Process Control Challenge Problems

R. Russell Rhinehart – 2018-11-15
russ@r3eda.com, rrr@okstate.edu

I. Why have a Process Control Challenge Problem?

Investigators use simulators to test control algorithms. Simulations are also useful to tune controllers off-line; to test methods for steady state detection, data reconciliation, dynamic model generation (such as FOPDT or FIR), steady state model generation, process and control strategy design; and many more tasks related to automation technique application, evaluation, and development. Challenge problems are also useful as education and training exercises; and desirably, a challenge problem is simple (to understand and to implement).

Yet, it should also be representative of typical challenges that applications provide. Each industry (or academic discipline alignment) has unique applications, with an associated range of challenges. Compared to electronic, mechatronic, and aerospace applications; chemical processes have slow dynamics. This would seem to ease the control challenge within the chemical process industries (CPI); but their applications are nonlinear, have variable deadtime, noise, unmeasured disturbances, constraints, interaction, measurement error, and are nonstationary (their gains, time-constants, noise amplitude change in time). Further, the economics of the CPI applications do not support either expensive control systems or academically credentialed operation personnel; so controllers must be relatively simple to compute and manage.

Challenge problems for process control should address such relevant aspects. Although the inverted pendulum is certainly a control challenge, and is relevant to many applications, it has little relevance to CPI applications.

The 1970s brought a significant interest in industrial process control as industry switched from analog single loop controllers to digital control in DCS (distributed control systems) or PLC (programmable logic control), and as vendors introduced model predictive control (MPC) and other computer-era products for monitoring and automation. (MPC is presently termed Advanced Process Control in industry, and has the acronym APC.) At that time, significant challenge problems were offered by industry to help guide academic research and industrial product development. Some of these have been termed the Eastman or Shell challenge problems, and the Wood-Berry column. There is a nice compilation of 13 challenge problems by the IFAC Theory Committee led by Davison and published in 1990 [1]. However, many of the challenge problems are related to mechatronic devices, not industrial processes; and, those that are related to the CPI are either fairly complicated, or linearized, or deterministic, or stationary. At the time these were created, the elementary problems were appropriate, and provided a welcomed direction for investigators. Since then, techniques have developed, and what used to be the challenge is not so much today. Today, new problems that include today's challenges for control of chemical processes are appropriate to guide research.

Most CPI challenge problems represents continuous flowing process. In contrast, a key characteristic of the chemical process industry are batch and fed-batch processes, which have unique control-relevant issues. Batch processes such as reactions or bio-production need continuous control of temperature and reagent additions; but, as the batch material transforms, the process changes in reaction volume, heat transfer characteristics, limits on mixing shear, change in reaction rates, nonlinearity of pH control, etc. These issues make batch control difficult; but in addition, batch-to-batch variation in feed often requires supervisory changes to the recipe to optimize outcomes for each batch.

Issues associated with auxiliary variables are often overlooked by investigations. Familiar auxiliary variable examples in automobile operation include engine oil pressure, coolant temperature, carburetor vacuum, gas tank level, road conditions, and other traffic. While the primary controlled variable is vehicle speed, a driver is usually monitoring the other variables, and will override the speed set point when an auxiliary variable is out of range. By contrast, investigators often pretend that distillation control is a 2x2 application with reflux and boil-up rates changed to control top and bottom product compositions. But, levels in the reboiler and condensate receiver, column pressure, column differential pressure, non-condensable gases, and boiler tube temperature must also be controlled. Often the auxiliary variables are as important as the primary variables for control, and an investigator's controller success on a trivialization of the application cannot be taken as a message of triumph to the community.

I am not against simple test cases. For classroom instruction about the basic concepts to learners, simple, 2x2, linear FOPDT models of a process are appropriate. The Wood-Berry distillation column model is a case in point. Simple linear system models are also appropriate for initial investigation of algorithms or for data confirmation of theoretical analysis. When one is revealing how to implement a controller, a simple application case study is also appropriate. I frequently use simple process models. Often, there is no need to use nonlinear, stochastic, extensive-time simulations.

However, often there is. For instance, when one is seeking to demonstrate for the world that they have devised a better controller, one should back up the claim with data from a credible simulation, not a trivial application to a linear, stationary, deterministic, low-order process that is perfectly known to the controller.

In my review of the scientific journal literature, I usually find that researchers use simple SISO or 2x2, linear, stationary, noiseless models for their studies; and representing the CPI, they seem to favor distillation or pH control applications. They find the model in the prior literature, which becomes their justification to use it, which promulgates justification for subsequent graduate students. But, such triviality in a testing basis, cannot support definitive claims to the real world announcing the solution for which we all have been awaiting. Something that is successful in demonstrating academic criteria for earning a PhD (mathematical prowess, integrating complexity, implementing the latest technique fashion) does not make it relevant to the practice.

My aim in this editorial is threefold: 1) To offer a guide to researchers regarding including reality in their process control investigations, 2) to offer challenge problems that I've developed, and 3) to solicit challenge problems from others to compile a new set for today's research. I welcome feedback to russ@r3eda.com, and you to visit www.r3eda.com to see several challenge problems.

There are nomenclature variations in the literature. In the following, MV means manipulated variable, the output of the controller, which is often termed controller output or control variable. Here, the acronym CV means controlled variable, which is the process variable (PV) that is to be kept at a desired set point. There are many other PVs that are monitored, and AV, means auxiliary variable, one that is monitored and action taken only when it is out of limits. Finally, here "system" means the controlled process. The controller and the process are separate entities within the system.

