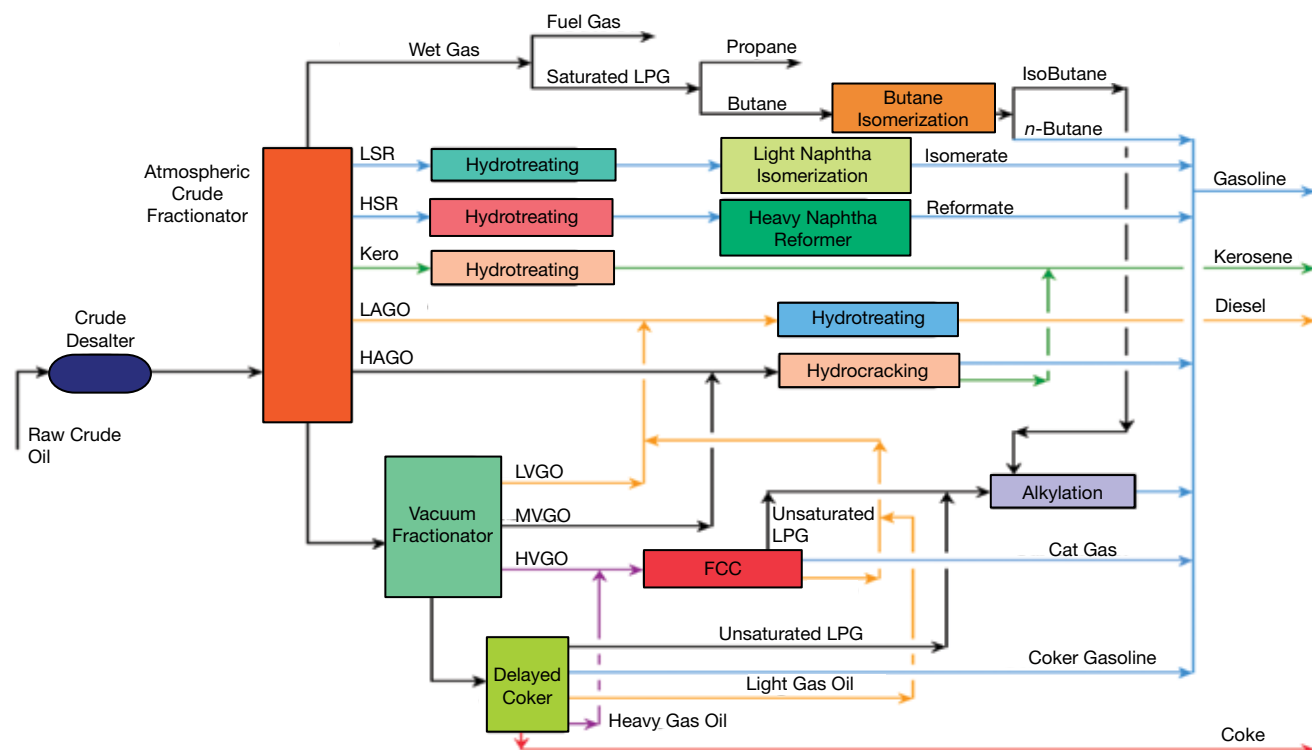
▲ **Figure 1.** This refinery flow diagram demonstrates how raw crude oil is converted into fuels.
LSR = light straight-run naphtha. HSR = heavy straight run naphtha. Kero = kerosene. LAGO = light atmospheric gas oil.
HAGO = heavy atmospheric gas oil. LVGO = light vacuum gas oil. MVGO = medium vacuum gas oil. HVGO = heavy vacuum gas oil.

steam to establish (or control) the IBP of the material leaving the bottom of the side-cut stripper. Steam enters the bottom of the stripper and reduces the hydrocarbon partial pressure. The process fluid partially vaporizes to reestablish vapor-liquid equilibrium. (The heat of vaporization comes from the process fluid itself, not from the stripping steam.) As mentioned earlier, these side draws are used to preheat the incoming crude oil. Typical side-draw cuts include light and heavy naphtha, kerosene or jet fuel, diesel, and light and heavy gas oil. The bottoms draw, referred to as atmospheric residuum, is a heavy, high-boiling-point oil, which is sent to the vacuum column for further separation.

### Vacuum fractionator

Downstream of the atmospheric crude fractionator (which operates at temperatures up to 650–700°F), the vacuum fractionator further separates the heavy, higher-boiling-point components under a vacuum to prevent thermal cracking. Because boiling points decrease with lower pressure, these separations can be achieved without initiating thermal cracking. Steam is added to the vacuum column's inlet to prevent fouling and to the vacuum fractionator bottoms to improve vaporization of these heavy components. The vacuum fractionator tends to be much larger in diameter than the atmospheric crude fractionator (a good way to tell the two apart from a distance) because under a vacuum the vaporized crude oil occupies a larger volume.

### Gas concentration unit

The light hydrocarbon gases that come off the top of the atmospheric crude unit are concentrated, separated, and distributed to where they are needed in what is often referred to as the saturated-gas concentration unit (not shown on Figure 1). Gases are typically separated into wet gases (propane and butane, or liquefied petroleum gas [LPG]) and dry gases (ethane and methane) through a series of distillation columns. The dry gases are primarily used as fuel throughout the refinery for the fired heaters and boilers. Purchased natural gas is used to supplement this fuel gas based on refinery fuel demands.
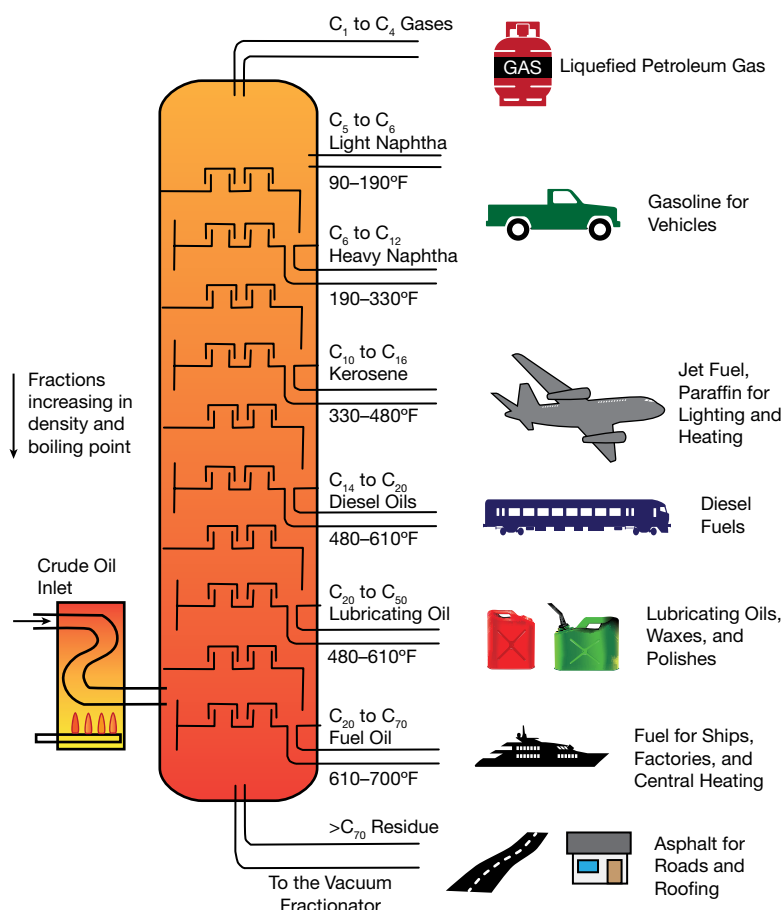
### Fluidized catalytic cracker

The fluidized catalytic cracker (FCC), illustrated in Figure 3, was developed collaboratively by refiners during World War II to help meet gasoline demands for the war. Its function is to break — or crack — long-chain hydrocarbons such as heavy gas oil into lighter, shorter, naphtha-boiling-range hydrocarbons. Its unique process dynamics can make the FCC unit one of the most challenging refinery units to operate.

The FCC gets its name because the catalyst is so fine that it flows like a liquid. The mixture of fluidized catalyst and feed is reacted at high temperatures (900–1,000°F). The cracking reaction is endothermic, and is fueled by the exothermic catalyst-regeneration reaction. Because about 5 wt% of the feed ends up as coke on the catalyst, the catalyst must be regenerated by introducing air into the regenerator to burn off the coke. The catalyst flowrate is about 1 ton per thousand barrels of oil feed.

The reactor and regenerator typically sit side by side in a common steel structure, which makes it easy to identify the FCC unit. The catalyst regenerator exit temperatures can range from 1,200°F to 1,500°F.

Downstream of the FCC reactor is the FCC main fractionator column, which is



$C_1$ to $C_4$ Gases

Liquefied Petroleum Gas

$C_5$ to $C_6$
Light Naphtha

90–190°F

Gasoline for Vehicles

$C_6$ to $C_{12}$
Heavy Naphtha

190–330°F

$C_{10}$ to $C_{16}$
Kerosene

330–480°F

Jet Fuel, Paraffin for Lighting and Heating

$C_{14}$ to $C_{20}$
Diesel Oils

480–610°F

Diesel Fuels

$C_{20}$ to $C_{50}$
Lubricating Oil

480–610°F

Lubricating Oils, Waxes, and Polishes

$C_{20}$ to $C_{70}$
Fuel Oil

610–700°F

Fuel for Ships, Factories, and Central Heating

>$C_{70}$ Residue

Asphalt for Roads and Roofing

Fractions increasing in density and boiling point

Crude Oil Inlet

To the Vacuum Fractionator

▲ **Figure 2.** One of the first major units in a refinery, the atmospheric crude fractionator typically operates at 90°F up to 700°F. It separates crude oil into components based on their boiling points.

the first step in the separation and recovery of the cracked hydrocarbon vapors from the reactor. It is similar to an atmospheric fractionator, but with two key differences: the vapors must be cooled before any fractionation can begin, and significant quantities of light gases pass overhead with the gasoline. The reaction products enter the main FCC fractionation column at high temperatures, 900–1,000°F. The products that are withdrawn from the fractionator include heavy cat naphtha (HCN), light cycle oil (LCO), heavy cycle oil (HCO), and slurry oil from the column bottoms.

Large quantities of heavy oil are circulated over a series of bottom fractionator trays to cool the vapors and wash down entrained catalyst. The heat removed by the main column bottoms and the heavy oil is used for feed preheating, steam generation, and reboiler heating in the unsaturated-gas concentration unit, or some combination of the three. Gas compression and cooling are used to separate the light hydrocarbons from the net gas as a liquid. The remaining vapor phase is passed through two absorbers to further recover light hydrocarbons into the liquid phases leaving the absorbers. A series of distillation columns processes the liquids from the overhead receiver of the main fractionator to remove light gases (*e.g.,* butane) from the gasoline cut, and then the remaining gases are separated into alkylation feed (*e.g.,* olefin LPGs), ethane, and fuel gas. This gasoline intermediate product is known as FCC gasoline or sometimes "cat gas."

## Butane isomerization unit and alkylation unit

The alkylation unit combines an isobutene feed with olefins (*e.g.,* propylene or butylene) produced in the FCC unit to form larger molecules known as alkylate, which is used in gasoline blending to raise the fuel's octane. The alkylate component of gasoline has the highest quality of all refinery products. It has high octane (both research octane number [RON] and motor octane number [MON]), low selectivity (*i.e.,* the difference between RON and MON), low RVP, no aromatics, and no sulfur. These qualities make alkylate a valuable gasoline blending component. Alkylate is a main component added to raise the octane of gasoline and produce a premium grade (*e.g.,* 93-octane, compared to 87- or 89-octane for regular grades).

Because isobutane gives the alkylate a much better octane rating than *n*-butane, butane isomerization is required. The butane isomerization unit converts *n*-butane into isobutane with excess hydrogen and catalyst. The hydrogen is added to avoid carbon deposition on

the catalyst and to shift the reaction to generate more of the desired components. Next, a large distillation column known as a de-isobutanizer (DIB) separates *n*-butane from isobutane. The DIB is often the tallest column in the refinery — the large size is a reflection of how difficult it is to separate these butane isomers because of their close boiling points.

The isobutane from the butane isomerization unit, the olefins from the FCC, and sometimes olefins from the coker (discussed later) are the feeds to the alkylation unit. Alkylation units require an acid catalyst, either sulfuric acid or hydrofluoric acid. Both catalysts operate at low temperatures and high isobutane-to-olefin ratios to reduce side reactions and acid (catalyst) consumption. Without this high ratio, the olefins could easily react with each other to combine into much larger (unwanted) molecules, such as C12s, C16s, C20s, etc.

When butane isomerization and olefin alkylation take place in a single unit, the bottom of the column serves as a product separator to capture the alkylate for gasoline blending, *n*-butane is taken as a side draw, and the overhead is isobutane.

*Article continues on next page*
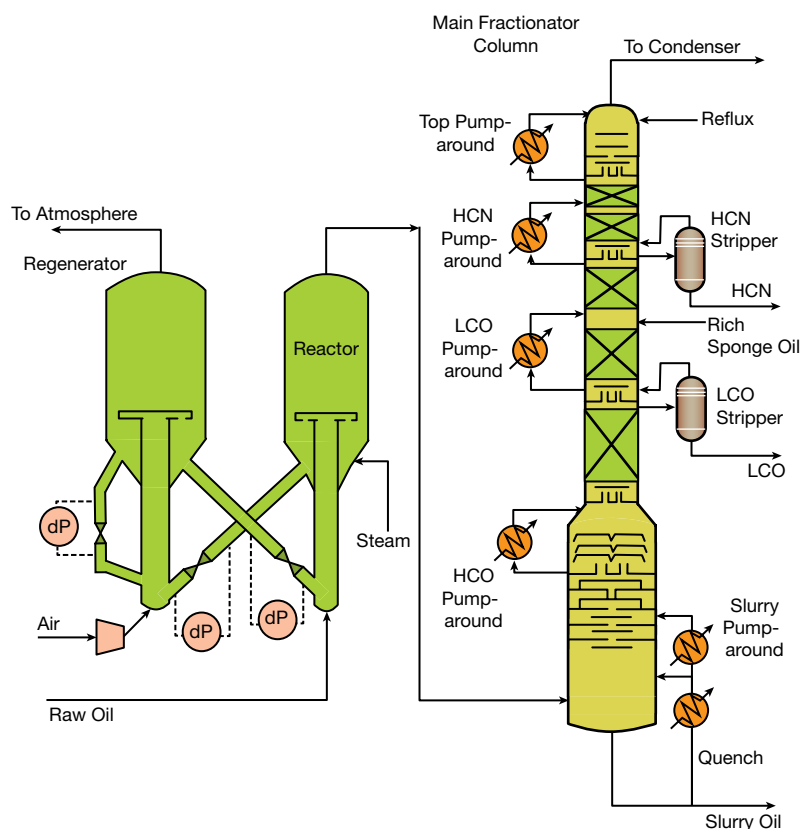
▲ **Figure 3.** The fluidized catalytic cracking (FCC) unit converts heavy vacuum gas oil (HVGO) from the vacuum fractionator into gasoline intermediates. The FCC is composed of a reactor, a catalyst regenerator, and a downstream main fractionating column, as well as a series of distillation columns.

## Light naphtha isomerization unit

The light naphtha isomerization unit saturates benzene and isomerizes light naphtha normal paraffins (pentanes and hexanes) into branched molecules that are higher-octane gasoline components for blending.

Like the butane isomerization unit, it converts *n*-pentane and *n*-hexane into isopentane and isohexane in the presence of excess hydrogen and catalyst. The light naphtha can be pretreated in a de-isopentanizer or de-isohexanizer column to separate the isoparaffins from the *n*-paraffins and shift the equilibrium reaction toward the isoparaffins.

After the isomerization reactors, a distillation column removes butane and lighter gases from the gasoline intermediate product, which is referred to as isomerate.
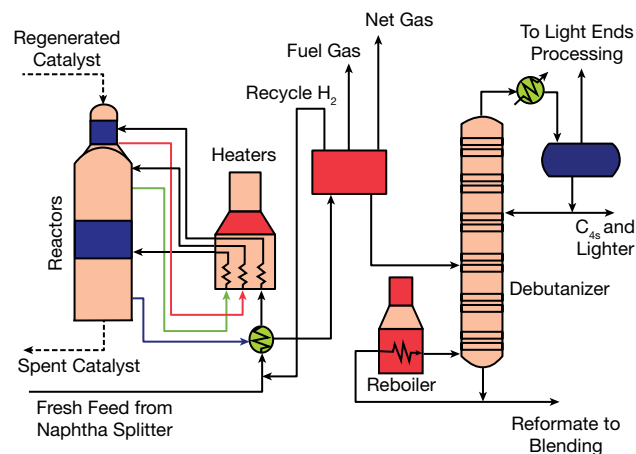


▲ **Figure 4.** The naphtha reformer converts naphthenes into aromatics to increase the gasoline octane rating or to obtain a feedstock for aromatics production. The reformer produces a gasoline intermediate known as the reformate.
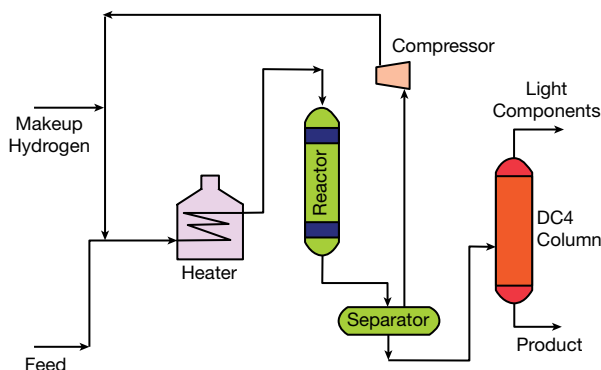


▲ **Figure 5.** The naphtha hydrotreater removes impurities from the heavy straight-run (HSR) naphtha prior to it entering the naphtha reformer. The DC4 column is a debutanizer that removes butanes from the naphtha product.