II. Control Challenge Attributes of Chemical Processes

Process simulators for chemical process control exploration should include attributes characteristic of CPI applications. Here is a list of CPI application attributes:

1. Nonlinear – Include enough of a gain change so that a linear controller tuned for one region does not work acceptably in another. Time-constants and delays also change with variations in process throughputs, making a tuned controller at one production rate not work well in another. Simulation of these process changes do not have to be excessive, just effective in troubling a linear controller. In the normal range of operation, process gain often has a 3:1 range, and time-constants and delay can be scaled with the inverse of process throughput rate. Without process experience, researchers often only see the one nominal operating state. I think it helps to relate personal experience, and I ask a reader to consider the range of speeds that they drive a car (from parking lots, to alleys, to mountain roads, to super highways) and the implications for steering and speed control.
2. Nonstationary – Properties or characteristics of the process or feed materials change in time. These could be due to heat exchanger fouling, loss of catalyst reactivity, or product purity. However, as well, these may be the result of operational change of flow paths, or switching of pumps, reactors, etc. Again, the extent of change of simulator coefficients or model should be enough to make a controller not work well under the new conditions. Easily, such factors for feed composition, reactivity, and pressure losses can change on a 2:1 range.
3. Disturbances – Include unmeasured external disturbances that change in time. For example, an elementary view of distillation is that reflux and boil-up rates are the MVs used to control top and bottom composition. But, feed flow rate, composition and temperature and column heat losses and reflux temperature are all continually changing unmeasured disturbances. These could be human-contrived events, if for instance maintenance is underway or a friend wants to wake you up by playing with a valve. I don't condone mistakes or practical jokes; but, they don't not happen just because they shouldn't.
4. Extra MVs – Beyond square MIMO studies, include extra MVs. Again, an elementary view of distillation is that reflux and boil-up rates are the MVs. But, column pressure can also be adjusted to control composition. It is an extra MV. There are desired resting values for the extra MVs, and control will switch to use the extra MVs when a primary MV is out of range.
5. Auxiliary Variables – For safety, for example, an AV may be monitored, and if that PV value gets out of range, an auxiliary controller may take over the MV from the primary controller. A classic example is that of a fired heater. As inflow rate increases, or inflow temperature decreases, or fouling increases, the firing rate must increase to keep the outflowing process fluid at set point. But, excessive firing can overheat the tubes; and to prevent failure, an auxiliary tube temperature controller will override the primary controller to reduce firing. As another example, in distillation differential pressure across the column is monitored as an indication of weeping or flooding; and when this AV is out of range, normal control is overridden. How does either the normal or override controller operate when not selected? How is the transfer managed to prevent a bump in a transfer of control? External reset feedback is a common approach with override switches in PID control, preventing integral windup and bumpless transition.
6. Interactive – Usually, each MV affects each CV, or at least some MVs affect several CVs. For instance, one might think that in mixing of hot and cold water (to achieve a desired total flow rate and mixed fluid temperature) the signal to the cold valve does not affect the hot water flow rate.

But it does. If the cold water flow rate changes, then the mixed flow rate changes, which affects the pressure drop on the mixed flow line, and consequently the hot flow rate. Unless, cascade structures are used, the manipulated variables are not the process conditions; they are signals to the valves that influence the process.

7. Constraints – These include limits on both value and rate-of-change on the MV, and limits on the CV and AVs. Action now can lead to a constraint violation in the future, so control should include the future impact of current actions on both CVs and AVs. Action on one MV (boil up rate) can change the constrained value on another MV (feasible reflux range). Often control is simply a response to current and past conditions, such as the logic in a proportional-integral controller. At least the D-action in PID has some anticipatory aspect; but, it does not forecast the impact on the AVs.
8. Noise – This can be considered as independent sampling-to-sampling perturbations on the measurement, such as from flow turbulence, imperfect mixing, mechanical vibration, or electronic interference. The distribution could be modeled as Gaussian, but often nonlinearities skew it. Many features create persistence of perturbations (mixing, thermal inventory, filtering), so the noise model might have autocorrelation. Include noise. If there is a derivative-equivalent aspect to your controller, the MV will be a joy to watch!
9. Spurious Signals – Occasional extreme values of either a process influence or measurement may happen because of a pulse (mechanical or electronic) or may appear to be because of a missed data transmission. These could be a one-sample event, or have a brief duration. A median filter is commonly used to reject spurious signal values, but this injects a delay in the measurement.
10. Sensor Error – This could be from calibration error, or calibration drift (due to age, dust, humidity, temperature, etc.). Alternately, it could be due to actual failure of the device (sample line plugging, broken glass membrane on a pH probe, blown fuse, etc.) Finally, a brief period of an erroneous signal may be the result of process maintenance. Include sensor error, do not pretend that your controller can know the true process variable value.
11. Delay – Often the value of the delay is variable and dependent on flow rates or analyzer duty. Any simulation should not pretend that the delay has a constant value.
12. Inverse Action – Here, the natural process response is initially in the wrong direction. The classic is the boiler swell effect, but this is also common in fluidized bed reactors. Arising from Laplace or Fourier transformed variable analysis, this is often termed non-minimum phase behavior. A Laplace transform might provide an efficient communication tool, but don't be thinking that doing so makes the process linear or low order. My experience has been with nonlinear integrating inverse responses.
13. Integrating Processes – Tanks and other inventory containers vessels have a fill or empty stage, they don't magically start at the level set point with inflow and outflow at steady state. As well, any inventory accumulation could be described as integrating, such as internal electronic heat generation in an insulated device, the odor accumulation in a sealed room, impurity buildup in a process with recycle, or the attrition loss of catalyst particles in a fluid bed. When the inventory threshold hits a limit, this triggers an event (like turning on the computer fan, or opening a purge vent) which then changes the inventory. A simulation should include such on-off cycling for AV control, because it affects the nominal control of the process. For instance, level in a tank affects discharge pressure which can change the gain of a valve or cavitation override on a pump.
14. Open Loop Unstable – This is characteristic of reactions that are autocatalytic. The faster it reacts, the hotter it gets, and the greater is the initiation of other elements to react. Here, rates and attributes (consumption, inventory, temperature, or pressure) would seem to rise exponentially. Consider nuclear reactors, biological production, or exothermic chemical reactions. An explosion could also be characterized this way, but is on a very rapid time scale. However, the open loop

unstable characteristic is only valid with an infinite inventory. In reality, after a period of nearly exponential acceleration, the process settles as inventory is consumed. Using the open loop unstable transfer function, $1/(\tau s - 1)$, is just a partial representation of the initial stage of a process behavior, not the entire story.

15. Inexpensive computers – The CPI cannot justify the expense of computers as can, for instance, aerospace, military, critical infrastructure utilities, or health-care applications.
16. Simplicity – Processes and controllers in the CPI are managed by operators, who might have an associate degree, and if they do, probably a degree that is not technical. Tuning and diagnosing abnormal situations must be easy – at the non-college bound, High School, infrequently refreshed, skill level.
17. Batch as well as continuous – Control during a batch reaction process must accommodate the nonlinear and nonstationary issues as the batch evolves and its properties change. This would be similar to all of the items above, which relate to continuous flow processes. With batch, however, there is the additional aspect that feed materials change from batch-to-batch, and the recipe needs to adapt from batch-to-batch. This more than tuning or modeling, this is an on-line changing of the models, set points, trigger values, and trigger times.
18. Transitions – Even continuous processes need to make transitions from one product grade to another. The transition might be in product viscosity, polymer molecular weight, or actual chemical makeup. In the extreme, transitions for a continuous process include startup and shutdown. The transition needs to happen as fast as possible and to minimize cross-contaminated waste; but also, it needs to be safely managed. Environmental, Health, Safety, and Loss Prevention (EHS&LP) are fundamental concerns. During transitions gain, time-constants, and deadtimes might change on a 5:1 range.
19. Bumpless Operation – Controllers start in the manual (MAN) mode, and are placed in automatic (AUTO) mode when the operator feels that the system is stable. Often, operators switch from AUTO to MAN to make their own corrections, then switch back to AUTO when comfortable. Bumpless transfer is the term to mean that the controller output and set point do not change to alternate values in the transition. The MAN mode is not the OFF mode; and in MAN, the controller model, set point, and model bias should be updated. Consider the cruise controller on your car. When you get up to speed then press the cruise control button, you don't want the MV to go to zero, or to its last use value. As an additional example for the need for bumpless operation, when tuning is adjusted, you only want it to affect new actions, not scale the past history. Consider a controller with an integral multiplied by the gain and divided by the time-constant. If the controller accumulates the integral $I_i = I_{i-1} + e_i$, then computes action by $u_i = K(e_i + I_i/\tau)$, which is how the PI equation and block diagram seems to indicate the code sctructure; then changing a coefficient value rescales the integral history, which makes a bump in the controller output. For bumpless tuning in the AUTO mode, the integral should be calculated as $B_i = B_{i-1} + Ke_i/\tau$, then compute action by $u_i = Ke_i + B_i$. The control action is identical under normal operation. To differentiate the approaches to computation, I renamed the scaled integral as controller bias, B. Test your controller to ensure that it provides bumpless operation.

Since I don't think simple-to-implement problems will include each of those 19 aspects, I suspect that controller testing will require several benchmark challenge problems. I am offering three on my web site [2], which are open code and free to use. I have used these for classroom teaching and control algorithm research. These are Hot & Cold mixing, pH in a CSTR [3], and an automobile speed controller. I am anticipating to post others in the future, and welcome suggestions or examples from readers.

III. Criteria for Assessing Controllers

How are controllers evaluated? This has two aspects, one is how the tests are performed, and the second is the choice of metrics that are chosen to relate to desirable and undesirable aspects.

First, how to test: Classic tuning criteria for PID controllers are based on step testing of the set point, and controllers are often evaluated on how well they make the process follow the set point. However, most chemical processes remain at a single set point, and the main controller job is to reject disturbances. Perhaps a set point is changed once per day, and the event lasts for an hour; so for 4% of the time the transition would be observed. But even during this time, the need for regulatory rejection of continuing disturbances remains. Goodness of control during the regulatory control period is more important to the CPI than set point following. So, investigators should use regulatory metrics, not set point following metrics. While the simulated process is being controlled, noise and disturbances should continually seek to torment the controller, as they do on real life. The demonstration should not be just a response to one step-and-hold in a disturbance. The simulation should be long enough to simulate the population of disturbance and noise confluences that a controller might encounter. It is good to explore set point responses, but emphasize regulatory performance in your testing investigations.

Second, how to assess goodness: Here is an overview of six aspects.

There are several key performance indicators of a controller. One is its ability to keep the controlled variable (CV) at set point. Another is the manipulated variable (MV) work that is required to do so. There are several acceptable choices of goodness of control metrics for either. For the first, I like time-normalized Integral of the Squared Error (nISE) as a CV performance metric, and for the second, time-normalized Travel (nTravel) as a control effort metric of the MV [5].