## Heavy naphtha reformer and hydrotreater

The heavy naphtha reformer (Figure 4) contributes to increasing the gasoline octane rating by converting naphthenes into aromatics and generating a reformate gasoline component. The reforming process is unusual in that the reaction is endothermic. It is carried out in a series of three or four reactors, each of which is preceded by a fired heater.

The catalyst needs to be regenerated. Two main process designs exist: semi-regeneration, which requires a shutdown to regenerate the catalyst *in situ* in the reactor; and continuous catalyst regeneration (CCR), in which flowing catalyst spheres circulate from the reactor to the regenerator and back. The CCR design is depicted in Figure 4.

The reforming process also generates the hydrogen needed by the refinery's hydrocracking and hydrotreating units, as well as benzene, toluene, and xylene (BTX) feedstocks for downstream petrochemical processing. Hydrotreaters (Figure 5) are located throughout the refinery. Their main purpose is to remove impurities such as sulfur, nitrogen, oxygen, metallic salts, olefins, and to a lesser extent aromatics (although saturating aromatics is usually not desired). They also serve as hydrodesulfurization units to remove sulfur from the FCC naphtha in order to meet regulatory requirements. The hydrotreating reaction is not intended to change the boiling range of the feed (unlike hydrocracking).

## Hydrocracking

Older hydrocracking units in the U.S. were designed primarily to maximize the production of gasoline and jet fuel. Newer hydrocracking units are more focused on maximizing ultra-low-sulfur diesel and jet fuel production, with less focus on gasoline (Figure 6).

Not all refiners have a hydrocracker, mainly because of the high capital costs associated with its construction and its high hydrogen consumption. Operating pressures can reach 3,000 psig, and piping, vessels, pumps, compressors, and instruments that can handle this high pressure are expensive. Additionally, a refiner would need to build a hydrogen generation plant or buy hydrogen for hydrocracker use.

Although the cracking reactions are endothermic, the overall process is highly exothermic due to the concurrent hydrotreating reactions, with reactor temperatures that can range from 550°F to 850°F. Cool hydrogen is added to the inlet of each catalyst bed in the reactor to control the reactor's

**TIM OLSEN** (Email: Tim.Olsen@emerson.com) has been with Emerson Process Management for 16 years. He is a refining consultant within the global refining industry solutions group, where he supports Emerson's technical and business strategy. He is a past chair of AIChE's Fuels and Petrochemicals Div., and was the Meeting Program Chair for the recent 2014 AIChE Spring Meeting in New Orleans, LA. He has a BS in chemical engineering from Iowa State Univ. with an emphasis in industrial engineering and process control, and an MBA from the Univ. of Iowa.

internal temperature. The addition of hydrogen increases the production of isoparaffins (which are desired over normal paraffins because they are higher-quality gasoline-blending components) and limits the production of olefins.

The hydrocracking unit includes a product fractionator to separate light gases, gasoline, and diesel. The bottoms product from that column is typically recycled for a second chance at reacting.

### Delayed coking and asphalt production

The delayed coker (Figure 7) thermally cracks heavy feedstocks to produce solid coke and lighter components that are blended into gasoline or fed to other processing units.

Delayed cokers are easy to identify from a distance — they look like a pair of oil derricks on top of a pair of vertical drums. Delayed cokers are always installed in pairs and their operation is cycled every 24 hr — one is online (as indicated by the solid lines in Figure 7) while the other is offline (dashed lines) so that the coke can be cut out of it.

The vacuum fractionator bottoms are sent with steam to a fired heater and then to the online coke drum, where the entrained liquid is thermally cracked to coke and other vapor products. As the coke drum fills, the cracked vapors

leave the top of the drum and are returned to the bottom of the fractionator, where the light intermediate products like gasoline separate as they travel up the column.

Refineries without thermal cracking units have the option of producing asphalt or selling the vacuum residuum to other refineries that can process it.

### Amine treating and sulfur recovery

An amine treating unit (not shown in Figure 1) captures the hydrogen sulfide from the refinery gas streams and concentrates it through absorption into an amine solution. These rich amine streams are routed to an amine regenerator column; the overhead vapor from this column contains the released hydrogen sulfide, which is converted into elemental sulfur in a sulfur recovery unit (also not shown in Figure 1). The bottoms from the amine regenerator is the lean amine, which is circulated back to the various amine absorbers in the refinery.

As regulations over the past few decades have placed stricter limits on sulfur emissions and reduced the amount of sulfur that fuel products may contain, the sulfur recovery units have required revamps for higher capacity and greater onstream availability.

*Article continues on next page*
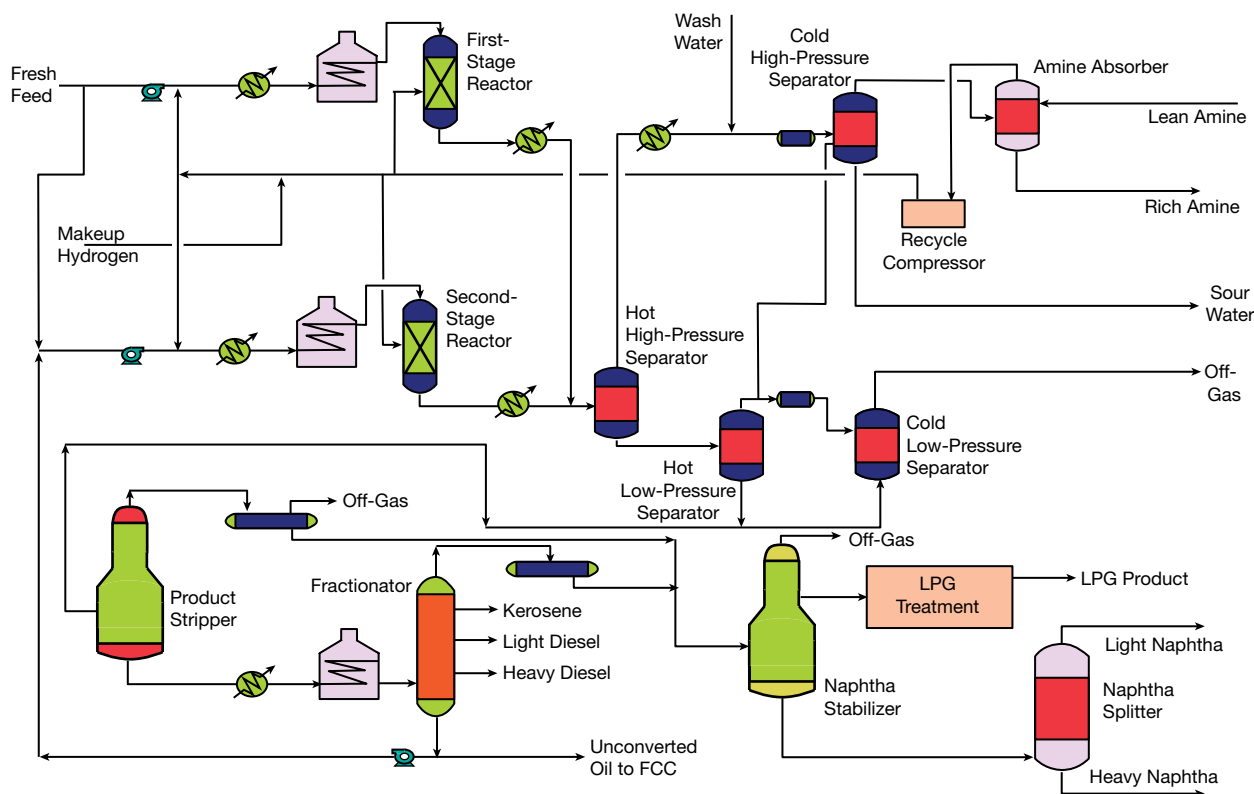


▲ **Figure 6.** Hydrocracking units operate at high pressures (up to 3,000 psig) and are big hydrogen consumers. They maximize ultra-low-sulfur diesel and jet fuel production.

## Blending

Various fuel components are blended into final commercial products that meet particular specifications based on the consumer's location. In the U.S., there are many different location-dependent gasoline specifications. For example, large urban areas typically have more-restrictive fuel specifications.

Intermediate products include alkylate, isomerate, reformate, cat gas, hydrocracker gas, and coker gas. Normal butane is also added as needed to meet RVP specifications, which are different in summer and winter, and in winter vary by location.

Diesel is also blended from each of the various units that produce intermediate products, such as the crude unit, FCC, hydrocracker, and delayed coker.

## Tank farm and distribution

Finally, transportation fuel products need to be distributed to the consumer. These products are typically transported by pipelines to terminals. At the terminals, ethanol and other biofuels, as well as various additives, are added, and then the gasoline or diesel is trucked to commercial distribution stations. The additives account for the differences among various brands of fuel.

The refinery's tank farm has various sizes of tanks to store crude oil coming into the refinery, intermediate products (*e.g.,* petrochemical feedstocks), liquefied petroleum gases, and fuel products.

The next time you fill up your vehicle, you will have a greater understanding of the path the crude oil had to travel to become a fuel.

CEP
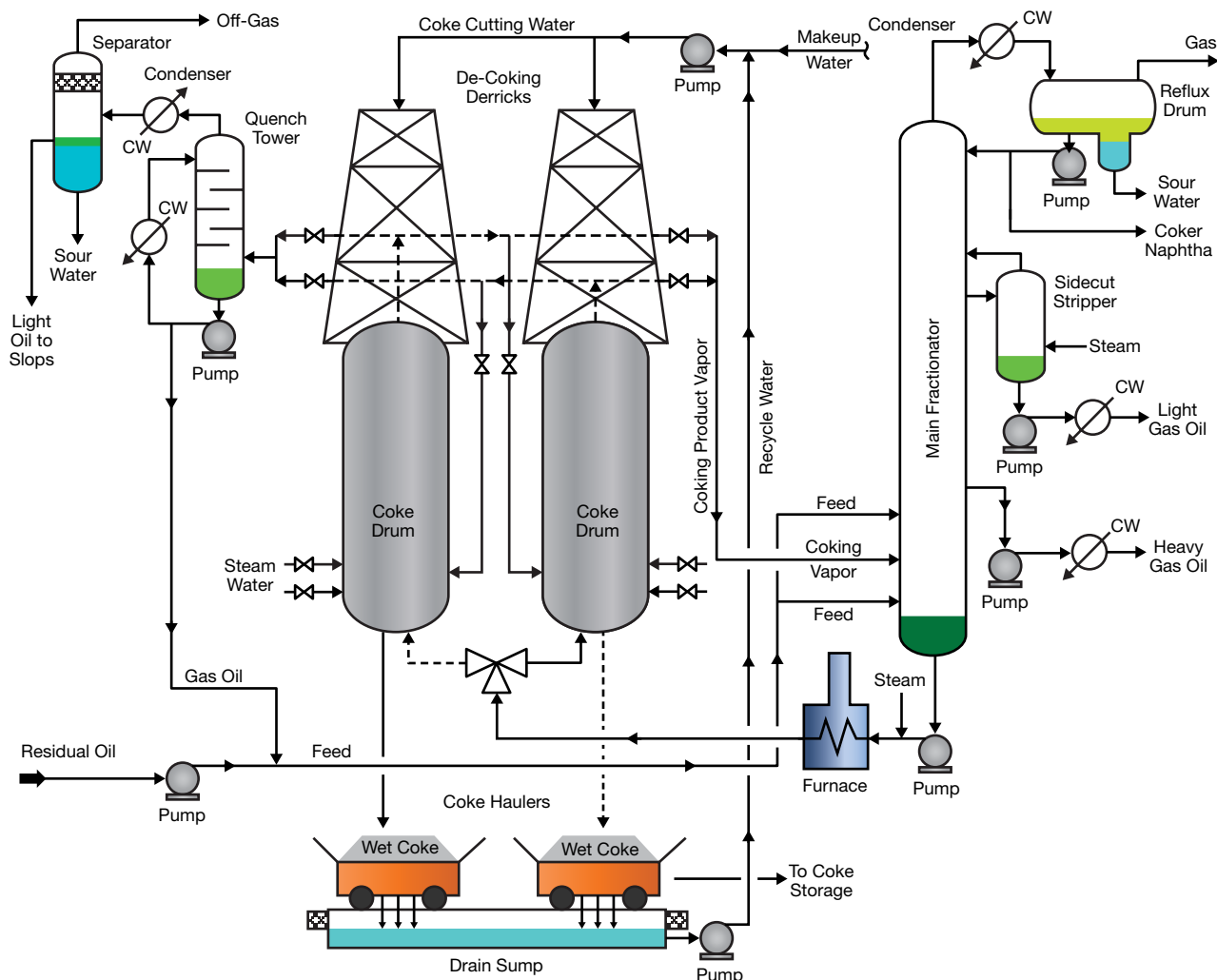


▲ **Figure 7.** The delayed coker thermally cracks heavy feedstocks from the vacuum fractionator to generate solid coke and gasoline intermediates. Delayed cokers are always installed in pairs. Their operation is cycled every 24 hr — one is online (as indicated by the solid lines) while the other is offline (dashed lines). The quench system in the upper left corner is used only when a drum is being steamed prior to de-coking. Image by Milton Beychok (http://en.wikipedia.org/wiki/File:Delayed_Coker.png).

# Fermentation Fundamentals: Brewing Bugs for Bioengineering

**Sujata K. Bhatia, M.D., P.E.**
Harvard Univ.

A fermentation process leverages live cells to manufacture a variety of products. Choosing the correct host cell, as well as the optimal production equipment and method, are the fundamental first steps of these complex processes.

Fermentation, broadly defined as the cultivation of live cells to produce useful molecules, is arguably the oldest chemical engineering process on earth. The word fermentation is derived from the Latin verb *fervere,* which means "to boil." Fermentation, with its bubbling culture media and animated cells, also conjures images of commotion, excitement, and agitation. Living cells are themselves miniature chemical reactors, capable of performing complex reactions and manufacturing valuable products, such as pharmaceuticals, plastics, and biofuels *(1)*.

Both prokaryotic (bacterial) cells and eukaryotic (yeast and mammalian) cells are incredibly versatile for chemical production. Yet live cells also demand exquisite control systems in order to optimize cellular growth, maximize the quantity of the desired output, ensure product quality and reliability, and minimize side reactions. The biochemical engineer must therefore carefully choose both the host cell and the fermentation process for cellular cultivation.

This article describes the fundamentals of host cell selection, process design considerations, and equipment selection for fermentation processes. It compares and contrasts bacterial cells, yeast cells, and mammalian cells in fermentation. It also discusses the U.S. Food and Drug Administration's (FDA) requirements for fermentation equipment in pharmaceutical and biologic production, according to principles of good manufacturing practice (GMP). Finally, the article discusses important challenges and future directions in research and development of fermentation processes, including algal fermentation.

## Choosing a host: The cellular reactors

Fermentation relies on live cells to carry out complex chemical synthesis and breakdown reactions. Bacteria, yeast, and mammalian host cells are all commonly employed in biochemical engineering processes, and each of these cell types has distinct capabilities, advantages, and disadvantages. In general, the biochemical engineer faces a trade-off between low-cost, well-characterized, conveniently grown prokaryotic cells on the one hand, and highly versatile, complex, reliable eukaryotic cells on the other. The desired product will ultimately guide the selection of the host cell.

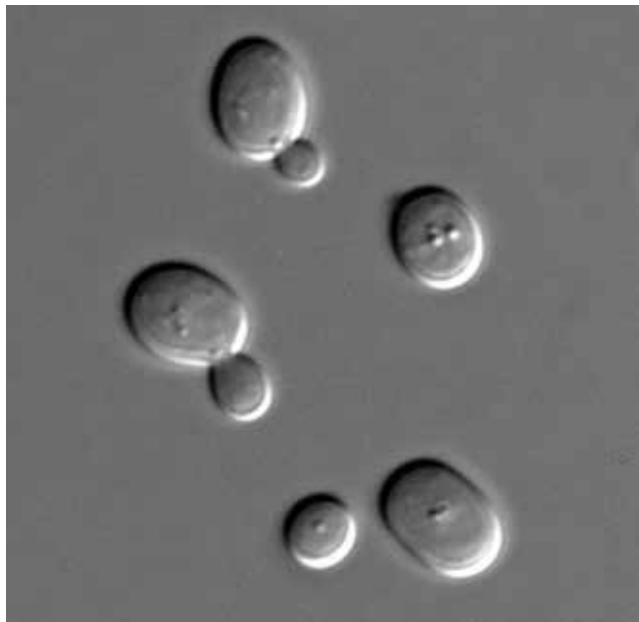*Bacterial cells* maximize convenience, speed, and
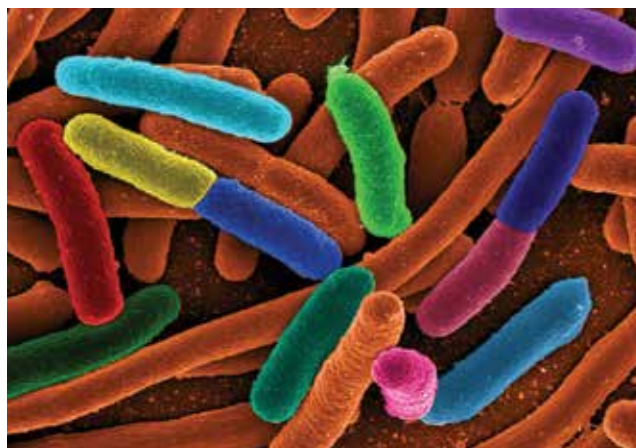


▲ **Figure 1.** *Escherichia coli* bacterial cells exhibit a relatively high growth rate and are the most frequently used bacteria for fermentation.

cost-effectiveness. The most frequently used bacterium for fermentation is *Escherichia coli* (Figure 1). These bacteria are advantageous because their physiology and genetics are far better understood than those of any other organism; their genome has been mapped *(2)*. *E. coli* exhibit a relatively high growth rate, doubling every 20 to 30 min, and survive on simple and inexpensive culture media. During fermentation, *E. coli* are able to grow to high cell concentrations (>50 g dry weight of cells per liter of culture media) and are extremely productive, with the ability to produce 25–50% or more of their total protein as the desired protein product *(3)*. *E. coli* can therefore achieve high productivity without sacrificing cell growth.