A third important CV metric would be measures of any violation of a constraint or specification. pH, for instance, has associated limits for discharge, usually between 6 and 9. For practical considerations, a controller that can keep the pH between 6.99 and 7.01 is no better than one that keeps the pH between 6.5 and 8. Both meet specification. Although an nISE of 0 beats an nISE of 5, the “improvement” may be purely academic, and inconsequential to the customer or to society if neither cause discharge violations. Although CV metrics like nISE are important, often a more important CV metric is related to the quantity (material or time) and degree of violation of a limit (product, regulatory, or safety). If out-of-specification, the material needs to be recycled, held, converted to waste, or the product downgraded. If safety or regulatory limits are exceeded there is the probability of an undesired event or fine. In any case, there is a cost penalty, and in any case override action needs to handle the material or control the process. So, a third metric would relate to the quantity and degree of limit violation. I calculate the product of quantity times the violation, and sum it over an extensive regulatory simulation to explore the population of events that Nature might contrive. As I’ll discuss later, process variance is related to nISE, and a strong indicator of quality give away and/or extent of specification violation.

Fourth, and often overlooked in assessing desirability of a controller is the operator and engineering requirements to implement, initialize and maintain a controller. Process engineers are typically entry-level BS graduates. They might setup a control strategy, and the design and modeling required needs to be consistent with that skill level. But, the day-shift engineer is off-duty ¾ of the process operation time, and when on-duty does not remain at the control panel. Usually control is maintained by the shift-to-shift on-duty operators. Some of these might have an associate degree, and likely several will have had training courses on PID. And, although they may have extensive and valuable process experience, their technical/mathematical skill level is nominally at a low High School level. Their solution to nearly any perceived problem with a controller is to place it in MAN mode, and wait for the responsible person to diagnose and fix it. The CPI is generally adverse to anything complicated, and the K.I.S.S. principle (keep it simple, safe) guides choices. Necessarily for this environment, simplicity of control concept, of obtaining models, of tuning, and bumpless MAN-AUTO transitions are essential. Accordingly, the fourth criteria, often overlooked in science or engineering articles or textbooks about control, is simplicity, related to every aspect of control strategy development, implementation, and operation.

Fifth, there is the cost of the system, the process and the control elements. This includes initial installed cost and annual maintenance and operational costs. Both of which are effectively proportional to the number of devices. Installation and maintenance of both the process and the devices in the control system are dominant issues. Measurements are not free. Advanced control strategies may require extensive process testing (with associated process and material costs) and front-end engineering (setup costs). Be sure to consider these. In addition operating set points need to be safely away from a limit or specification to prevent violation, and the operation at higher purity costs in energy or capacity. Any specification violation generates waste, or material to be recycled, or a hazard; each having a cost impact. Consider all economic indicators.

Sixth, and finally, all of the above listed 19 attributes of a CPI application are important. Since a single challenge problem will likely not include all of those events, it is likely that several simulators for testing would be needed to assess a control algorithm.

All six of these aspects of desirability or undesirability are important. One cannot claim a control strategy is better if only assessing one aspect, or a process that expresses just a few of the 19 attributes.

Since the metrics are diverse (closeness of CV to set point, controller cost, K.I.S.S. rating) I would suggest evaluating control performance in a Pareto analysis method.

Here are some guides on creating and analyzing metrics:

Assessing CV performance:

Although integral of the CV error is valid when overshoots counter undershoots by downstream blending; in my experience, usually one party (customer or manufacturer) is concerned about excess and the other about deficiency in a quality. The customer may be the downstream recipient of material within the same company, and although the upstream unit may be happy about energy conservation, the downstream unit may be unhappy with the lack of transfer-product purity. Alternately, the manufacturer may be concerned about both the excess needed to make something above specifications as well as the waste if it does not meet specs. Since the importance of a deviation normally scales with the square of the deviation, I like ISE, integral of the squared error, where “error” is the CV deviation from set point. However doubling the length of a run, would double the ISE, so I like normalizing ISE by the run time, nISE. Of course, one cannot analytically integrate a discontinuous signal, but if the rectangle rule of integration is used, then mathematically, nISE is effectively the process variance in a regulatory period.

$$nISE = \frac{1}{T}\int_0^T e^2 dt \cong \frac{1}{N}\sum e^2 \qquad (1)$$

The square-root of process variance, the standard deviation, is a critical indicator of quality give-away when set points need to be shifted from a specification to prevent violation of the "spec". If one wishes to be within spec 99.9% of the time then the set point needs to be 3 sigma from the specification (if fluctuation is Gaussian); which is another reason I like nISE. nISE is the variance, and the square root of nISE is the standard deviation (if Gaussian).

Although, related to variance during a regulatory period, nISE can also be used to assess CV performance during set point changes. However, I believe set point following metrics are of secondary importance to application functionality in the CPI.

However, other CV metrics are valid and defensible. For instance if in-line blending tempers overshoots and undershoots, the integral of the error (IE) is a right metric.

If rapid set point following response and rapid settling are both important then minimizing the integral of the time-weighted absolute value of the error (ITAE) is relevant. Normally, ITAE is measured for a deterministic (no noise, no disturbance changes) after a step-and-hold change in the set point, with time zero at the event. ITAE can also be used by measuring time-deviation after a step-and-hold event in a disturbance. Classically, quarter-amplitude damped (QAD) response to a step-and-hold in the set point was the criteria for tuning controllers in the pre-electronic age of pneumatic-mechanical devices. In my opinion, however, ITAE or QAD criteria make controllers uncomfortably aggressive for CPI use, and they lose fidelity to the reality of the process owner's concerns for chemical process operation in the regulatory response to continuing disturbances.

Assessing MV Performance:

Although control energy (the sum of squared MV changes) is often used to measure the work, I like travel, the sum of absolute value of MV changes because I think it better relates to the manufacturer's view of the impact of changing valves or other final element states (such as speed). Normalized travel is:

$$nTravel = \frac{1}{N}\sum|\Delta u| \qquad (2)$$

In any case, assess both MV and CV performance, and emphasize regulatory control to continual disturbances over set point tracking.

Noise and Disturbances:

Noise is considered to be fluctuations on any measurement, which are random and independent at each sampling. Perhaps the fluctuation is due to mechanical vibrations, stray electronic interference, mixing imperfections, the statistics of small numbers (such as in a x-ray analysis), flow turbulence, or sample-to-sample preparation errors. Regardless, it is often considered to be Gaussian, normally and independently distributed, with a mean of zero and an amplitude indicated by the standard deviation, $NID(0, \sigma)$. My favorite method to simulate such is the Box-Muller method:

$$x_{measured,i} = x_{simulator,i} + \sigma\sqrt{-2\ln(r_{1,i})}\sin(2\pi r_{2,i}) \qquad (3)$$

Here $x_{simulator}$ is the PV value generated by the simulator, the true but unknowable value. And, r_1 and r_2 are uniformly distributed random numbers on the interval $0 < r \leq 1$. The sampling interval is indicated by the subscript i.

However, noise might also be considered autocorrelated, if for instance the sensor has a filter to temper noise, or the sources are considered to have some persistence. Autocorrelated noise can be generated by an autoregressive moving average model that is first order in both, $ARMA(1,1)$, driven by an $NID(0,\sigma)$ input.

$$n_i = (1-\lambda)n_{i-1} + \lambda\sigma_d\sqrt{-2\ln(r_1)}\sin(2\pi r_2) \quad (4)$$
$$x_{measured,i} = x_{simulator,i} + n_i$$

Here, n_i represents the noise fluctuation at each sample, and λ, the filter factor, is related to the sampling interval, and the time constant for persistence of the noise. $\lambda = 1 - e^{-\Delta t/\tau}$. Choose a fitting time-constant for the disruptive event persistence (perhaps one or two sampling intervals), then calculate λ. Here, however, the amplitude of the noise will be attenuated by the filter, so to have a desired standard deviation on the series of n_i values, the value of σ in the driver needs to be higher. $\sigma_d = \sigma_n\sqrt{(2-\lambda)/\lambda}$.

Sensor calibration drifts in time, perhaps due the influence of changes in humidity, temperature, fouling, age in lights, dust, etc. This generates a systematic bias on a measurement, which can also be modeled with an $ARMA(1,1)$ model driven by an $NID(0,\sigma)$ input.

$$d_i = (1-\lambda)d_{i-1} + \lambda\sigma_d\sqrt{-2\ln(r_1)}\sin(2\pi r_2) \quad (5)$$

Again, choose values for the time-constant of the calibration drift (perhaps on the order of hours) and a standard deviation for the calibration error (perhaps 2% of the calibration range, then determine values for λ and σ_d as above. If both drift and autocorrelated noise are considered, then

$$x_{measured,i} = x_{simulator,i} + d_i + n_i \quad (6)$$

Finally, no process is left un-tormented by Nature. In driving a car, the road turns, goes uphill and down, the pavement changes roughness, the wind blows in gusts and changes direction as you pass hills and buildings, and changes properties with temperature and humidity as you pass over hill or through dells. These all change the resistance to car motion, which require continual accelerator pedal changes. A chemical process is also continually affected by a plethora of issues. Some of these influences are considered as properties of the material inflow or other external environmental disturbances: inflow velocity, inflow temperature, raw material composition, heat loads resulting from wind rain and cloud changes, etc. And others can be considered as in-process properties: heat exchanger fouling, reactivity poisoning of catalysts, silt or air buildup in pipes, etc. You might classify these as either model input values or as model coefficient values. The same $ARMA(1,1)$ model driven by an $NID(0,\sigma)$ input can be used to mimic such environmental disturbances.

$$e_i = (1-\lambda)e_{i-1} + \lambda\sigma_d\sqrt{-2\ln(r_1)}\sin(2\pi r_2) \quad (7)$$

Again, choose values for the time-constant of the environmental drift (perhaps on the order 15 minutes) and a standard deviation for the calibration error (perhaps 5% of the nominal value, then determine values for λ and σ_d as above. Then use the continually changing value, not the nominal value, as the model input or coefficient.

$$c_i = c_{nominal,i} + e_i \tag{8}$$

This should be done independently for each model input and each model coefficient that is considered to be affected over time.

Off-Specification Cost:

A deterministic simulator, with no noise or changing disturbances, would settle to a single value. However, when subject to measurement noise and disturbances, the output of a simulator continually changes and the simulated measurement has even greater variation. Even with active control the measurement does not settle to one value, but continually fluctuates about the set point. If the set point is close to the specification limit, then the measured value will occasionally violate the limit. The extent of violation could be a count of instances when the CV is out of range, or the sum of the quantity of material out of range, or of the quantity weighted by the violation. My preference is the weighted quantity.

$$P = \sum \Delta t v_i F_i \tag{9}$$

Where P is the simulator penalty, when the violation is v [$v = 0$ unless the CV is beyond the specification, then $v = Abs(CV - spec)$] and the flow rate of the material is F. The sum should be over a long enough simulated time duration to effectively encounter the population of all disturbance realizations.

A good controller will have a penalty of zero. But, a bad controller can also have a penalty of zero, if the set point is far enough from the limit. And, a good controller can have a significant penalty value of the set point is set at the specification.

Quality Give-Away:

If the set point is not at the limit, so that violations are minimized, the deviation between set point and specification is termed quality give-away. On average the quality value is at the set point, and on average the quality value will be better than the specification. Here, the manufacturing process must average some quality measure that is better than the spec, and gives away quality to ensure no violations (or infrequent and of minimal magnitude).