However, these cells do not normally secrete and release the proteins they produce; this lack of secretion complicates recovery and purification of desired products. *E. coli* may actually degrade proteins or form insoluble bodies that further complicate product recovery. Most importantly, *E. coli* and other bacterial hosts cannot perform glycosylation reactions or other modifications to proteins after translation has occurred.

Bacterial cells are typically chosen as fermentation hosts for the production of specialty chemicals, biofuels, and other non-food, non-biologic products.

*Yeast cells* are able to carry out more complex reactions than bacterial cells, but there are sacrifices with regard to speed and cost. *Saccharomyces cerevisiae*, commonly called baker's yeast, are frequently employed in fermentation (Figure 2). A particular advantage of *S. cerevisiae* yeast cells

is that the FDA considers these organisms to be generally regarded as safe (*i.e.,* they have GRAS status) *(4)*, so that food additives, food ingredients, and nutritional supplements manufactured via yeast fermentation may readily gain FDA approval. Yeast cells are relatively low-cost and have a reasonably high growth rate, doubling every 1 to 2 hr, and they can perform simple glycosylation reactions.

However, it is difficult to achieve protein production levels in yeast that are as high as those in *E. coli*. In addition, bottlenecks in secretion can occur with yeast production hosts, limiting cell productivity and complicating product recovery *(5)*.

Because of their GRAS status, yeast cells are often chosen as fermentation hosts for food and nutritional products.

*Mammalian cells* for fermentation processes are typically selected from cell lines that have been immortalized. Normally, mammalian cells do not proliferate indefinitely, but immortalized cell lines carry mutations that allow for an indefinite number of cycles of cell division and proliferation. This enables the cultivation of immortalized cell lines for prolonged periods *in vitro*. Immortalized cell lines for mammalian fermentation cultures include Chinese hamster ovary (CHO) cells (Figure 3) and baby hamster kidney fibroblast (BHK-21) cells. Mammalian cells can perform complex glycosylation reactions, and readily excrete most proteins, which avoids the bottlenecks associated with yeast. Of all the fermentation hosts discussed in this article, mammalian cells produce a product that most closely resembles its natural counterpart.

Mammalian cells, however, are by no means convenient for fermentation. Mammalian cells grow slowly, doubling approximately every 24 hr, and the media ingredients for mammalian culture are very expensive. Mammalian cells cannot achieve high protein production levels — less than

▲ **Figure 2.** *Saccharomyces cerevisiae* yeast cells can carry out more-complex reactions than bacterial cells and are classified as generally regarded as safe (GRAS).
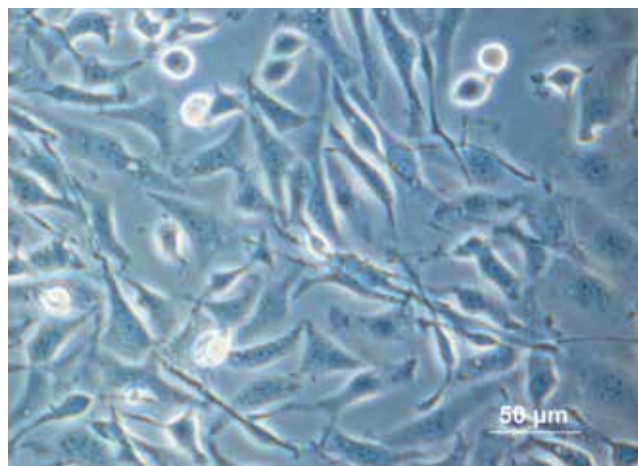
▲ **Figure 3.** These Chinese hamster ovary (CHO) mammalian cells are part of an immortalized cell line — they carry mutations that allow them to divide and proliferate indefinitely.

5% of the total protein they produce is the desired protein product *(3)*. Finally, since immortalized cells are mutants, these mammalian cells require extreme care during purification processes to avoid any contamination of the desired product with genetic material from the mutated cells. This makes mammalian cell fermentations incredibly costly in terms of both time and resources.

Mammalian cells are chosen as fermentation hosts when the authenticity of the product must be complete and guaranteed. Mammalian cells should be seriously considered for production of human therapeutic proteins and other biologic therapeutics.

### Scaling it up: The fermentation reactor

Fermentation reactors, often called bioreactors, can operate at the laboratory scale (Figure 4) up to the pilot-plant or plant scale (Figure 5). The traditional fermenter is a stirred tank reactor. A sparger supplies gas under pressure to the tank, and an impeller disperses gas bubbles throughout the tank. A disc stirrer or turbine stirrer is used for mixing, and baffles are included to facilitate gas dispersion and mixing *(3)*. The fermenter will usually contain four baffles, each with a width of 8% to 10% of the reactor diameter. Fermen-

ters are typically constructed with a height-to-diameter ratio of 2:3, and the impeller diameter is usually 30% to 40% of the tank diameter.

Fermentation processes can be designed for batch, fed-batch, or continuous operation. A batch process, with no inflows of nutrients nor outflows of waste products, allows maximum containment of the live cells, but minimal control over cellular proliferation. During batch fermentation, cell growth rates start off slow in a lag phase, and then accelerate quickly into an exponentially increasing growth phase. Eventually the cells exhaust the available nutrients, and cellular growth levels off into a final stationary phase.

Batch processes are associated with significant periods of nonproductive time, as the batch culture must be stopped once the cells reach the stationary phase. The product must be removed and the media must be replaced before the fermentation can be restarted. The environment within a batch reactor is difficult to control as well: heat output, acid or base production, and oxygen consumption vary from low rates at the beginning of the fermentation to very high rates during the exponential growth phase *(6)*. Finally, batch fermentation carries the risk that the desired product may itself be depleted, as cells may utilize the product as an



▲ **Figure 4.** Mammalian cell culture in a bench-scale bioreactor.



▲ **Figure 5.** This mobile pilot-plant fermenter has a 90-L capacity for the production of cellulosic ethanol. Source: U.S. Dept. of Energy.

energy source once the available nutrient supplies have been exhausted.

A continuous fermentation process, with both inflow of nutrients and outflow of waste products, has very significant productivity advantages. The continuous fermentation operates at steady state: production and consumption rates of all materials are constant. Since the continuous process avoids the stationary phase of cell growth, there is no need to stop and restart the process; continuous operation thus minimizes nonproductive time. For a culture with a 1.5-hr doubling time and a 20-hr batch cycle, a continuous system has a 14-fold productivity advantage over a batch system *(3)*. Moreover, a continuous fermenter does not experience swings in heat output, acid or base production, or oxygen consumption by cells, so the environment within a continuous bioreactor is much more easily controlled.

However, continuous bioreactors do have notable disadvantages. While the production of desirable products is more efficient in the continuous reactor, the production of growth-associated byproducts is also more efficient *(6)*. Continuous systems may also be more susceptible to contamination, and the cells themselves may be more difficult to contain.

A fed-batch process is intermediate between a batch process and a continuous process. Fed-batch fermenters, which enable inflow of fresh nutrients with continuous or periodic withdrawal of broth, can overcome the major limitations of batch processes without the disadvantages of continuous processes. Fed-batch fermentation can extend the productive period of a traditional batch process and achieve a substantial level of process control, yet it does not require any additional special equipment beyond that required for batch fermentation *(7)*. The fed-batch process can be designed to keep the cellular growth rate constant, and to maintain the nutrient supply at a low, constant concentration. The constantly replenished nutrient supply enables cells to proliferate and reach high cell densities within the bioreactor, and controls deviations in the growth pattern of the cells. Fed-batch fermentations also limit byproduct formation, control product concentration, and avoid product depletion. Most importantly, fed-batch processes allow containment of organisms and avoid contamination.

## Making it right: Good manufacturing practice

The goal of GMPs is to ensure the quality of medicines, as well as some food products *(8)*. Because fermentation is increasingly being applied to the production of small-molecule drugs, biologic therapeutics, vaccines, nutritional supplements, and other medicinal and food products, it is critical for chemical engineers to consider GMP guidelines when selecting fermentation equipment.

GMP regulations are issued by the FDA and are laid out in Section 21 of the Code of Federal Regulations, Parts 210 and 211. They are enforced via inspections of manufacturing facilities; failure to comply with GMP requirements can result in regulatory actions against manufacturers, and can even jeopardize FDA approval of a new drug.

Chemical engineers must recognize that GMP regulations are continually evolving to meet the demands of new technologies; for this reason, GMP is often denoted as cGMP, meaning current good manufacturing practice. Also keep in mind that GMP regulations represent the minimum requirements for a compliant process; many companies choose to exceed these standards.

To allow manufacturers the maximum flexibility in equipment selection and process design, the FDA does not

### LITERATURE CITED

1. **Bhatia, S. K.,** "Biology as a Basis for Biochemical Engineering," *Chem. Eng. Progress,* **109** (7), pp. 40–44 (July 2013).
2. **Perna, N. T.,** *et al.,* "Genome Sequence of Enterohaemorrhagic *Escherichia coli* O157:H7," *Nature,* **409,** pp. 529–533 (Jan. 25, 2001).
3. **Shuler, M. L., and F. Kargi,** "Bioprocess Engineering: Basic Concepts," Prentice-Hall, Englewood Cliffs, NJ (1992).
4. **U.S. Food and Drug Administration,** "Microorganisms and Microbial-Derived Ingredients Used in Food (Partial List)," FDA (Aug. 21, 2013).
5. **Niebauer, R. T., and A. S. Robinson,** "*Saccharomyces cerevisiae* Protein Expression: From Protein Production to Protein Engineering," in "Protein Expression Technologies: Current Status and Future Trends," Fracois Baneyx, ed., Taylor and Francis, Norfolk, U.K. (2004).
6. **Stanbury, P. F.,** *et al.,* "Principles of Fermentation Technology," 2nd Ed., Butterworth Heinemann, New York, NY (1999).
7. **Longobardi, G. P.,** "Fed-Batch versus Batch Fermentation," *Bioprocess Engineering,* **10,** pp. 185–194 (May 1994).
8. **U.S. Food and Drug Administration,** "Facts About Current Good Manufacturing Practices (cGMP)," www.fda.gov/drugs/developmentapprovalprocess/manufacturing/ucm169105.htm (May 2, 2013).
9. **U.S. Food and Drug Administration,** "Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance – Equipment," www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm124777.htm (Sept. 25, 2013).
10. **U.S. Food and Drug Administration,** "Code of Federal Regulations, Title 21, Part 211, Section 211.65," www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=211.65 (April 1, 2013).
11. **Muchemu, D. N.,** "Change Control for FDA Regulated Industries: A Risk Assessment Approach," AuthorHouse, Bloomington, IN (2007).
12. **Pienkos, P. T., and A. Darzins,** "The Promise and Challenges of Microalgal-Derived Biofuels," *Biofuels, Bioproducts and Biorefining,* **3,** pp. 431–440 (May 28, 2009).
13. **Wilson, S. A., and S. C. Roberts,** "Recent Advances towards Development and Commercialization of Plant Cell Culture Processes for the Synthesis of Biomolecules," *Plant Biotechnology Journal,* **10,** pp. 249–268 (April 2012).

maintain a list of approved cGMP manufacturing equipment. Instead, the cGMP standards require that equipment be appropriately designed for its intended use, and that equipment be designed for thorough cleaning and maintenance *(9)*. The equipment surfaces in contact with the starting materials, in-process materials, or products must be nonreactive, nonadditive, and nonabsorptive to ensure that the equipment surfaces do not "alter the safety, identity, strength, quality, or purity of the drug product beyond the official or other established requirements" *(10)*. To guarantee inertness and reliable cleaning of equipment, engineers typically choose stainless steel, Hastelloy alloy, or glass-lined steel units as cGMP fermentation reactors.

GMP also requires documentation of any changes to the fermentation process or equipment; this is known as change control. Change control procedures apply to changes in operating conditions, standard operating procedures, manufacturing facilities, raw materials, production equipment, technical specifications, software, and quality assurance protocols *(11)*. A rule of thumb is that change control applies to any change that affects one of the five inputs of a process, also known as the five M's: man, material, method, machine, and Mother Nature. The goal of change control procedures is to limit risk by assessing the impacts of any process changes. Whenever engineers introduce a process alteration, the change must be documented and reported, and the adverse impacts on the safety, quality, efficacy, potency, and purity of the product must be evaluated and appropriately mitigated.

## Forthcoming advances in fermentation

A well-designed, well-controlled fermentation process leverages live cells to manufacture specialty chemicals, polymers, therapeutic drugs, biologics, and foods and nutritional supplements. Depending on the desired product, the chemical engineer may select bacteria, yeast, or mammalian cells as production hosts, and can choose to design a batch, fed-batch, or continuous fermentation process.

Increasingly, fermentation processes are being recognized as critical components for the establishment of a sustainable bio-economy, and fermentation operations are being designed to tackle issues of energy and the environment. For instance, algal fermentation processes (Figure 6) are under development for the production of biodiesel, green diesel, green jet fuel, and green gasoline *(12)*. Plant cell fermentation processes are also under consideration for the manufacture of complex drugs and proteins *(13)*.

The main challenges for chemical engineers will be to increase the yield of these processes, improve the cost-effectiveness, and ensure that the manufacturing operations are scalable and reliable. With appropriate attention to engineering design, fermentation processes will address needs in healthcare, nutrition, and sustainability.  **CEP**

▲ **Figure 6.** Green algae in a fermentation bioreactor at the laboratory scale. Image courtesy of Umberto Salvagnin, Dept. of Sustainable Agro-Ecosystems and Bioresources, Centro Ricerca e Innovazione, Italy.

**SUJATA K. BHATIA, M.D., P.E.,** is a physician and bioengineer who serves on the biomedical engineering teaching faculty at Harvard Univ. (Phone: (617) 496-2840; Email: sbhatia@seas.harvard.edu). She is the Assistant Director for Undergraduate Studies in Biomedical Engineering at Harvard. She is also an Associate of the Harvard Kennedy School of Government for the Science, Technology, and Globalization Project, as well as a faculty member in the Harvard Kennedy School Executive Education program on Innovation for Economic Development. She received bachelor's degrees in biology, biochemistry, and chemical engineering, and a master's degree in chemical engineering, from the Univ. of Delaware, and she received an M.D. and a PhD in bioengineering, both from the Univ. of Pennsylvania. Prior to joining Harvard, she was a principal investigator at the DuPont Co., where her projects included the development of bioadhesives for wound closure and the development of minimally invasive medical devices. She has written two books, *Biomaterials for Clinical Applications* (a textbook that discusses opportunities for both biomaterials scientists and physicians to alleviate diseases worldwide) and *Engineering Biomaterials for Regenerative Medicine*. She received an award from the Harvard Univ. President's Innovation Fund for Faculty in recognition of her innovative approaches to biomedical engineering education, the John R. Marquand Award for Exceptional Advising and Counseling of Harvard Students, and the Capers and Marion McDonald Award for Excellence in Mentoring and Advising. She is a member of AIChE and is a registered P.E. in the state of Massachusetts.

# Compressed Air Basics

**Michael L. Stowe, P.E.**
Advanced Energy

Compressing air is an inefficient and expensive process. This article offers some tactics to optimize your compressed air system, which can help reduce energy costs at your plant.

Compressed air is an integral part of most manufacturing processes. However, it is one of the most inefficient, expensive, and misused utilities in manufacturing plants. Compressed air is often considered a free commodity at the point of use. But by the time air is compressed, cooled, dried, transported, regulated, and then finally used, it is anything but free. Electric costs for air compressors can account for up to 30% of a manufacturing site's total electric bill.

If you must use compressed air, there are certain things that should be done to ensure the most efficient compressed air system possible for your circumstances.

This article explains why compressing air is so inefficient and why the costs for this process are so high. It describes the basics of optimizing compressed air systems using tactics such as reducing air leaks, properly training operators, sustaining proper maintenance practices, and conducting an end-use survey.

## Compressing air: An inefficient process

An air compressor increases the pressure of inlet air by reducing its volume. The majority of air compressors have, at their core, either centrifugal impellers or rotary screws (Figure 1) that compress the air.

By the very physics and thermodynamics involved, compressing air is naturally inefficient. Most of the electrical power consumed by the motor that drives the air compressor heats the air. The air that exits the compressor must then be cooled, which requires a fan and an air- or water-cooled heat exchanger, which consumes even more energy.

Once the air is compressed, it has to be delivered at a certain pressure to the end user. As the air is transported, losses occur and inefficiencies arise along the way. At the end of the line, the compressed air is often used incorrectly, or for purposes not intended by the original designer.

In the end, only about 5–10% of the original energy input performs useful work in manufacturing processes. Figure 2 breaks down the energy consumption for a 100-kW energy input *(1)*. In addition to the inefficiencies inherent in the physics and thermodynamics of air compression, there are many other system inefficiencies, which are explored later in this article.

*Example cost breakdown.* Let's consider a 100-hp air compressor that consumes 0.746 kW per horsepower. It runs for 8,760 hr/yr (the plant operates 24 hr/day, 365 days per year), with a 75% load factor (*i.e.,* fully loaded 75% of the time) and a 93% electric motor efficiency. If energy costs $0.075/kW on average, what is the annual cost of running this compressor?

$$\text{Cost} = \frac{(100 \text{ hp})(0.746 \text{ kW/hp})(8{,}760 \text{ hr/yr})(0.75 \text{ load factor})(\$0.075/\text{kWh})}{0.93 \text{ motor efficiency}}$$

$$= \$39{,}526/\text{yr}$$

In round numbers, this 100-hp air compressor would cost approximately $40,000/yr just for the electricity. A plant with 1,000 horsepower of air compressors operating at the same conditions would spend around $400,000/yr.

To help make air compression somewhat more economical, consider setting your machines to switch off when they are not being used. Depending on your shift pattern, turning your compressors off during evenings and weekends could reduce your energy bills significantly *(2)*.

Figure 2 shows that most of the energy spent to run a compressor creates heat. As much as 90% of the heat in compressed air can be recovered for such uses as preheating water for hot water heaters or supplemental building heat *(2)*.

## Stop misusing compressed air

Misuse of compressed air contributes to the inefficiency and the expenses involved. Operators on the plant floor might think of compressed air as a free commodity and, using their creativity, think of all sorts of things to do with it. Compressed air is used in many applications even though a different method would be far more economical. And, many applications can be done more effectively or more efficiently using a method other than compressed air *(2)*.

Compressed air is frequently used to blow water or dirt off of manufactured parts, similar to the way a leaf blower is used to clean a driveway. These compressed air clean-ups, called "blow-offs," typically include a pipe or a hose supplied with compressed air that feeds to one or more outlets or nozzles. Blow-offs are basically just big air leaks.

Based on useful work delivered, a 1-hp compressed air motor consumes seven times as much energy as a comparable 1-hp electric motor. Thus, the electrical costs of using compressed air can be significantly more than those of an alternative method, as the following two examples illustrate.

*Example 1: Personal cooling.* I have seen compressed air used for personal cooling in some plants. In one case, a worker in a hot area took a compressed air source and directed the output onto himself to keep cool (Figure 3). To calculate the personal cooling cost of compressed air, assume:

• as a rule of thumb, a 1/8-in. hole in a 100-psi compressed air line costs $1,000/yr for the electricity

• five 1/8-in. holes are used for the personal cooling

• the personal cooling is used during half of the year, in the hot months.

$$\text{Cost}_{\text{PC-CA}} = \left(5 \text{ holes}\right)\left(\$1{,}000/\text{yr per hole}\right)\left(0.5 \text{ yr}\right)$$
$$= \$2{,}500/\text{yr}$$
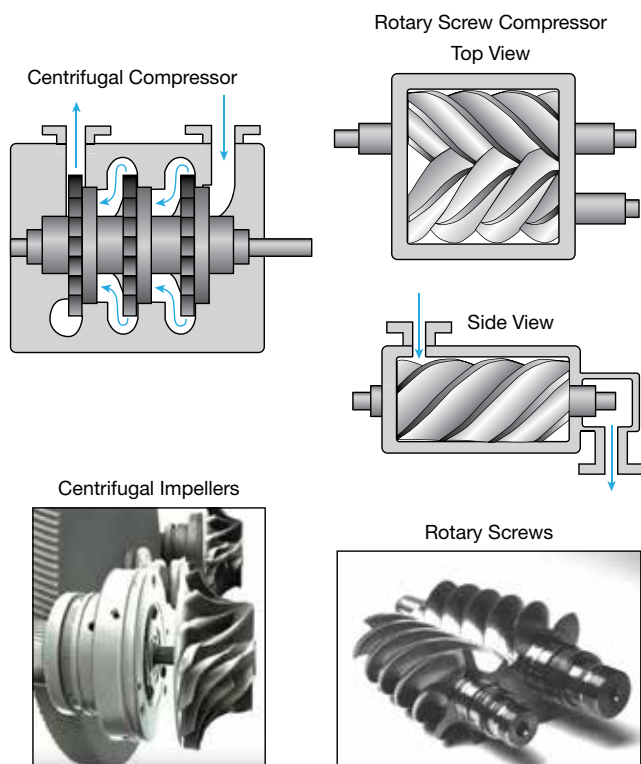


▲ **Figure 1.** Most air compressors use either centrifugal impellers (left) or rotary screws (right) to compress air.
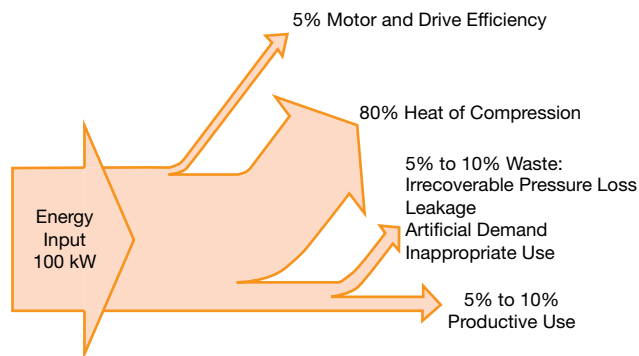


▲ **Figure 2.** Only about 5–10% of the energy input to a compressor goes toward productive use of compressed air.
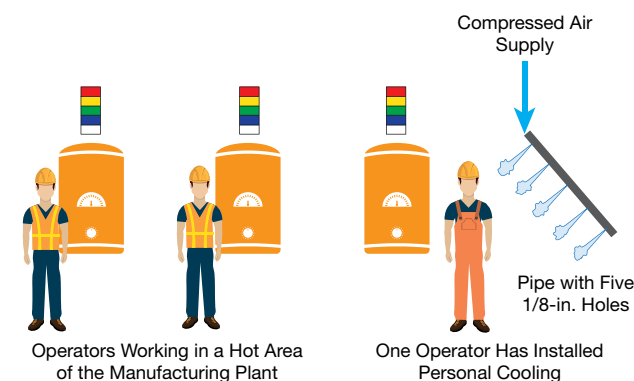


▲ **Figure 3.** Using compressed air for personal cooling will cost the example plant $2,500/yr on average. In comparison, using an electric fan for cooling will cost only about $150/yr.

In comparison, to calculate the cost of an electric fan for personal cooling, assume:

- a 0.50-hp fan motor
- the motor is 82.5% efficient
- the motor has a conversion factor of 0.746 kW/hp
- the fan runs during half of the year, 4,380 hr/yr
- electricity costs $0.075/kWh.

$$\text{Cost}_{\text{PC-EF}} = \frac{(0.5 \text{ hp})(0.746 \text{ kW/hp})(4,380 \text{ hr/yr})(\$0.075/\text{kWh})}{0.825 \text{ motor efficiency}}$$
$$= \$149/\text{yr}$$

The electrical energy cost for personal cooling with compressed air is nearly 17 times higher than that of cooling using an electric fan.

*Example 2: Blow-off.* Blow-offs are frequently used to remove moisture or debris from parts during the manufacturing process (Figure 4). To calculate the cost of a blow-off using compressed air, assume:

- as a rule of thumb, a 1/8-in. hole in a 100-psi compressed air line costs $1,000/yr for the electricity
- five 1/8-in. holes are used for the blow-off
- the blow-off is in operation on all shifts, all year.

$$\text{Cost}_{\text{BO-CA}} = (5 \text{ holes})(\$1,000/\text{yr per hole})(1 \text{ yr})$$
$$= \$5,000/\text{yr}$$

There are alternatives to using compressed air for a parts blow-off. A small, low-pressure blower may be sufficient and much more economical. In comparison, to calculate the cost of a low-pressure blower for a blow-off, assume:

- a 2-hp low-pressure blower motor
- the motor is 82.5% efficient
- the motor has a conversion factor of 0.746 kW/hp
- the blower runs almost all the time, 8,000 hr/yr
- electricity costs $0.075/kWh.
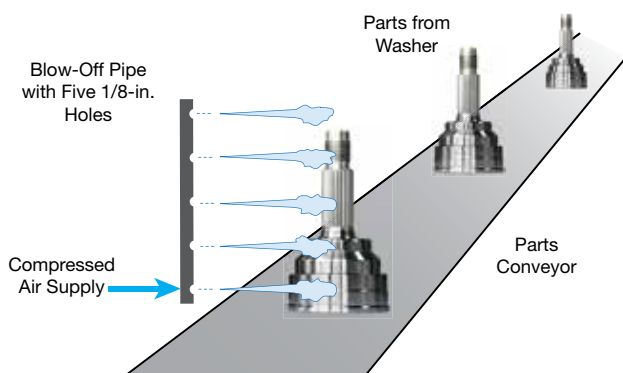


▲ **Figure 4.** Using compressed air for a part blow-off will cost this plant $5,000/yr. However, using a low-pressure blower will only cost the plant $1,085/yr.

$$\text{Cost}_{\text{BO-LPB}} = \frac{(2 \text{ hp})(0.746 \text{ kW/hp})(8,000 \text{ hr/yr})(\$0.075/\text{kWh})}{0.825 \text{ motor efficiency}}$$
$$= \$1,085/\text{yr}$$

The electrical energy cost for the blow-off with compressed air is just over 4.5 times higher than that of using a low-pressure blower. It is clear from these two examples that it is much more economical to use direct electrically driven devices than compressed air wherever possible.

Table 1 lists some other possible misuses of compressed air and suggests alternatives *(3)*.

### Take a systems approach to compressed air

We have clearly established that compressed air systems are inefficient, expensive, and often misused. Taking a systems approach can help you devise solutions to these challenges.

A systems approach entails looking at the entire compressed air system, from start to finish, when considering optimization. Just looking at a single air compressor or an end use alone may be beneficial, but an individual improve-

| Table 1. Many plants use compressed air inappropriately. Fortunately, alternatives are available. | |
|---|---|
| **Potentially Inappropriate Uses** | **Suggested Alternatives** |
| Clean-up, drying, process cooling | Low-pressure blowers, electric fans, brooms, engineered nozzles |
| Sparging | Low-pressure blowers and mixers |
| Aspirating, atomizing | Low-pressure blowers |
| Padding | Low- to medium-pressure blowers |
| Vacuum generator | Dedicated vacuum pump or central vacuum system |
| Personal cooling | Electric fans |
| Open-tube, compressed-air-operated vortex coolers without thermostats | Air-to-air heat exchanger or air conditioner; add thermostats to vortex cooler |
| Air-motor-driven mixer | Electric-motor-driven mixer |
| Air-operated diaphragm pumps | Proper regulator and speed control; electric pump |
| Idle equipment* | Put an air-stop valve at the compressed air inlet |
| Abandoned equipment** | Disconnect air supply to equipment |
| * Equipment that is temporarily not in use during the production cycle.<br>** Equipment that is no longer in use, either due to a process change or malfunction | |

ment opportunity may be masked by a problem in another part of the system. The best plan is to look at the compressed air system from the air inlet of the compressor all the way to the end use of the air on the plant floor, including everything in between (Figure 5).

Taking a systems approach can help eliminate ongoing system problems. For example, one plant had a history of poor air pressure on the plant floor, especially in the areas that were farthest from the compressor room. The plant had tried to remedy this problem by adding additional air compressors. At the time of my visit, they had 1,800 horsepower of air compressors, but the poor air delivery performance problem persisted.

A detailed total system review revealed that even though more air compressors had been added, the piping delivery system had never been upgraded. This plant was trying to push 1,800 horsepower worth of compressed air through a piping system that was originally designed for 600 horsepower. If they had taken a total systems approach before adding additional air compressors, they would have realized a piping system upgrade was necessary.

## Understand demand and supply

Both the supply and demand sides of a compressed air system are important, and they should be considered together when making decisions about the compressed air system (such as choosing the compressor type and capacity, location, receiver tanks, piping size and arrangement, etc.). For example, if you are purchasing an air compressor for the supply side, you should consider the end-use loads, filters, and delivery piping on the demand side.

An understanding of how the components on the supply and demand sides interact is essential to the design and operation of an efficient compressed air system. Here are some points to keep in mind.

*Supply-side: Air compressor.* Evaluate the compressor installation's location, connection to cooling water, and ventilation. Because compressors are often located in boiler rooms or rooms with inadequate ventilation, it is essential that adequate ventilation be arranged to prevent the accumulation of heat within the room.

It is often recommended that ambient air going to the compressor inlet be taken from outside the building. The coolest location is normally a north-facing wall, because it has the least direct sun, although this may not be the most convenient location. In addition, a means of preventing rainwater entry should be included *(3)*.

*Supply-side: Aftercooler.* An increase in relative humidity can increase the amount of condensate that must be handled by aftercoolers. The aftercooler and separator efficiencies, cooling effectiveness, and condensate separation effectiveness should all be taken into account *(3)*.

*Supply-side: Dryer.* The dryer size, pressure drop, and efficiency should be measured and evaluated, then compared to the current application. A common recommendation is to add a filter upstream of the compressed air dryer, to prevent particulates and slugs of liquid condensate from entering the dryer. In the case of a regenerative-desiccant dryer, placing a coalescing filter before the dryer also minimizes contamination of the desiccant bed by lubricant carryover *(3)*.

*Supply side: Automatic drains.* Evaluate the location, application, and effectiveness of both supply-side and demand-side drains *(3)*.

*Storage: Receiver tanks.* Consider the volume and locations of tanks for compressed air storage. This can be highly process-dependent. There should be good primary storage on the supply side, properly sized for the air compressors. Secondary storage may also be needed throughout the plant, as well as at some high-demand locations.

*Demand side: Piping distribution system.* Consider the overall layout of the distribution system piping. Ensure that the piping size and arrangement match the compressor sizes on the supply side. Sometimes, a loop distribution pipeline makes more sense than a straight-line distribution pipe. Specifying a 3-in.-dia. pipe instead of a 2-in.-dia. pipe can reduce pressure drops by up to 50% *(2)*. Shortening the distance air has to travel can further reduce pressure drops by about 20–40% *(2)*.

The effectiveness of the condensate removal system should also be evaluated. The main compressed air headers should be slightly sloped, with a pitch of about one inch
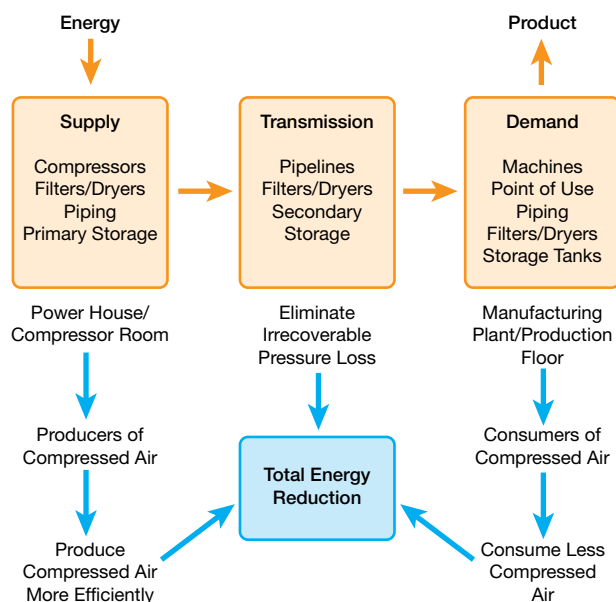


▲ **Figure 5.** Considering the entire compressed air system, rather than just a single compressor or a single piece of process equipment, can help plants reduce energy use *(4)*.

per ten feet of pipe to allow water and condensate to drain out. Drains should be located at low points in the header. Feeder pipes to equipment should come out of the top of the supply header to help prevent moisture from reaching the point of use (Figure 6).

*Demand side: Load profile.* Estimate the compressed air load profile, *i.e.,* how the demand in cubic feet per minute (CFM) changes over time. A facility with a varying load profile will likely benefit from advanced control strategies. A facility with short periods of heavy demand may benefit from implementing storage options.

*Demand side: Air treatment.* Compressed air systems require an appropriate level of air treatment for proper operation of the end-use equipment. Air treatment includes filtration and drying. Different manufacturing processes require different levels of air treatment. Air quality levels should be measured at critical points in the system. Air that contains excessive moisture or that is not filtered properly can damage equipment and reduce system efficiency. In some cases, only certain end-use equipment requires highly treated air, so you might need a system that provides different levels of treatment at different points in the system *(3)*.

### Controls and variable-frequency drives

Matching the compressed air system supply and demand can be a tricky problem. Ideally, the air compressors would be controlled and optimized to run absolutely as little as possible, while at the same time providing high-quality air at the correct pressure to all demand-side end users. The overall efficiency of a plant's compressed air system is directly related to the sequencing and control of the individual air compressors.

Air compressors are generally most efficient when they are running fully loaded. Running multiple air compressors at partial loading is very inefficient. An ideal situation is to have one air compressor (or several) fully loaded and one additional air compressor with a variable-frequency drive



▲ **Figure 6.** The main compressed air headers should be sloped to allow condensate to drain. The compressed air piping that is being fed to the end user should come out of the top of the header to further eliminate entrained moisture in the air.

(VFD) to handle the variable load on the demand side.

Information on the demand load profile is very valuable for setting up the air compressor control system and determining the feasibility of using a VFD.

The existing control system should be evaluated to determine whether it is appropriate for the system demand profile. Estimate possible performance gains that could be achieved by operating the system in a different mode or using an alternative control strategy *(3)*. For example, check to see if your control system could be optimized to operate in one mode on first and second shift and in another mode on third shift and the weekends. Your compressed air control system should be based on your operational schedule and your compressed air demand profile.

### Maintaining compressed air systems

Maintenance is key to sustaining successful long-term operation and maximizing equipment life. Proper air compressor maintenance can cut energy costs by approximately 1% and help prevent breakdowns that result in downtime and lost production *(2)*.

Here are a few things to consider for the maintenance of your compressed air system.

*Filters.* Just as you change the oil filter in your car at scheduled intervals to ensure optimum performance, be sure to change the filters in your air compressor and air system regularly to ensure air quality and to prevent excessive pressure drops *(2)*.

*Condensate drains.* Are your condensate drains stuck open? If so, you could be wasting compressed air.

Go one step further and replace timer drains with zero-loss drains to save more air *(2)*. Timer drains use a solenoid valve that is opened by a timer for a few seconds every hour or so to bleed out moisture. This type of drain opens directly to the compressed air lines and air is lost during each cycle. Zero-loss condensate drains operate similarly, but do not release any air during each drain cycle — they only release the collected moisture and are therefore much more economical. For both types, if something prevents the solenoid valve from reseating, or if the open signal stays engaged, a significant amount of air can be lost. Both types should be checked frequently.

*Cooling water quality.* Various cooling water sources are found in industrial plants. Where city water is used, the quality of the water is generally good, but in some areas, treatment for hardness may be needed to avoid fouling of cooling surfaces. Water analysis should be conducted to determine any treatment needs *(3)*.

*Motor and air-end bearings.* Make sure all bearings are greased properly.