If the perturbations on the CV are Gaussian, then nISE from the regulatory period can be used to calculate standard deviation, and the likelihood of having a spec violation used to determine the quality give away, the distance of the set point from the spec. If the fluctuations are not Gaussian, then this is an estimate, and the actual penalty for deviations needs to be calculated empirically from Equation (9).

The penalty for off-spec material may be inconsequential, such as a temporary hold-up of material in a tank. For instance in pH control of wastewater, if the pH violates the specification simply shut off the discharge and accumulate material in the tank until it can be returned within pH limits. However, the penalty may relate to the cost of recycling or reprocessing material, and amount to the manufacturing added value. Or, if it cannot be diverted, the penalty may be related to discharge fines or customer complaints, which may be ten times the value of the product. The impact depends on the situation, and the amount depends on the choice of quality give-away.

The assessment may have too many economic aspects. But in any case, it is strongly related to nISE. The smaller the variance in the regulatory period, the smaller is the necessary quality give-away, and for any set point choice the smaller is the penalty of off-spec material. So, although these are two very important considerations, they both can be related to nISE.

MIMO:

For a MIMO process, quality metrics for each CV and each MV need to be assessed. It would be convenient to have an overall CV and overall MV performance metric. In my hot & cold water mixing simulator, I used equal concern factors to combine the impact [5]. For example, a 5 °C deviation in temperature set point could be considered as having the same concern as a 1 L/s deviation in flow rate. So, the combined CV metric would be $\frac{nISE_1}{{EC_1}^2}+\frac{nISE_2}{{EC_2}^2}=\sum\frac{nISE_i}{{EC_i}^2}$. Since the MVs range from 0 to 100%, it is likely that the weighting on the $nTravel$ is identical for the two MVs. So, the overall MV metric would be $\sum nTravel_i$. But, if changing one particular utility has adverse impact on other processes, then there might also be equal-concern weighting on the MVs, $\sum\frac{nTravel_i}{EC_i}$. Note: The EC value is squared in the CV metric, but not in the MV metric, to match the numerator dimensions.

Normally, greater MV work will improve the CV closeness to set point. However, excessive MV work upsets utilities, creates operator discomfort, and hurts process efficiency. The process owner seeks a balance, minimizing both $nISE$ and $nTravel$, but without excess in either. One could weight the importance of the two and combine them, $w_{CV}\sum\frac{nISE_i}{{EC_i}^2}+w_{MV}\sum\frac{nTravel_i}{EC_i}$. Or create any number of joint combinations, such as a product, $\left(\sum\frac{nISE_i}{{EC_i}^2}\right)\left(\sum\frac{nTravel_i}{EC_i}\right)$. My personal preference has been to use a Pareto analysis: Generate data for $\left(\sum\frac{nISE_i}{{EC_i}^2}\right)$ w.r.t. $\left(\sum\frac{nTravel_i}{EC_i}\right)$ for multiple experimental conditions, and compare controllers on a dominance basis.

Regulatory and Set Point Following:

Both CV and MV quality metrics need to be assessed for both set point changes and regulatory periods. Often researchers use deterministic simulators, which settle to a steady value. However, the reality is that processes are continually influenced by input variation, disturbances, property changes, measurement noise, and calibration drifts. There is no settling to a deterministic steady value. In a chemical process, there might be a set point change once a shift or once a day, then for the remaining 99.9% of the time, the control is in a regulatory period. Set point response is important, and has evolved to be fundamental to nearly every tuning approach. However, the brief period after a set point change cannot dominate the importance of the regulatory period when assessing controllers.

Cascade and ratio are common CPI control strategies, in which the secondary controller set point is changed by the supervisory (primary) controller. One might consider that the set point following response of a secondary controller can be defined by observing a step-and-hold change in the set point. However, the reality is the set point to a secondary controller is the MV from the primary controller, and it changes continuously in small increments. Step-and-hold set point responses are important, but since this may only happen infrequently, it is much less than even half of the story, and should not dominate analysis.

Over the Entire Operating Range:

Both CV and MV metrics need to be assessed at all of the operating conditions. Chemical processes are nonlinear. An investigator could tune a linear controller for great performance at one set of conditions, and demonstrate how good it is. However, the controller may become dysfunctional at other conditions.

In the CPI, de-tuning of a controller is common, perhaps by lowering the gain, or increasing the integral time, or both, to retain stability at other process operating conditions. De-tuning is also a trick to dampen interaction between controllers. It seems common in the CPI that the process gain will change over a 3:1 range as a result of operating conditions. It is not uncommon to have process gain change over a 10:1 range, and even switch sign. In pH control, the process gain could change by two orders of magnitude.

So, for a two CV process there would be 2x2=4 regulatory periods (if only High-High, H-Low, L-H, and L-L CV combinations are considered). And, there would be set point changes to each of the 4 state combinations, making 4 set point changes, totaling 8 testing conditions that should be evaluated.

Replicates:

Since the vagaries of the environmental influences are random, for any test I typically use 100 replicate trials and a simulation period of a day or so each. Averaging nISE and nTravel over that many trials keeps the uncertainty on the quality metric to less than about 1% of its value, which I think is adequate to differentiate controllers. If performance of one method is within a few percent of another, there is not adequate justification to claim one is better.

Constraints:

One should also test control transitions on and off of constraints. These are important. Controllers regularly would like to push the MV to beyond its 0 and 100% feasible values in response to any of several situations. The individual set points may be infeasible (such as asking my car to reach a speed of 1,000 mph). Alternately a combination of individually feasible set points may not be jointly feasible. (Consider mixing hot and cold water: A moderately high mixed water flow rate may be jointly achievable with both hot and cold valves close to fully open. But, along with a very hot or very cool temperature set point, it would require nearly all of the flow to be provided by one stream, not both, which may not be feasible.) Additionally, it could be that the set point combinations are normally feasible, but a confluence of environmental influences prevent the conditions from being achievable during a particular operational period. Alternately, the controller may encounter constraints because an operator has opened a bypass valve, and even at a 0% signal to the valve, there is still flow. This feasible limit could be any value within the 0 to 100% range, not just the nominal 0 and 100% limits, resulting from the override of the signal to the valve by an auxiliary controller or an operator, or a disconnection in the control communication (such as when a valve is being repaired). Finally, an auxiliary controller may have overridden the primary controller signal to the MV, or an operator may have done the same.

A classic problem for integral control action is windup. A good controller should not windup, but should keep the integral value at the point where it can immediately let the MV respond when the infeasibility is removed. Whether the controller explicitly has an integral or implicitly has the equivalent action (for instance, by using an incremental accumulation for the MV, or in adapting a model coefficient), the controller can windup. If a controller winds-up to an excessive value, then when the constraint is removed, it remains at the limit until it can wind-down. This is detected by a delay in the controller acting after the constraint is removed.

The issue of constraint handling is larger than just anti-windup features in the controller. A question is, if one valve hits a constraint, does the controller use the other valve to keep one variable at set point, or adjust the unconstrained valve to balance deviations of both CVs from set points. An alternate strategy could be to use the remaining degree of freedom (or both valves) to balance both CV deviations, based on the relative importance of both. What does your control strategy do? Does it ignore the one deviation, or seek to compromise both set points.

Additionally, about constraints, it is not uncommon for operations to desire a rate-of-change constraint on the MV. Extreme instantaneous changes in the utilities stress equipment and lower energy efficiency. One might want to temper the control action. This could be done by detuning the controller; but alternately, it is occasionally effected by limiting the sampling-to-sampling change in the controller output. How does your controller implement a rate-of-change limit?

Finally, a constraint issue that is important in model-predictive control (alternately termed Advanced Process Control) is the avoiding of a constraint on an auxiliary variable. The auxiliary variable might be related to safety (such as coil temperature in a heater, or a lower explosive limit in vapor composition) or operation (such as cavitation in a valve, or condensate freezing on a chiller unit, or the "hold point" of an impurity in cement production). This auxiliary constraint may require tempering control action on the CVs to prevent the auxiliary variable from violating a limit. Further, as done in MPC, rather than waiting to see if a constraint is violated, control action at the current sampling is tempered to prevent a future constraint violation as predicted by the model.

Including investigation of constraints to the trials (4 regulatory, and 4 set point following), the testing should also reveal several constraint events. Perhaps 12 trials, each with 100 replicates. The comprehensive evaluation of a control approach is not just assessed by one CV metric in response to a single set point step-and-hold test.

Implementation and Maintenance Cost:

Often missed in assessing controllers is the cost associated with engineering design, installation of equipment, and subsequent maintenance.

A nominal cost of any installed device is $10k. This includes communication lines; and is applicable to temperature sensor-transmitter, flow rate sensor-transmitter, valve, or controller. For a simple 2 CV SISO application there are two measurements, two controllers, and two valves totaling $60k. If you want to use a simple ratio scheme, the additional flow rate measurements and controller cost an additional $30k to install. The maintenance and calibration could be considered as scaling with the number of devices. So, a simple ratio strategy could be considered to have 1.5 times the annualized expense of a simple SISO loops for two CVs. It may be difficult to exactly quantify the costs. For a nuclear application the installation and maintenance cost might be several times greater than that of a field gas-liquid separation unit. But for any application, the installation and maintenance cost is proportional to the number of devices, so one could count controllers, measurements, and final elements to compare one control strategy to another.

If one is considering a sophisticated model-based controller, one also needs to include the additional costs associated with engineering investment to generate the controller, solutions for data acquisition issues (such as sensor fault detection and correction), operational cost associated with step-testing on the process, and training of operations personnel to take ownership. Something that seems simple and obvious to a PhD researcher, might seem very complicated to a process engineer, and not be comprehensible to the series of operators needed to manage it over years of operation. The cost of a control system is more than just the cost of installed devices. Your assessment should include the fees or salaries required for continual supervision of the strategy. But this might be difficult to assess. I would offer the engineering and setup cost is proportional to the complexity of the algorithm, so that one could multiply the installation and maintenance cost of a standard PID strategy by the complexity of the proposed strategy over PID. You could count lines of code, number of user adjusted coefficients, stages in the set-up procedure, and time extent of process testing and scale this my the corresponding aspects of a standard PID strategy as the complexity multiplier for the number of devices.

Overview:

Any tests of a controller should be grounded in assessment of all of the important factors of 1) CV performance, 2) MV performance, 3) constraint handling, 4) cost (devices and people), 5) quality give away and penalty for specification violations, and 6) any of the 19 issues that are relevant process-specific characteristics. And, each should be considered over the entire range of operating conditions (including constraints) and the vagaries that Nature will use to taunt the controller.

IV. Terminology About Modeling for Control Testing

The term “model” has many meanings. This can create communication misperceptions.

I’ll use the term model to mean mathematical model, the calculus-algebra equation representation. Note: the term “model” also refers to the mental understanding that leads to the derivation of those equations; and to differentiate meanings, I’ll term that the “conceptual model”.

Alternately, the term model also refers to a small-scale version of the real process. These are important in lab-scale testing because of the relatively low cost in operating them. And, I applaud those who demonstrate controllers on lab-scale or pilot-scale real processes. But, in most studies researchers simulate the processes with mathematical models.