*Air/oil separator.* The air/oil separator, which is typically located at the exit of the air compressor, should be
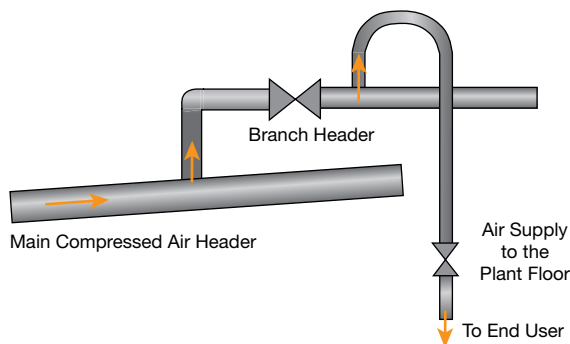
> Air leaks deserve the same level of attention and quick repair as any other type of leak.

cleaned periodically. This unit separates the compressor's lubricating oil from the compressed air. It is a source of pressure drop, and if not maintained properly, pressure drop will increase, requiring more energy from the air compressor motor.

*Heat exchangers and coolers.* All coolers and heat exchangers on the supply side should be cleaned periodically.

*Controls.* Verify that the sequencing and operation of the air compressors matches the operational compressed air demand profile.

Again, the better you maintain your total compressed air system, the more efficient and cost effective it will be.

### Eliminate air leaks

Do you hear hissing when you walk through your plant? That is the sound of money going down the drain — or rather venting to atmosphere. It costs about $1,000 per year in electricity to compress the amount of air that escapes through a 1/8-in.-dia. air leak. In a large manufacturing plant, that can add up to a lot of wasted energy and money.

At one plant, a compressed air system leak test during a shutdown revealed that 450 horsepower of air compressors were needed just to hold the compressed air header pressure at 100 psi with no demand on the system. That translates to about $180,000/yr in electricity cost just to supply the air leaks in the plant. That is a huge and completely avoidable cost.

Leaks are common at many locations in a plant, such as:
• threaded pipe fittings
• condensate drains that are stuck open
• compressor blow-downs that are stuck open
• hoses and fittings
• air-driven hand tool connections
• valve stems
• air-driven spray applicators
• plastic bowls on the bottom of filter/regulator/lubricators (FRLs) (Figure 7).

There are several approaches you can take to remedy air leaks. First, establish a formal air leak search, tag, and repair program. Prioritize leak repairs by size, and repair or fix the biggest ones first. Make all operators and maintenance personnel aware of the significant cost of compressed air leaks, and set up a recognition and reward program to encourage employees to find and repair leaks. Keep your compressed air dry and filtered to help eliminate dust and sludge, which helps prevent the formation of new leaks *(2)*.

Approximately 80% of air leaks are not audible, so you may need a third-party auditor to help detect leaks using an ultrasonic leak detector (or you can buy one and use it yourself) *(2)*.

Why do air leaks tend to get ignored? I think it is because they are somewhat invisible. If water or hydraulic oil leaks, it creates a visible spot and a possible safety hazard on the floor. Liquid leaks tend to get cleaned up immediately and repaired quickly. Try to make everyone aware that air leaks deserve the same level of attention and quick repair. Make sure air leaks have the same level of importance as any other type of leak.

### Conduct an end-use survey

To verify that the total compressed air system is as efficient as possible, you must ensure that it is being used for appropriate applications. Conduct a plant-wide compressed air end-use survey and identify each end use. Rank each end use by size, in terms of the required volume and the required supply pressure (psi). You can use this information to help you optimize the compressed air system performance.

Compressed air is a very expensive resource, and one of the best ways to save energy is to operate the air compressor at the lowest possible discharge pressure. A facility can save 1% in compressed air energy costs for every 2 psi reduction in the compressor discharge pressure *(4)*.

One plant had an operation that required 125-psi compressed air supply pressure in order to operate properly. The plant ran the entire compressed air system at 125 psi, which was very expensive. One recommendation was to buy a small standalone air compressor specifically for the machine that required the 125-psi air and reduce the pressure of the plant-wide compressed air system to 100 psi. This small change would save an estimated 12.5% in air compressor energy costs per year.

Use your end-use survey to see if it makes sense for your plant to have one or more discrete standalone compressors or storage tanks for specific end uses so that you can optimize the rest of your compressed air system.

▶ **Figure 7.** The plastic bowls on the bottom of filter/regulator/lubricators (FRLs) are notorious for leaks. Establishing a formal air leak search, tag, and repair program can help document and remedy FRL leaks.

*Article continues on next page*

Plants typically have hoses with nozzles for various functions. Try to eliminate these completely, but if you must use them, consider using engineered nozzles for blow-offs. Engineered nozzles use much less air than regular nozzles. Some electric utilities pay an incentive for installation of engineered nozzles for compressed air end use.

### Involve everyone in compressed air improvements

The front line in compressed air system improvements is not the maintenance department. It is the operators who use the air on a daily basis.

Most manufacturing plants provide annual training on various topics, particularly safety and hazardous materials. Make compressed air systems a part of annual training. Encourage operators to look for ways to reduce demand for compressed air. Emphasize the following points to all employees:
- Compressed air is not free.
- A 1/8-in.-dia. air leak or equivalent air misuse costs about $1,000/yr in electricity.
- Compressed air should not be the default choice — always look for alternatives.

Train operators on your air leak search, tag, and repair program, and seek suggestions for improvements from the operators. Involve everyone in the plant in reducing compressed air demand, including the accounting, purchasing, and product design departments, as well as manufacturing engineers, operators, maintenance technicians, and facilities technicians.

While there are hundreds of articles that deal with the technical specifics of compressed air improvement, the business side must always be considered. Make sure that compressed air system optimization projects get into the capital budget cycle and do your homework to show the costs and benefits of an optimized compressed air system.

### Next steps

Think about where you and your plant are with regard to your compressed air system. Are you in pretty good shape, or could you make some improvements? In either case, here are some things to consider doing:
- Conduct a compressed air system audit.
- Set up a formal leak detection and repair program.
- Make sure you are looking at both the supply side and demand side.
- Conduct a plant-wide compressed air end-use survey.
- Eliminate improper uses of compressed air.
- Conduct compressed air awareness training.

This article has just touched the tip of the iceberg of what you can do to optimize your compressed air system. Many resources and consultants are available to help you further improve your compressed air system. I highly recommend subscribing to the magazine *Compressed Air Best Practices* and its e-newsletters. This publication has many good articles on ongoing compressed air system improvements. Also, check out the U.S. Dept. of Energy's Compressed Air tools and resources *(1, 3–5)*.

There are many things that you can do before you have to spend the capital for adding an air compressor to your system. Make choosing air as an energy source a last resort that is given serious consideration by all involved, not the default choice just because it seems to be a free commodity. Compressed air is anything but free. **CEP**

### Literature Cited

1. **U.S. Dept. of Energy,** "Compressed Air Assessment Basics," Federal Energy Management Program, Office of Energy Efficiency and Renewal Energy, www1.eere.energy.gov/femp/pdfs/compair_webinar_05162011.pdf (2011).

2. **Arfalk, E.,** "Energy Savings Checklist for Your Compressed Air System: Top 10 Ways to Increase Compressed Air Energy Efficiency," 5th Utility, Atlas Copco Compressors, www.thecompressedairblog.com/energy-savings-checklist-for-your-compressed-air-system (Mar. 2015).

3. **U.S. Dept. of Energy,** "Energy Tips — Compressed Air: Eliminate Inappropriate Uses of Compressed Air," Compressed Air Tip Sheet #2, Industrial Technologies Program, Office of Energy Efficiency and Renewable Energy, https://energy.gov/sites/prod/files/2014/05/f16/compressed_air2.pdf (Aug. 2004).

4. **U.S. Dept. of Energy,** "Energy Tips — Compressed Air: Compressed Air System Control Strategies," Compressed Air Tip Sheet #7, Industrial Technologies Program, Office of Energy Efficiency and Renewable Energy, https://energy.gov/eere/amo/downloads/compressed-air-system-control-strategies (Aug. 2004).

5. **U.S. Dept. of Energy,** "Improving Compressed Air System Performance: A Sourcebook for Industry," 3rd ed., Office of Energy Efficiency and Renewable Energy, https://energy.gov/eere/amo/downloads/improving-compressed-air-system-performance-sourcebook-industry-third-edition (2016).

**MICHAEL L. STOWE, P.E., CEM, PEM,** is a senior energy engineer with Advanced Energy in Raleigh, NC (www.advancedenergy.org; Email: mstowe@advancedenergy.org; Phone: (919) 857-9043), where he works with utilities, industrial equipment vendors, and manufacturing plant teams to find the best technical and most energy efficient solutions for industrial processes. He has over 28 years of experience with many industrial manufacturing processes and plants, including roles as production manager, maintenance manager, and plant engineer. He has developed and delivered numerous presentations on industrial energy efficiency and related topics for utilities, industrial trade associations, and technical groups. He is a licensed Professional Engineer in North Carolina, a Certified Energy Manager (CEM), a Professional Energy Manager (PEM), a Certified Practitioner Energy Management Systems-Industrial (CP EnMS-Industrial), and a Superior Energy Performance, Performance Verifier (SEP PV). He is a member of several trade associations: American Foundry Society (AFS), ASM Heat Treat Society (ASM-HTS), Association of Energy Engineers (AEE), Chemical Coaters Association International (CCAI), Industrial Heating Equipment Association (IHEA), and RADTECH International. He holds a BS in mechanical engineering from Georgia Institute of Technology, and attended U.S. Navy, Nuclear Power School and Submarine Officer Training.

# An Introduction to Mammalian Cell Culture

Tung S. Le
Meghan McCann
Samira M. Azarin
Wei-Shou Hu
Univ. of Minnesota

**Mammalian cell culture is at the core of biomanufacturing therapeutic proteins and viral vaccines. Find out how mammalian cells are derived and cultivated, and what opportunities this field holds.**

Cells isolated from animal tissues can be expanded in culture for use as a research tool, for the production of virus vaccines and various therapeutic proteins, and to generate functional cells or tissue analogues for regenerative medicine. Chemical engineers are actively involved in harnessing the full potential of mammalian cells, especially with regard to process design and optimization.

In the past quarter century, cells derived from animals, especially mammals like rodents and humans, have become a major vehicle for producing biologics, a class of medications that includes vaccines and various proteins used in treating cancer, genetic diseases, and other ailments.

Mammalian cells can be made to produce vaccines through viral infection, and therapeutic proteins through genetic engineering. Many of these medicines are necessary for patients who either lack the normal form of a protein or cannot produce it in sufficient quantity. For example, patients with Gaucher's disease, a congenic disorder characterized by a lack of the functional enzyme β-glucocerebrosidase, can be treated with Cerezyme, a recombinant enzyme produced in mammalian cells *(1)*.

Other therapeutic proteins include antibodies and specific binding proteins that neutralize disease-causing molecules within the body. For example, the drug Etanercept (trade name Enbrel) binds to tumor necrosis factor (TNF), thus preventing it from causing an inflammatory reaction in rheumatoid arthritis patients *(2)*.

Human cells, in particular, are poised to enable opportunities in cell-based therapy and regenerative medicine. We now have the capability to derive stem cells from many sources and guide them to become specific cell types for clinical applications.

This article describes how mammalian cells are derived, their utility, and the processes that harness their full potential.

## Normal life span and cell line derivation

During the early stages of development, animal cells undergo extensive proliferation and differentiation while developing into different tissues and organs. In an adult, the vast majority of cells are quiescent; although they are metabolically active and perform their physiological roles, such as filtration in the kidneys or synthesis and chemical transformation in the liver, most are not actively dividing. Most normal adult cells only divide in response to stimuli to replenish old or damaged cells. Only cells in specific tissues, such as skin or epithelial intestinal cells, divide regularly.

The body has over 200 different types of cells, many of which cannot be excised and grown in culture. Cells that are more amenable to culture include fibroblasts and certain epithelial cells. A first step in cell isolation is to explant a tissue in a physical and chemical environment suitable for those cells to survive and proliferate.

A permissive environment for cell growth requires a complex mixture of nutrients, including sugars, amino acids, vitamins, minerals, and growth factors such as insulin. Except for certain cell types in blood, cells derived from tissues are anchorage-dependent, meaning they do not grow as free-floating individual cells. Therefore, after being released from the tissue environment, cells require a surface on which they can attach, otherwise they will fail to survive and divide.

After attachment, cells grow and expand onto empty surfaces until the entire surface is covered in a layer that is one cell thick (*i.e.,* a monolayer). At this point, they stop dividing and reach a state called contact inhibition. Next, an enzyme, such as trypsin, is used to degrade the proteins that "glue" the cells to the surface, thereby releasing the cells into solution. Once detached, the cells can be transferred to a culture vessel with a larger surface area to resume growth.

This cycle of attachment, cell expansion, and detachment can repeat many times, with each cycle comprised of multiple cell divisions. However, most normal cells have an internal clock that counts their own doublings. Cell division stops once the so-called Hayflick limit is reached *(3, 4)*. Most cells derived from tissues can divide up to 40–60 times before ceasing to proliferate (becoming senescent) and exhibiting abnormal appearance. Nevertheless, the number of doublings that these cells can sustain in culture is sufficient for vaccine production applications.

The senescence and contact inhibition exhibited by these cells are hallmarks of cells from normal tissues. Certain cells isolated from cancers, however, are immortal and can overcome contact inhibition. More than a half-century ago, scientists succeeded in isolating cells that survived senescence *(5, 6)*. These cells continued to divide after all others died. Interestingly, unlike cancer cells, some of the surviving cells still obeyed contact inhibition and looked morphologically normal.

These immortal cells that bypass the Hayflick limit and continue to divide are called cell lines and are immortal in culture, unlike cell strains isolated from normal tissues. Cell strains and cell lines differ in another important way: All of the cells in cell strains have normal chromosomes with two sets per cell, while cells from cell lines do not typically have two sets of chromosomes, even if they are normal morphologically.

Many cell lines induce tumor formation when injected into immunocompromised mice. However, because they can be cultured forever, cell lines can be genetically engineered to produce a product in virtually unlimited quantities. For this reason, all of the therapeutic proteins produced in mammalian cells employ cell lines.

## Stem cells

With the appropriate chemical and physical environment, cells isolated from different tissues can retain their important functional properties. For example, for a period of time, cultured hepatocytes isolated from livers can continue to produce albumin and other proteins, as well as metabolize some xenobiotics. However, most of these differentiated cells have a very limited capacity for expansion in culture. They are valuable research tools, but their limited expansion capability diminishes their utility for transplantation back into patients with the goal of augmenting or replacing an ailing organ.

Many organs in the human body have a small amount of stem cells that can differentiate to become mature cells within the organ. Although few, these stem cells provide the body with some capacity for repair, maintenance, or even regeneration of tissue. Stem cells in adult tissues are somewhat restricted, as they can typically differentiate only to cells of their own lineage.

The most notable stem cells are the hematopoietic stem cells in bone marrow. Four decades ago, scientists realized that the stem cells or progenitor cells in a patient's bone marrow could be transplanted into another patient to repopulate the constituent cells in the recipient *(7)*.

During a very early stage of embryo development, when the embryo contains only 70–100 cells, there is a special population of cells (stem cells) that have the capacity to differentiate into any cell type in adult organs. This quality, known as pluripotency, occurs before the embryo has implanted. These highly potent cells are transient; following a short period of limited expansion, they continue to differentiate into all of the cell types in the developing embryo, and thus lose their pluripotency.

In the early 1980s, scientists succeeded in isolating pluripotent cells from early-stage mouse embryos and were able to culture and maintain them in a state that preserved their potency *(8)*. These embryonic stem cells can be expanded in an undifferentiated state over tens of cell divisions, while maintaining the capacity to differentiate to all developmental lineages. Their chromosomes were normal and diploid, like classical cell strains, but they represented a distinct new class of cells.

It took over a decade before human embryonic stem cells were isolated and grown in culture *(9)*. This advance offered great potential for regenerative medicine; scientists hoped that diseased tissues could be replaced with differentiated

cells or engineered tissues derived from stem cells. However, this technology faced significant ethical challenges, as the cells are derived from very early-stage human embryos, typically fertilized embryos that remain unused after *in vitro* fertilization procedures.

In spite of the controversy, our understanding of the human genome and epigenome has advanced over the past decades. Every cell in a person's body has the same genome, but cells in different tissues express their genes at different levels, giving them tissue-specific physiological activities and functions. The cues for using or expressing the genome are embedded in epigenetics, which describes the different ways DNA can be packed, and specific chemical modifications that can occur on different genomic locations *(10, 11)*. Only some regions of the genome are accessible for expression, whereas other regions are condensed and inaccessible.

In 2006, Shinya Yamanaka (Kyoto Univ., Japan) and his colleagues succeeded in changing the epigenetic state of differentiated cells to a pluripotent state by introducing four genes into the recipient cell *(12)*. This new class of stem cells is called induced pluripotent stem cells (iPSC) *(13)*. Since then, many methods of "reprogramming" various differentiated cells to pluripotent cells have been reported. This new technology of generating pluripotent cells without the use of embryos has altered the landscape of regenerative medicine. The prospect of deriving stem cells from adults for possible cell transplantation or other therapeutic uses is now possible.

### Cells as production vehicles: Therapeutic proteins and viral vaccines

The commercial value of biologics for both vaccines and therapeutic proteins exceeds US$100 billion/yr worldwide.

Most viral vaccines are produced by infecting cells with a virus, so that the cells produce large quantities of that virus. After isolation and product purification, but before formulation and final packaging, the virus is inactivated by treatment with formalin. Although the virus has been inactivated, its surface antigen can still elicit immune response in humans, thus providing protection against future viral infection. In some cases, the virus is engineered so that it is not infectious to humans and no inactivation step is required during production.

Although the growth rate of mammalian cells is two orders of magnitude slower than that of fast-growing bacteria, mammalian cells are employed for protein production when the protein requires complex modifications (*e.g.,* oligosaccharide addition or the formation of complex protein structures) that cannot be accomplished in a microorganism. To produce a recombinant protein, a cell is first engineered to incorporate the product gene into its

genome. After the introduction of the product gene, extensive selection and screening isolate a high-production cell line for manufacturing. Today, a top cell line can produce a protein at levels that rival the protein secretory cells in our body, such as liver cells with albumin secretion or plasma cells with antibodies.

Viruses are also used as delivery vehicles that shuttle DNA or RNA into cells of target tissues as a therapy to rectify genetic deficiencies in cells. The virus produced in culture is injected into the patient and delivered to a target tissue or organ where it enters the cells and produces the desired protein from the engineered virus' genome. The exogenous protein, which, for example, may be an enzyme lacking in the patient, enables the recipient cells to perform normal physiological functions.

For such gene therapy applications, the amount of virus required is substantially larger than that for vaccination. The road to gene therapy has been arduous because a large dose of virus carries many off-target effects. Nevertheless, regulatory approval of gene therapy has been granted in the European Union. Accordingly, the use of cell culture for gene therapy applications will likely increase in the near future.

### Cells as a product themselves: Cellular therapy

While cell bioprocessing has traditionally focused on cells as production vehicles for proteins or viruses, there is a growing number of applications in which the cells themselves are delivered as the therapy. The concept of cellular therapy has been around for many decades; bone marrow transplants are performed routinely and involve the transfer of bone marrow stem cells from a donor to the recipient.

Many other tissues in the body contain stem cells that are not pluripotent, but rather multipotent, meaning although they cannot differentiate into any type of cell in the body, they are capable of producing multiple cell types within a specific lineage. These adult stem cells have a higher proliferative capacity than fully differentiated cells within the tissue, though their proliferative capacity is still lower than that of pluripotent stem cells.

As the tools for isolating and controlling the behavior of iPSCs and adult stem cells improve, the use of these cell types as sources for cellular therapy will become more widespread. In 2013, there were 69 cell therapies at various stages of clinical trials *(14),* and this number will likely continue to grow. Given the unique bioprocessing challenges presented by growth, differentiation, and purification requirements of cell-based therapeutics, chemical engineers have a critical role to play in moving these therapies from the benchtop to industry.

## Surface requirements for cultivation

The majority of cells isolated from tissues require a surface for adhesion, although some are capable of growing in suspension, especially those isolated from blood (Figure 1). Although cultivating adherent cells is rather straightforward in the lab using Petri dishes or flat flasks, it can be a challenge in large-scale operations. For this reason, cells that are used to manufacture therapeutic proteins are modified to enable them to grow in suspension.

Taking cues from blood cells that can grow in suspension, scientists realized that cell lines of significance in the production of therapeutic proteins can be made to grow in suspension if they are gradually deprived of a surface and forced to grow. This realization meant that stirred-tank bioreactors could be leveraged for cell cultivation. A typical manufacturing process involves growing cells in suspension within a series of reactors with increasing scale to increase the number of cells until they reach the production scale, at which point the cells are maintained at a high enough concentration to allow for product accumulation.

Unfortunately, many cells cannot be forced to grow in suspension and still maintain their differentiated properties or potency. For products manufactured in smaller quanti-
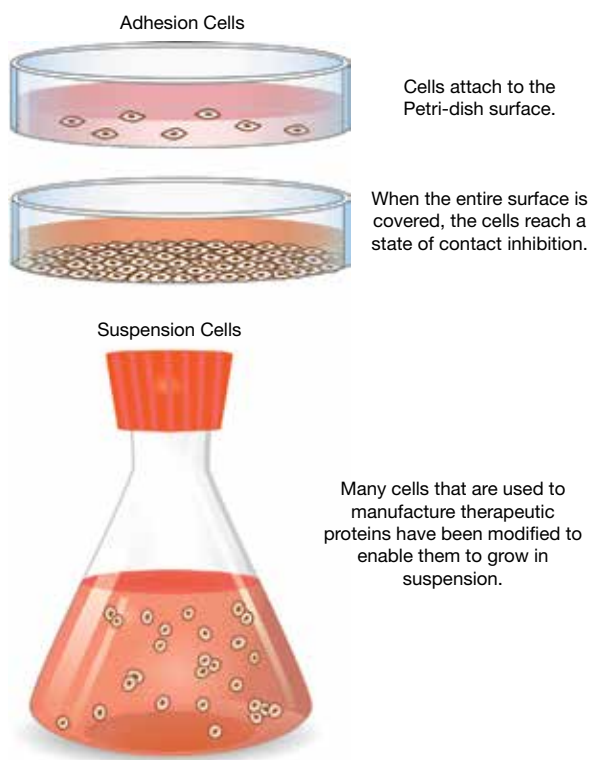
ties, flask-based or large flat-surface-based culture methods are still manageable and employed. This is the case for many vaccines and for patient-specific applications that use a patient's own cells. For many other processes, a scalable cultivation method must be developed.

One method of growing anchorage-dependent in-suspension cells employs small beads called microcarriers. These microcarriers provide a surface for cell attachment and growth while allowing the bead-and-cell complexes to be suspended in a stirred-tank bioreactor (Figure 2). Due to their size — as well as their material composition and structure — the beads have a very slow terminal settling velocity and therefore require minimal agitation to remain in suspension, which minimizes mechanical damage to the cells. Microcarriers are widely used in vaccine production and are becoming more useful for cell therapy applications.

## Product purification standards

Traditionally, cells are grown in a medium containing sugar, amino acids, and balanced salts, and are supplemented with animal serum, often fetal bovine serum.

Within a vaccine, the amount of antigen (*i.e.,* virus) needed for each immunization is very small, just enough to elicit an immunogenic response. Given the small size of the required dose, any trace quantities of impurities or contaminating materials carried along with the product can be tolerated, eliminating the need for extensive purification prior to formulation of the final product.



Adhesion Cells

Cells attach to the Petri-dish surface.

When the entire surface is covered, the cells reach a state of contact inhibition.

Suspension Cells

Many cells that are used to manufacture therapeutic proteins have been modified to enable them to grow in suspension.

▲ **Figure 1.** Most cells isolated from tissues are anchorage-dependent and require a surface for adhesion. Cells that are used to manufacture therapeutic proteins are modified to eliminate their need for surface attachment.
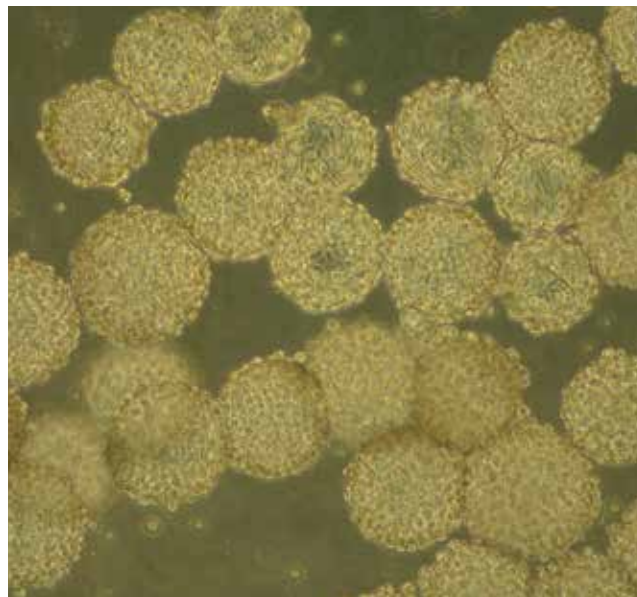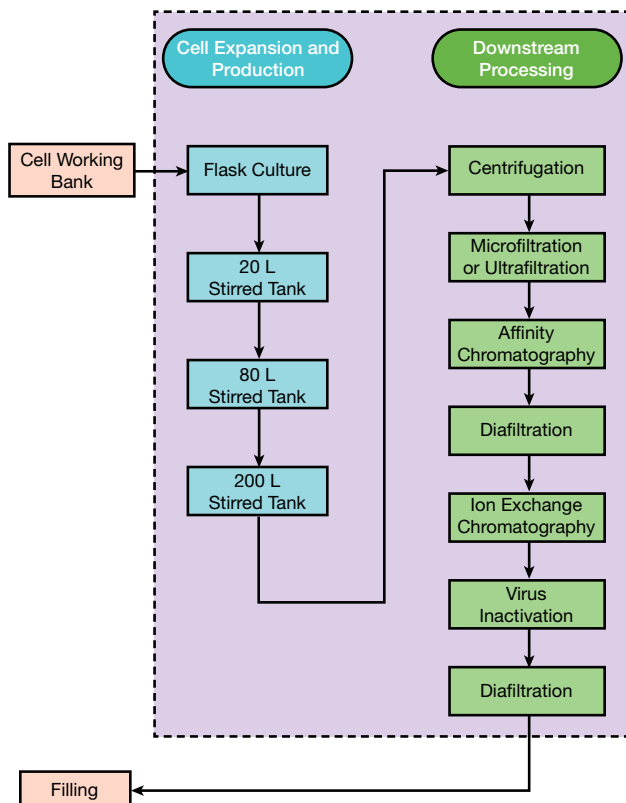


▲ **Figure 2.** Human embryonic stem cells can be grown on beads called microcarriers, which provide a surface for cell attachment and growth that can be suspended in a stirred-tank bioreactor. In this figure, the microcarriers are completely covered by the stem cells.
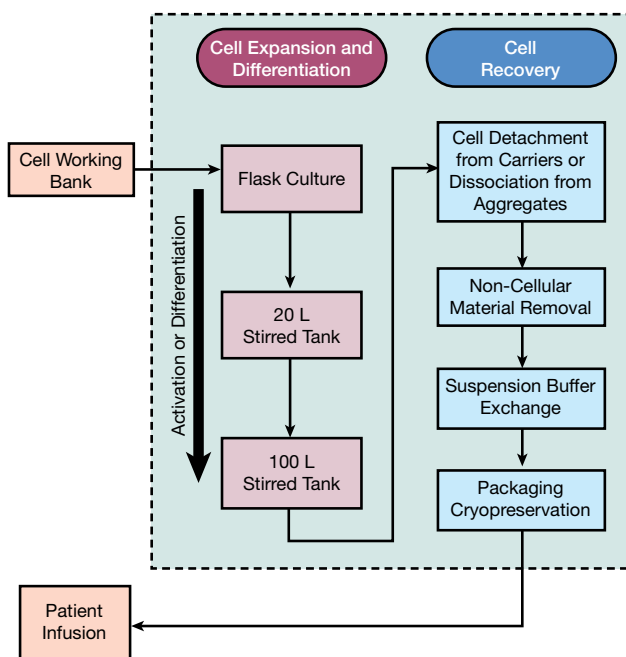
# Back to Basics

### (a) Workflow for conventional therapeutic protein production



### (b) Workflow for cellular therapy products



**◄ Figure 3.** Mammalian cell-culture processes include many similar steps. This is an example workflow for the production and recovery of (a) therapeutic proteins and (b) cellular therapy products. While similar bioreactor operations can be used to increase the cell population for both processes, the maximum scale for cellular therapy products is much lower than that of therapeutic proteins. Additionally, products for cellular therapy will often undergo differentiation and/or activation steps to become a specific cell type (such as skin cells) or to become activated. Though therapeutic protein products are recovered and purified through a series of unit operations (*i.e.,* centrifugation, membrane, and chromatographic processes), cell products can only be minimally processed to maintain acceptable viability.

This is not the case for therapeutic proteins. Often, these proteins are injected in large doses, up to grams of protein at a time. With a larger dose, the potential to carry over an intolerable amount of processing materials is higher. These materials may originate from medium components or be released by the cells. Thus, the biopharmaceutical industry has expended much effort to develop a completely chemically-defined (CD) medium, by not only removing serum (a regulatory requirement introduced two decades ago), but also by replacing any complex additives with structurally defined compounds. These products also undergo an extensive purification process to meet purity requirements.

Product purification is different for regenerative medicine applications. In this case, healthy and potent cells are the products. These cells cannot undergo many unit operation steps or prolonged treatment because such processes may reduce the cells' viability and potency. The key to achieving high purity is to prevent undesired contaminants from entering the process. In addition to eliminating traditional contaminants, process purification must remove any cells that do not possess the desired functionality or that are still undifferentiated and could possess tumorigenic potential. Manufacturing cells destined for therapeutic use is, in many respects, far more challenging than therapeutic protein production.

## Biomanufacturing processes

*Fixed reactor processes.* The processes for generating various cell culture products bear much resemblance to one another (Figure 3). Once a cell line for production is established, it is expanded and distributed into thousands of individual ampules that can be frozen and stored until needed for protein manufacturing. For each manufacturing run, an ampule from the cell bank is withdrawn from frozen storage and thawed to initiate a culture. As the culture grows, it is successively expanded into larger volume reactors until the production scale is reached. After the production phase is over, the contents of the production reactor are filtered to remove cells, and the broth is subjected to recovery steps for the isolation of purified product.

A typical manufacturing plant takes over two years to plan and construct, with an additional year required to

prepare and train workers to ensure smooth operation. In between each run, the entire facility is thoroughly cleaned, examined, and verified; the reactors and related equipment must be steam sterilized. The entire core manufacturing area must meet the quality criteria before another run can be initiated.

The capital investment for a manufacturing plant is high, often in the hundreds of millions of dollars, because of the high cost of permanent fixtures such as reactors and their auxiliary equipment (Figure 4). The cost of reconditioning the facility prior to each operation cycle is also very high. Steep costs have prompted the search for alternative approaches to cell culture manufacturing.

*Single-use, modular equipment processes.* In the past decade, a trend of employing disposable reactors and auxiliary equipment has taken hold. These single-use reactors and supplies are prefabricated and certified by their commercial supplier. They can be designed as modules, with separate single-use equipment for each unit operation ready to be assembled and put into operation. In some cases, the entire manufacturing line can be assembled from single-use equipment. Modularity offers a higher degree of operational flexibility, lower capital investment, and faster turnaround time, as well as a shortened lead-time for plant construction.

*Continuous processes.* Traditional cell culture processes are run in batch mode, sometimes with periodic additions of medium or nutrient mixtures during the process. For vaccine manufacturing, after reaching a high concentration, cells are infected with a virus to initiate virus propagation. This infection and subsequent virus replication often causes the cells to become sick and leads to cell lysis and decreased cell viability. The culture process is operated for a period of time to allow for virus accumulation before product recovery. For protein manufacturing, concentrated nutrients are added to the culture to prolong the amount of time that cells continue to secrete the product.

A typical manufacturing plant has multiple reactors. Each reactor can be sized between 10 $m^3$ and 30 $m^3$ and operated on a 10–15 day cycle. New single-use bioreactors are fabricated with reinforced polymers and coated with materials to prevent the leaching of plasticizer during manufacturing. The largest available single-use reactor that can sustain mechanical stability during stirred-tank operation is about 2 $m^3$. This volume is not large enough to meet manufacturing demands for biologics unless a large number of reactors is used. However, using many reactors will increase operating costs, and thus makes single-use reactors uneconomical.

A possible solution is to operate the smaller disposable reactors in a continuous mode at much higher cell concentrations. Continuous processing may also have the advantage of steady-state operation, which could generate product of more consistent quality. Efforts to develop continuous operations using single-use biomanufacturing equipment are ongoing, but will likely become the norm for cell-based therapy processes. The scale for these applications will likely be smaller and well within the range of disposable supplies. For more information on continuous processing in biopharmaceutical production, see the SBE Supplement: Biopharmaceuticals (*CEP*, Dec. 2015, pp. 29–46).

## The opportunities

In the past decade, the manufacturers of therapeutic proteins have adopted standardized biomanufacturing platforms. By standardizing the way cell lines, media, and processes are developed, all new products can be produced using the platform protocol with relatively small tweaks.

Advances in therapeutic protein manufacturing technology aim to enhance robustness in product quality. Over the past three decades, tens of millions of doses of protein therapeutics have been administered to patients with no deaths due to biomanufacturing process irregularity. This excellent product safety record is now being extended to the control of product quality attributes, such as the structure of glycan (sugars) on the protein. Glycan structure can improve drug efficacy so that less of the active ingredient can be used within each dose, and it can improve the half-life of the drugs so they can be administered less frequently.

For new protein biomanufacturing processes, robust cell lines, media, and processes are the norm. Medium components are extensively screened based on design of experiment (DOE) and quality by design (QbD) principles.

For protein biomanufacturing, opportunities exist that employ genome engineering tools to drastically alter cells



▲ **Figure 4.** A typical biomanufacturing plant can take over two years to design and construct, and must be thoroughly cleaned in between every product run. The manufacturing plant shown here produces therapeutic proteins in bioreactors that measure 25 $m^3$. Image courtesy of Bristol-Myers Squibb.

to best suit process needs. Because cells have evolved to serve as members of the organism from which they are derived, their genetic, biochemical, and structural machinery is not optimized for biomanufacturing. In fact, their metabolic behavior is unfavorable for optimal reactor performance. Genome engineering tools give engineers the ability to modify genes and genomes to tune their metabolic characteristics and synthetic capacity for robust manufacturing.

The manufacturing process for cell-based therapies is still in its infancy. The development of cell biomanufacturing processes can potentially benefit from the advances made in protein biomanufacturing.

Cell biomanufacturing processes face additional challenges because the cells themselves are the product, and they require complex nutrients, growth factors, and surface adhesion. For applications involving a patient's own cells (autologous cell therapy), each individual's cells are likely to behave differently. A rapid process development scheme that can be adapted for individual differences must be created. At the end of cell expansion and recovery, cells intended for cellular therapy must be in a state ready for rapid use; if they need to be cryopreserved, any recovery process must minimize traumatic shock to the cells. Ideally, the cells should also be in a mixture amenable for delivery to patients.

**TUNG S. LE** is a PhD candidate in chemical engineering at the Univ. of Minnesota. His research involves cell culture engineering for the production of biologics, and his area of focus is elucidation of cellular metabolism via transcriptome analysis and mathematical modeling. He has a BS in chemical engineering from Hanoi Univ. of Science and Technology, Vietnam.

**MEGHAN McCANN** is a graduate student in chemical engineering at the Univ. of Minnesota. Her research involves cell culture engineering, focusing on genetically engineering cells for the production of biologics. She has a BS in chemical engineering from North Carolina State Univ.

**SAMIRA M. AZARIN, PhD,** is an assistant professor in the Dept. of Chemical Engineering and Materials Science at the Univ. of Minnesota. Her research focuses on engineering the cell microenvironment for applications in regenerative medicine and cancer therapeutics. She has a BS in chemical engineering from the Massachusetts Institute of Technology and a PhD in chemical engineering from the Univ. of Wisconsin-Madison.

**WEI-SHOU HU, PhD,** is the Distinguished McKnight University Professor in the Dept. of Chemical Engineering and Materials Science at the Univ. of Minnesota (Email: acre@cems.umn.edu). His research involves cell culture engineering, including both stem cells and recombinant cells for the production of biologics, and his areas of focus include the elucidation of gene-trait relationships through transcriptome/genome analysis and the control of cell metabolism via model-assisted genetic and environmental intervention. He has a BS in agricultural chemistry from National Taiwan Univ., and SM and PhD degrees in biochemical engineering from the Massachusetts Institute of Technology. He is a member of AIChE and SBE, and currently serves on SBE's Advisory Board.

## Concluding remarks

Mammalian cell culture is central to biomanufacturing of therapeutic proteins and viral vaccines. Given current advances in cell-based therapy, the role of mammalian-cell bioprocessing in therapeutic applications will only increase in the next decade. The genome engineering toolkit will empower engineers and scientists to generate cells with more desirable biological traits, providing better control of product quality and enhanced productivity. This evolution requires an in-depth and systematic understanding of a cell's genome and intrinsic regulatory mechanisms. Chemical engineers are poised to help advance this trend by contributing their knowledge of process control and model-based optimization principles. **CEP**

## LITERATURE CITED

1. **Deegan, P. B., and T. M. Cox,** "Imiglucerase in the Treatment of Gaucher Disease: A History and Perspective," *Journal of Drug Design, Development and Therapy,* **6**, pp. 81–106 (2012).
2. **Goffe, B., and J. C. Cather,** "Etanercept: An Overview," *Journal of the American Academy of Dermatology,* **49** (2 Suppl), pp. S105–111 (Aug. 2003).
3. **Shay, J. W., and W. E. Wright,** "Hayflick, His Limit, and Cellular Ageing," *Nature Reviews Molecular Cell Biology,* **1** (1), pp. 72–76 (Oct. 2000).
4. **Hayflick, L.,** "The Limited *In Vitro* Lifetime of Human Diploid Cell Strains," *Experimental Cell Research,* **37**, pp. 614–636 (Mar. 1965).
5. **Todaro, G. J., and H. Green,** "Quantitative Studies of the Growth of Mouse Embryo Cells in Culture and Their Development into Established Lines," *The Journal of Cell Biology,* **17**, pp. 299–313 (May 1963).
6. **Wells, W. A.,** "A Cell Line That is Under Control," *The Journal of Cell Biology,* **168** (7), pp. 988–989 (Mar. 28, 2005).
7. **Thomas, E. D., et al.,** "Bone-Marrow Transplantation," *New England Journal of Medicine,* **292** (16), pp. 832–843 (1975).
8. **Evans, M. J., and M. H. Kaufman,** "Establishment in Culture of Pluripotential Cells from Mouse Embryos," *Nature,* **292** (5819), pp. 154–156 (July 9, 1981).
9. **Thomson, J. A., et al.,** "Embryonic Stem Cell Lines Derived from Human Blastocysts," *Science,* **282** (5391), pp. 1145–1147 (1998).
10. **Goldberg, A. D., et al.,** "Epigenetics: A Landscape Takes Shape," *Cell,* **128** (4), pp. 635–638 (Feb. 23, 2007).
11. **Dawson, M. A., and T. Kouzarides,** "Cancer Epigenetics: From Mechanism to Therapy," *Cell,* **150** (1), pp. 12–27 (July 6, 2012).
12. **Takahashi, K., and S. Yamanaka,** "Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors," *Cell,* **126** (4), pp. 663–676 (Aug. 25, 2006).
13. **Yamanaka, S.,** "Induced Pluripotent Stem Cells: Past, Present, and Future," *Cell Stem Cell,* **10** (6), pp. 678–684 (June 14, 2012).
14. **Pharmaceutical Research and Manufacturers of America,** "2013 Report: Medicines in Development — Biologics," www.phrma.org/sites/default/files/pdf/biologics2013.pdf, Washington, DC (Feb. 7, 2013).

# Lessons Learned from Recent Process Safety Incidents

**ALBERT NESS**
CENTER FOR CHEMICAL PROCESS SAFETY

These five incidents demonstrate how seemingly small mistakes or temporary lapses in judgment can result in disaster.

Failures of process safety management (PSM) systems are deadly and costly. Major accidents have emphasized the need for process safety within the chemical and petrochemical industries. For example, the founding of the Center for Chemical Process Safety (CCPS) was a response by industry to the methyl isocyanate release at Bhopal, India, in 1984 that killed over 2,000 people and injured tens of thousands. A fire and explosion at a PEMEX LPG terminal in Mexico City, also in 1984, killed more than 600 people and injured around 7,000.

Major environmental damage has also been caused by process safety incidents. The firefighting efforts during a fire in a Sandoz warehouse in Basel, Switzerland, in 1986 caused the release of many different chemicals, including pesticides, because responders failed to contain the water runoff. The release caused massive destruction to aquatic life in the Rhine River as far as 250 miles away; fishing was banned for six months. The environmental consequences of the Exxon Valdez spill in 1989 and the Deepwater Horizon incident in 2010 have been well-documented.

Engineers and students can obtain safety incident reports from several sources. One useful source is the U.S. Chemical Safety and Hazard Investigation Board (CSB). The CSB is a government agency charged with investigating chemical accidents at industrial facilities. The reports of their investigations are available for download from the CSB website (www.csb.gov). Additionally, the CSB has created a series of videos about many process safety incidents.

The CCPS book *Incidents that Define Process Safety (1)* discusses many more events. The book also describes events from industries other than chemicals and petrochemicals, demonstrating that many PSM concepts are universal in their relevance to safe operations.

This article describes five accidents that have occurred over the past 25 years that are less well-known than the Bhopal and Deepwater Horizon incidents. Major disasters — like the one that occurred in Bhopal — are sometimes referred to as black swans *(2)*. These types of incidents have a low likelihood of occurring again in our lifetime.

Accidents such as those discussed in this article are more likely to occur. These incidents demonstrate that even small mistakes can have disastrous consequences. Therefore, it is imperative that engineers learn from previous incidents to reduce their likelihood of recurring.

## Swiss cheese model

Accidents almost always have more than one cause. For many years, safety experts have used the Swiss cheese model *(3)* to help managers and workers in the process industries understand the events, failures, and decisions that can lead to a catastrophic incident or near miss. According to this model (Figure 1), each layer of protection is depicted as a slice of Swiss cheese, and the holes in the cheese represent potential failures in the protection layers, such as:

• human errors
• management decisions
• single-point equipment failures or malfunctions
• knowledge deficiencies
• management system inadequacies, such as a failure to perform hazard analyses, failure to recognize and manage

---

This article is based on "Chapter 3: The Need for Process Safety," of the *Student Handbook for Process Safety,* a Center for Chemical Process Safety (CCPS) book due to be published later in 2015. For more information on these and many more process safety incidents, please see that book.

changes, or inadequate follow-up on previously experienced incident warning signs.

Incidents are typically the result of multiple failures to address hazards effectively — represented by the holes in successive slices aligning. A management system may include physical safety devices or planned activities that protect and guard against failure. An effective PSM system has the effect of reducing the number of holes and the sizes of the holes in each of the system's layers, thereby reducing the likelihood that they will align.

## 1. ARCO Channelview explosion: Use MOC for wastewater tank maintenance

A wastewater tank at the ARCO chemical plant in Channelview, TX, exploded during the restart of a compressor on July 5, 1990. The nitrogen purge had been significantly reduced during maintenance, and a temporary oxygen analyzer failed to detect the buildup of a flammable atmosphere in the tank. When the compressor was restarted, flammable vapors were sucked into the compressor and ignited. The flashback of the flame into the headspace of the tank caused an explosion that killed 17 people. Damages were estimated to be $100 million *(4)*.

*Event details.* The 900,000-gal wastewater tank contained process wastewater from propylene oxide and styrene processes (Figure 2). There were thousands of feet of piping upstream of the tank where peroxides and caustic could mix. A nitrogen purge kept the vapor space inert, and an off-gas compressor drew the hydrocarbon vapors off before the waste was disposed of in a deep well.
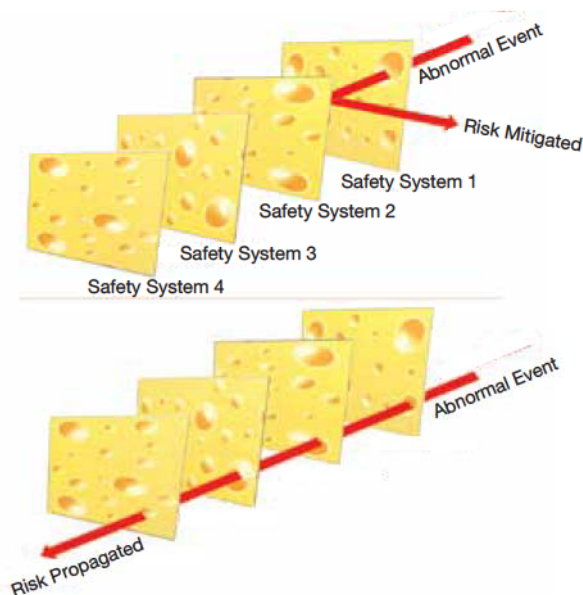


▲ **Figure 1.** The Swiss cheese model illustrates that when the weaknesses in layers of protection align, an accident can occur.

The tank was taken out of service so the nitrogen blanket compressor could be repaired. During this time, the normal flow of nitrogen purge gas to the tank was reduced to a minimum. Workers installed a temporary oxygen analyzer between two roof beams in the tank, and they planned to add a nitrogen purge if a high oxygen level was detected.

Within the tank, decomposition of the peroxides generated dangerous levels of oxygen. However, the air in the headspace of the tank was stagnant and the analyzer did not detect the oxygen buildup because it was in a dead zone. Occasional nitrogen purging was inadequate to prevent the formation of a flammable atmosphere in the headspace and in the piping to the compressor.

After maintenance was completed, a failed attempt to restart the compressor drew flammable vapors into the suction line of the compressor. When the compressor was successfully restarted, the flammable vapors were ignited and flame flashed back to the tank, causing an explosion.

*Causes.* The wastewater tank was not considered part of the operating plant. Hence, the management and workers did not understand that a chemical reaction was taking place in the tank, generating oxygen. The lack of understanding enabled a series of poor decisions, such as discontinuing the nitrogen purge, poor design and location of the temporary oxygen probe, no management of change (MOC) review of the these decisions, and no pre-startup safety review.

*Key lessons.* Keep in mind that the chemicals that enter any wastewater tank are still prone to reaction. Ensure that proper MOC procedures are followed before any maintenance work is performed. In this incident, the workers did not know that a chemical reaction that could produce an oxygen buildup was taking place in the tank. Therefore, they did not comprehend the importance of continuing an effective nitrogen purge.

## 2. Terra Industries ammonium nitrate explosion: Monitor equipment during shutdown

On Dec. 13, 1994, a massive explosion occurred in the ammonium nitrate (AN) portion of Terra Industries' fertilizer plant in Port Neal, IA (Figure 3). The explosion occurred after the process had been shut down and ammonium nitrate solution was left in several vessels. Multiple factors contributed to the explosion, including strongly acidic conditions in the neutralizer, application of 200-psig steam to the neutralizer vessel, and lack of monitoring of the plant when the process was shut down with materials in the process vessels. Four people were killed and 18 injured. Serious damage to other parts of the plant caused the release of nitric acid into the ground and anhydrous ammonia into the air *(5)*.

*Event details.* The plant produced nitric acid, ammonia, ammonium nitrate, urea, and urea-ammonium nitrate. Ammonia from the urea plant off-gas or from ammonia

storage tanks was added to the neutralizer through a sparger in the bottom of the vessel, and 55% nitric acid was added through a sparging ring in the middle of the vessel. The product, 83% AN, was sent to a rundown tank via an over-flow line for transfer to storage. A pH probe located in the overflow line controlled the nitric acid flow to the neutralizer to maintain the pH at 5.5–6.5. The temperature in the neu-tralizer was maintained at about 267°F. Both the neutralizer and rundown tank were vented to a scrubber, where the vapors were absorbed by 55–65% nitric acid and makeup water to produce 50% ammonium nitrate. A stream of 50% AN was recycled back to the neutralizer.

About two weeks prior to the event, the pH probe in the overflow line was found to be defective, at which time the plant switched to manual pH sampling. Two days prior to the event, the pH was measured as 1.5 and was not brought into the acceptable range until about 1:00 am on Dec. 12.

The AN plant was shut down at about 3:00 pm on Dec. 12 because the nitric acid plant was out of service. At about 3:30 pm, operators purged the nitric acid feed line to the neutralizer with air. At about 7:00 pm, operators pumped the scrubber solution to the neutralizer. Then, 200-psig steam (which is around 387°F) was applied through the nitric acid feed line to the nitric acid sparger to prevent backflow of AN into the nitric acid line. The explosion occurred at about 6:00 am on Dec. 13.

AN is known to become more sensitive to decomposi-tion, deflagration, and detonation at low pH levels, at high temperatures, in low-density areas (e.g., in areas contain-ing gas bubbles), in confined spaces, and in the presence of contaminants, such as chlorides. Calculations showed that the nitric acid line clearing would have lowered the pH at the time of the shutdown to about 0.8. The steam sparge was left on for 9 hr, providing enough heat to raise the solution to its boiling point in about 2 hr. The air and steam sparge created gas bubbles in the solution. Chlorides, carried over from the nitric acid plant, were also present in the AN solution.
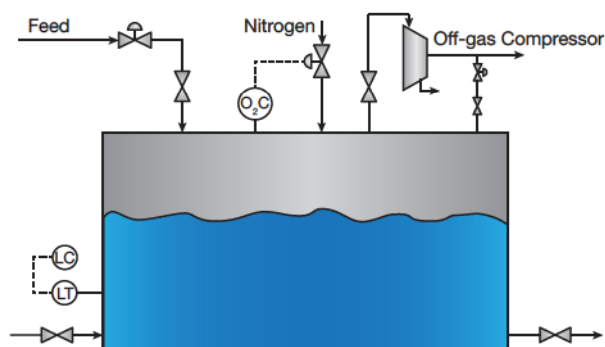
*Causes.* The U.S. Environmental Protection Agency (EPA) investigation concluded that the conditions that led to the explosion occurred due to the lack of safe operating procedures. There were no procedures for putting the vessels into a safe state at shutdown, or for monitoring the process vessels during shutdown. The EPA found that other produc-ers either emptied the process vessels during a shutdown or maintained the pH above 6.0. Also, other producers either did not allow steam sparges or, if steam sparges were used, they were conducted under direct supervision of operators.
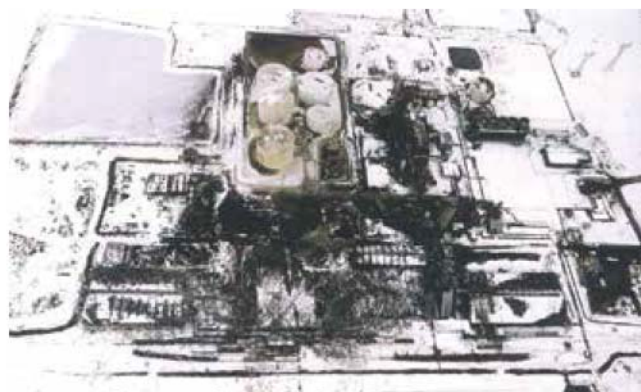
The EPA also noted that no hazard analysis had been done on the AN plant, and that personnel interviewed "indicated they were not aware of many of the hazards of ammonium nitrate" (5).

*Key lessons.* Operating procedures need to cover all phases of operation. In this event, the lack of procedures for shutdown and monitoring the equipment during shutdown led operators to perform actions that sensitized the AN solution and provided energy to initiate the decomposition reaction.

Because there had been no hazard identification study, personnel did not know about the conditions that sensitize AN to decomposition. A hazard assessment of the shutdown step would have revealed that the pH of the neutralizer could not be measured if there was no solution flowing through the overflow line, and that the temperature of the neutralizer could not be accurately measured without any circulation in the tank. A complete hazard identification study would have covered backflow of ammonium nitrate into the nitric acid line, and better design solutions could have been identified.

### 3. Partridge-Raleigh oilfield explosion: Beware of hot work and flammable gases

On June 5, 2006, three contract workers were killed and a fourth worker was seriously injured in an explosion and fire at the Partridge-Raleigh oilfield in Mississippi. The con-tractors, who were employees of Stringer Oilfield Services, were tasked with installing a pipe between two oil produc-

▲ **Figure 2.** A 900,000-gal wastewater tank had a nitrogen purge to keep the vapor space inert and an off-gas compressor to draw off hydrocarbon vapors before waste disposal.

▲ **Figure 3.** A massive explosion devastated this ammonium nitrate fertilizer plant. Source: (5).

tion tanks (Figure 4). Welding sparks ignited flammable vapor that was escaping from an open-ended pipe near the welding activity *(6)*.

*Event details.* Contract workers were connecting piping between two recently moved tanks (Tanks 3 and 4 in Figure 4). Several days earlier, crude oil residue was removed from Tank 4 and the tank was flushed with water. However, the contractors did not clean out or purge the crude oil residue from Tank 2 or Tank 3.

Before starting to weld, the welder checked for flammable vapors in Tank 4 by inserting a lit welding torch into it, an unsafe act known as flashing the tank. Then, as the CSB report *(6)* states, "The foreman climbed to the top of Tank 4. Two other maintenance workers climbed on top of Tank 3; they then laid a ladder on the tank roof, extending it across the 4-ft space between Tanks 3 and 4, and held the ladder steady for the welder. The welder attached his safety harness to the top of Tank 4 and positioned himself on the ladder *(6)*." Figure 5 illustrates the workers' locations.

Almost immediately after the welder started welding, flammable hydrocarbon vapor that was venting from the open-ended pipe attached to Tank 3 ignited. The fire flashed back into Tank 3, spread through the overflow connecting pipe from Tank 3 to Tank 2, and caused Tank 2 to explode. The lids of both tanks were blown off and the two maintenance workers and foreman were thrown off the tanks to the ground. The welder was thrown off the ladder, but his harness prevented him from falling to the ground.

*Causes.* The root cause of this incident was hot work being conducted in the presence of a flammable atmosphere without using any safe work permitting procedure. A gas detector should have been used to test for flammable vapor. The open pipe on Tank 3 was not capped or isolated. All of the tanks were interconnected, and some of the tanks still contained flammable residue and crude oil.

*Key lessons.* Safe work practices, such as hot work permits, are necessary to ensure a safe work environment
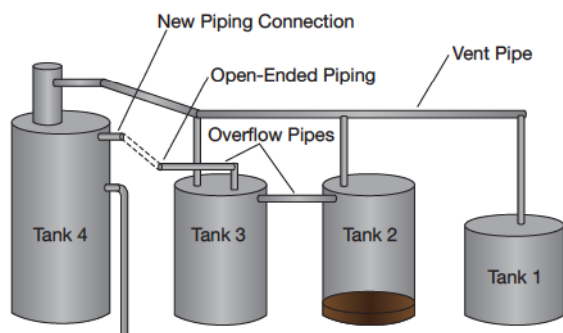
when hazardous chemicals, in this case flammable vapors, are present. The contractor, Stringer's Oilfield Services, did not require the use of safe work procedures, specifically hot work permits in this case.

Contractors need to be managed in such a way as to ensure they know about and use safe work practices. The owner of the wells and tanks, Partridge-Raleigh, relied on contractors to do most of its well commissioning work, such as installing tanks, pumps, and piping — this is a common practice. Partridge-Raleigh did not, however, manage the contractors to make sure they used safe work practices.

Companies need to be aware of and follow best industry practices. Several National Fire Protection Association (NFPA) and American Petroleum Institute (API) guidelines cover this situation. If Partridge-Raleigh or Stringer's Oilfield Services had adopted any of these industry standards, this incident could have been prevented:
• NFPA 326, "Standard for the Safeguarding of Tanks and Containers for Entry, Cleaning, or Repair" (2005)
• NFPA 51B, "Standard for Fire Prevention During Welding, Cutting, and Other Hot Work" (2003)
• API Recommended Practice 2009, "Safe Welding, Cutting and Hot Work Practices in the Petroleum and Petrochemical Industries" (2002)
• API 74, "Recommended Practice for Occupational Safety for Onshore Oil and Gas Production Operations" (2001).

## 4. Formosa Plastics vinyl chloride release: Follow correct operating procedures and protocols

On April 23, 2004, an explosion and fire at the Formosa Plastics Corp. plant in Illiopolis, IL, killed five workers and seriously injured two others. The event destroyed most of the polyvinyl chloride (PVC) manufacturing facility and ignited PVC resins stored in an adjacent warehouse *(7)*. Concerns



▲ **Figure 5.** The workers killed and injured in the Partridge-Raleigh oilfield tank explosion were on top of two tanks with a ladder stretching between them. The foreman (F) stood on Tank 4, while the two maintenance workers (M) stood on Tank 3, and the welder (W) positioned himself on the ladder. Source: *(6)*.



▲ **Figure 4.** Contractors were tasked with installing a pipe between two tanks. Vapors escaping from Tank 3 ignited and caused an explosion. Source: Adapted from *(6)*.

about the ensuing smoke from the fire forced a two-day community evacuation.

Vinyl chloride monomer (VCM) — a highly flammable chemical and known carcinogen — is the primary raw material in the PVC manufacturing process. The Formosa Plastics facility used VCM to manufacture PVC resins. VCM served as the fuel for the initial explosion and fire.

*Event details.* The facility produced PVC by heating VCM, water, suspending agents, and reactor initiators under pressure in a batch reactor. There were 24 reactors in a building, and the reactors were put in groups of 4, with a control station for every two reactors (Figure 6). When a reaction was complete, the PVC solution was transferred through the bottom valve to a vessel for the next step in the process.

After the transfer, the reactor was purged of hazardous gases and cleaned by power washing through an open manway. The wash water was emptied to a drain through the reactor's bottom valve and a drain valve. All of these steps were done manually.

On the day of the incident, the reaction and the power washing had been completed in reactor D306 and the operator went downstairs to drain the reactor. It is believed that, at the bottom of the stairway, he turned in the wrong direction, toward an identical set of four reactors that were in the reaction phase of the process (Figure 7). By mistake, the operator likely attempted to empty reactor D310 by opening the bottom and drain valves. The bottom valve, however, was interlocked to remain closed when the reactor pressure was above 10 psi. Because this tank was currently processing a batch of PVC at high pressure, the valve did not open.

In case of an emergency (such as reactor overpressure), operators could follow an emergency transfer procedure that required them to open the bottom valve and the transfer valve to connect the reactor to an empty reactor. However, during an emergency transfer, the reactor pressure is greater than 10 psi, and the safety interlock would prevent the

opening of the bottom valve. Therefore, the company added a manual interlock bypass so that operators could open the valve and reduce reactor pressure in an emergency. The bypass incorporated quick-connect fittings on air hoses so that operators could disconnect the valve actuator from its controller and open the valve by connecting an emergency air hose directly to the actuator.

It is likely that the operator thought he was at the correct reactor (D306) and that its bottom valve was not functioning. When the bottom valve did not open, he switched to the backup air supply and overrode the interlock. He did not contact the upstairs reactor operator or shift foreman to check the status of the reactor before doing this.

Once the bottom valve was opened, VCM poured out of the reactor and the building rapidly filled with liquid and vapor. A deluge system in the building activated and a shift supervisor came to the area to investigate. The VCM detectors in the building were reading above their maximum measurable levels. The shift foreman and reactor operators took measures to slow the release, rather than evacuate. The VCM vapors found an ignition source and several explosions occurred. The ensuing fire spread to the PVC warehouse and burned for hours, sending a plume of acrid smoke into a nearby community.

*Causes.* The operator overrode an interlock, which led to a release of hot, pressurized VCM. Formosa Plastics did not have comprehensive written standards, such as requiring shift supervisor approval, for managing interlocks on the vessels. Employees were unprepared for a major accident at the facility.

Several factors made this incident more likely to occur:

• The reactor groupings had similar layouts (Figure 7).

• The operators on the lower levels were not given radios, which would have made communication with the reactor control operators on the upper level easier. (Similar Formosa plants had radios or an intercom system.)

• Formosa eliminated an operator group leader position and gave its responsibilities to the shift supervisors, who
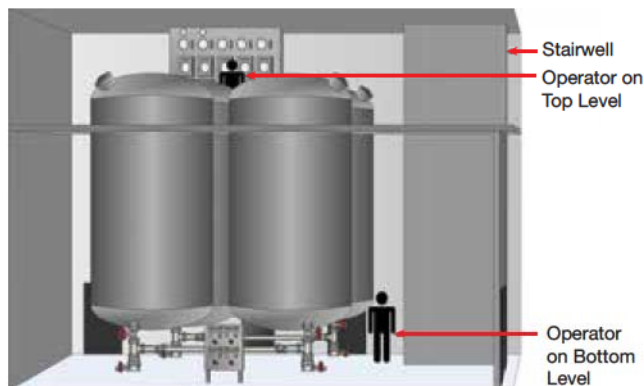


▲ **Figure 6.** Reactors at the PVC plant were arranged in groups of four. One operator monitored the reactors from the top level, while another operator, on the lower level, drained the tanks. Source: *(7)*.



▲ **Figure 7.** Reactors D306 and D310 occupied the same place in their respective, identical tank rooms. At the bottom of the stairway, the operator became confused and went into the wrong room, where he emptied the wrong reactor. Source: *(7)*.

were not always as available as the group leaders used to be.

*Key lessons.* Operators and engineers must follow operating procedures and protocols intelligently, and, when the process moves outside the operating envelope, stop work, get experienced advice as needed, and shut down as appropriate. The Formosa operator should have obtained supervisory approval to override the interlock.

Furthermore, in this event, the operators had to cope with an error-prone design — the reactor layout made it easier for a mix-up to occur. An emergency transfer procedure required bypassing the bottom valve interlock, so an easy means was provided to do this. Engineers who design and run plants should try to provide engineering controls and monitor shift notes and logs for instances of interlock bypassing. In this case, a reactor status indication on the operating floor could have been provided, and more-rigorous enforcement of operating procedures and interlock management implemented.

Operators were not given tools (radios for communication between floors) to make it easier for them to follow their procedures. It is management's responsibility to provide the tools and controls necessary for operators to do their jobs safely.

When Formosa Plastics took over the plant, it made staffing changes, such as reductions in staff and changes in responsibilities. It did not conduct a formal management of organizational change review to analyze the impact of these changes.

This explosion also illustrates the importance of emergency response planning. When the VCM release occurred, gas detectors in the building and a deluge system were activated. Operators responded by trying to mitigate the release. The proper response to these activations would have been to evacuate.

## 5. Hoeganaes combustible dust flash fires: Make housekeeping a priority

In 2011, a series of iron dust flash fires and a hydrogen explosion occurred at the Hoeganaes facility located in Gallatin, TN. The plant specialized in melting and converting scrap metal to various metal powders. These three incidents killed a total five people and injured three others.

The Hoeganaes facility's main product is a powder that is 99% iron. The process involves melting the iron, then cooling and milling it into a coarse powder. The powder is sent through an annealing furnace on a 100-ft-long conveyor belt. The furnace has a hydrogen atmosphere to reduce oxides and prevent oxidation. Hydrogen is supplied through pipes located in a trench in the floor, which is covered by metal plates. The product from the furnace, called a cake, is sent to a cake breaker and then crushed into a powder with a particle size of 45–150 µm.

*First incident.* On Jan. 31, 2011, operators thought that a bucket elevator used to transfer the powder was off track (*i.e.,* the belt had become misaligned, which can cause the motor to overheat due to the increased torque). After shutting down the motor, a maintenance mechanic and an electrician inspected the equipment. They did not believe the belt was off track and requested the operator to restart the motor. When the motor started, the vibrations dispersed powder that was on the equipment and floor (Figure 8). A flash fire occurred almost immediately and engulfed the two workers, killing both.

*Second incident.* On March 29, 2011, a Hoeganaes engineer and a contractor were replacing igniters on an annealing furnace. They had difficulty reconnecting a gas line, and the engineer used a hammer to force the connection. Large amounts of dust on surrounding surfaces were dispersed by the hammering and ignited almost immediately. The engineer suffered first- and second-degree burns, while the contractor was able to escape. The engineer was wearing flame-resistant clothing (FRC), which may have helped prevent more serious burns. Figure 9 is a photo taken at the Hoeganaes plant on Feb. 3, 2011, about two months before this incident *(8)*. This photo shows how much dust had piled up on the plant's surfaces.

*Third incident.* On May 27, 2011, operators near an annealing furnace identified a gas leak coming from a trench that contained hydrogen, nitrogen, and cooling water runoff pipes, in addition to a vent pipe for the furnaces. Mechanics were dispatched to find and repair the leak. One area operator stood by as the mechanics searched for the source of the leak. Although maintenance personnel knew that hydrogen piping was in the same trench, they presumed that the leak was nonflammable nitrogen because of a recent leak in a nitrogen pipe elsewhere in the plant. However, in this case the source of the leak was a line containing hydrogen.

The trench covers were too difficult to lift without
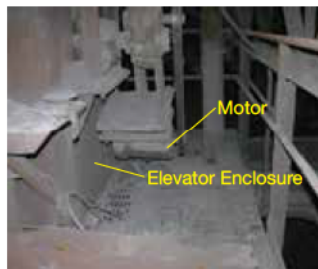
◀ **Figure 8.** Two workers went to inspect a bucket-elevator motor, which happened to be surrounded by piles of iron dust. During motor restart, vibrations dispersed powder on the floor and equipment. The dust ignited and created a flash fire, killing both workers. Source: *(8)*.

▶ **Figure 9.** Mounds of iron dust lurked on elevated surfaces at a scrap-metal processing plant. Source: *(8)*.

machinery, so a forklift was used to lift a cover near the leak. As the cover was pulled up by the forklift, friction created sparks and an explosion ensued. The hydrogen explosion dispersed large quantities of iron dust from rafters and other surfaces in the upper reaches of the building (Figure 9). Portions of this dust ignited, creating multiple dust flash fires in the area. Three employees died from the burns they suffered in the fire.

*Key lessons.* Understanding hazards and risks is one of the pillars of risk-based PSM *(9)*. After the incidents, combustibility tests indicated that the iron dust was a weak explosion hazard and relatively hard to ignite. These findings were similar to results Hoeganaes obtained after an insurance audit in 2008. A lesson here is that even a weakly explosive and hard-to-ignite dust is still combustible, and therefore, still hazardous and capable of causing fatalities when ignited. In this case, even though the company had the necessary information, personnel did not fully understand the hazards and risks of combustible dusts.

Learning from experience is another pillar of risk-based PSM *(9)*. The plant experienced an incident in 1992 that was very similar to the third incident in 2011. A hydrogen explosion in a furnace dispersed accumulated dust and created a flash fire that severely burned an employee (burns covered more than 90% of his body, and he spent a year in a burn unit). Hoeganaes did not learn from its own incident.

The importance of housekeeping in a facility that handles solids cannot be overstated. All three of these incidents were exacerbated by the large quantities of combustible dust present (Figures 8 and 9). Poor housekeeping has been involved in most, if not all, high-consequence dust explosions *(10)*. At the Hoeganaes plant, control of dust emissions and housekeeping were ineffective. Baghouse filtration systems that were installed to control dust were frequently out of service, and the CSB investigators observed that the baghouses leaked when the bags were pulsed. The 2008 insurance audit also noted that housekeeping needed to be improved in several areas. The ineffective dust control and housekeeping enabled dust layers with more than enough dust to fuel the flash fires to accumulate. These deficiencies were contributing factors to all three incidents.

## Closing thoughts

These five lesser-known incidents demonstrate the importance of good PSM. Many engineers have learned these lessons the hard way, but their mistakes can help you to avoid similar situations in the future. Trevor Kletz, a world-renowned expert in process safety, is often quoted as saying, "Organizations don't have memory — only people do" *(11)*. By providing these examples, this article is helping you to collect and recall the necessary memories to prevent future accidents.

Most processes are designed with more than one layer of protection. However, no protection or safeguard is 100% perfect, and, like slices of Swiss cheese, there are holes in every layer. Incidents occur when multiple failures — or holes — line up. The goal of PSM is to make the holes as small and as few as possible.

As many of these incidents show, technical competence is not enough to prevent an accident — management systems and company culture also play a key role in process safety. **CEP**

**ALBERT NESS** joined CCPS in 2013 as a process safety writer, and was elected a CCPS Fellow in 2014. Prior to that, he had 39 years of experience with Rohm and Haas, GE Plastics, the Dow Chemical Co., and ABS Consulting as a research and development engineer in plastics, agricultural chemicals, and ion exchange resins, and then, starting in 1989, as a process safety specialist. He served as secretary/treasurer and director of the Safety and Health Div. of AIChE, and a member of the NFPA 654 technical committee for combustible dusts as a representative from Rohm and Haas and then Dow Chemical. Ness has a BS from the Univ. of Arizona and an MS from the Univ. of Illinois, both in chemical engineering.

## LITERATURE CITED

1. Atherton, J., and G. Frederic, "Incidents that Define Process Safety," American Institute of Chemical Engineers, New York, NY, and Wiley, Hoboken, NJ (2008).

2. Nara, L., "The Importance of Considering the Improbable (a.k.a. Black Swan Events," *Chemical Engineering Progress,* **109** (2), p. 20 (Feb. 2013).

3. Reason, J., "The Contribution of Latent Human Failures to the Breakdown of Complex Systems," *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences,* **327** (1241), pp. 475–484 (Apr. 12, 1990).

4. ARCO Chemical Co., "A Briefing on the ARCO Chemical Channelview Plant July 5, 1990 Incident" (Jan. 1991).

5. U.S. Environmental Protection Agency, "Chemical Accident Investigation Report: Terra Industries Inc. Nitrogen Fertilizer Facility, Port Neal, Iowa," EPA (Sept. 1996).

6. U.S. Chemical Safety and Hazard Investigation Board, "Hot Work Control and Safe Work Practices at Oil and Gas Production Wells: Partridge-Raleigh Smith County Oilfield," Report No. 2006-07-I-MS, Raleigh, MS (June 5, 2006).

7. U.S. Chemical Safety and Hazard Investigation Board, "Investigation Report: Vinyl Chloride Monomer Explosion," Report No. 2004-10-I-IL, Formosa Plastics Corp., Illiopolis, IL (Mar. 2007).

8. U.S. Chemical Safety and Hazard Investigation Board, "Hoeganaes Corporation: Metal Dust Flash Fires and Hydrogen Explosion," Report No. 2011-4-1-TN, Gallatin, TN (Dec. 2011).

9. Center for Chemical Process Safety, "Guidelines for Risk Based Process Safety," American Institute of Chemical Engineers, New York, NY, and Wiley, Hoboken, NJ (2007).

10. Frank, W. L., "Dust Explosion Prevention and the Critical Importance of Housekeeping," *Process Safety Progress,* **24** (3), pp. 175–184 (Sept. 2004).

11. Kletz, T., "Lessons from Disaster: How Organizations Have No Memory and Accidents Recur," Institution of Chemical Engineers, Warwickshire, U.K., pp. 1–3 (1993).

**Don't forget these excellent AIChE resources for First Year Professionals:**

Online Leadership Training

YP  Article Bundles and Webinars

The Young Professionals Committee

The ChEnected Blog

Conferences

**AIChE**
*young* **PROFESSIONALS**

**Visit www.aiche.org/yp for more information.**