The term “simulator” will be used to represent the thing being controlled. You might want to call it the process, or the plant, or device, or system. But, in any case, investigator is likely using a simulation of the real thing, not the real thing. Accordingly, it should be called the simulator, or perhaps the simulated process.

The simulator is based on executable code. Although it is supposed to represent the model, it is written in a structured language not equations. If there is an error in the translation, it is will not represent the model. Usually substantial effort is required to verify the code, to ensure that it is true to the model. Further, the simulator actively presents data revealing how the model would behave in time. The model is just the equations, and describes the arithmetic procedure needed to calculate the variable values. The model does not do anything. It does not actively calculate the variable values. The model is just a set of directions. The simulator executes them.

Often the models are stated as differential equations, and numerical methods are used to approximate them. The models may contain nonlinear relations requiring numerical root-finding. In either case the incremental time and convergence criteria need to be small enough to make the numerical error of the solution to be orders of magnitude smaller (invisible) than what will be evidenced in graphs or assessment data.

The simulator of a process is not the process. The simulator is the computational execution of someone's mental construct of a process. If an investigator is using a simulator, even if it is intended to mimic some real process, it should not be claimed as a real process.

In control testing, we need a process. Typically, we use a simulator of the process. Unfortunately, with a simulator you can see the model behind the code, you can see the code, and you can see all the values that the code is using. The reality of Nature, by contrast, is that you cannot see the truth about Nature. Your conceptual model is not the truth. We don't know the truth. For instance, we have many thermodynamic models (VLE, EoS, properties, mixing rules) that seek to provide an adequate representation for design and optimization. But, none are perfect. We commonly use the ideal, inviscid fluid, squared "Bernoulli" relation between flow energy, $\frac{1}{2}\rho v^2$, and potential energy (pressure); and for convenience, apply it to the average velocity, keeping the squared power in friction factor relations and the orifice equation, even when data indicates that 1.8 might be a better value than 2, $\frac{1}{2}\rho \bar{v}^{1.8}$ [6]. The message is, if you use a simulator to substitute for the process in controller testing, your model for the controller should not be the same model or have the same coefficient or initial or disturbance values that the simulator represents.

If you use the same model for controller and simulated process, that would be cheating, that would be pretending that you can know Nature exactly. Certainly, such cheating is a good thing for the validating and verification stages of developing a controller and test procedures, but don't think that such cheating provides credible demonstration of practicability.

Finally note: Within model-based control, there are several models. The first model is used for one-step prediction of what the process was expected to have done in response to the actual MV implementation. I call this the past-to-now-model (PTNM). Typically, the PTNM is initialized with the prior PTNM prediction value, uses the past control action as its influence, and predicts what the modeled process should have done. (The past control action may not be what the controller calculated if something else overrode that value. Use the actual MV implementation to the process.) The residual, the process-to-model mismatch, the difference between the model expectation and the process measurement, is then used to adjust something. In MPC this is typically to shift the model values over the future horizon. However, in other control approaches, such as IMC or the Smith Predictor or PMBC, it is used to bias the set point.

The second use of a model in model-based control is to calculate the control action. Sometimes this is termed the inverse of the model, because it seeks to determine the control action that would generate the desired control response of the model. In IMC it is the realizable mathematical inverse of the transfer function (with delay and nom-minimum-phase behaviors deleted). However, in MPC this second use of the model is to forecast what the model would do over time given a MV sequence, so that the optimizer in the controller can discover the MV values that best shape the model response. I term this model the now-to-future-model (NTFM). In this case, the model needs to be initialized with the most recent PTNM prediction values. The NTFM is iteratively called by the optimizer to explore trial solutions, and each time, it needs to be initialized with the PTNM prediction.

Clearly, the model used in control is not the process. Unfortunately, custom in control has had us term the model of the process as "the process". Consider the subscripts of empirical FOPDT model coefficients used for reaction curve tuning, and of those in the model of an IMC structure. The controller model coefficients have the subscripts "p" (K_p, τ_p, θ_p). However, the model for control is not the process. Just because you choose a FOPDT model to mimic the process, does not mean that the real process is FOPDT. Hopefully, the model for control is a good representation of the process. But, we cannot see the truth about Nature, so we cannot pretend that we can. Don't succumb to the misperception that you can, because you can see the process model and the simulator code. Your controller model, should not duplicate the process model of the simulator. The controller model should be simpler, functionally dissimilar, and not use the same coefficient or parameter of disturbance values.

V. References

[1] Davison, E. J., "Benchmark Problems for Control System Design", Report of the IFAC Theory Committee, Edited by Edward J. Davison, Committee Chair, International Federation of Automatic Control, https://www.ifac-control.org/, 1990.

[2] www.r3eda.com, "Process Control, Challenge Problems"

[3] Choi, J. Y., H. G. Pandit, R. R. Rhinehart, and R. J. Farrell, "A Process Simulator for pH Control Studies," Computers & Chemical Engineering, Vol. 19, No 5, pp. 527-540, 1995

[4] Rhinehart, R. R., "Criteria and Procedure for Controller Tuning", Develop Your Potential Series in CONTROL magazine, January, 2017, Vol. 30, No. 1, pp 54-55. https://www.controlglobal.com/articles/2017/criteria-and-procedure-for-controller-tuning/

[5] Rhinehart, R. R., Engineering Optimization: Applications, Methods, and Analysis, 2018, John Wiley & Sons, ISBN-13: 978-1118936337, ISBN-10:1118936337, 776 pages with companion web site www.r3eda.com.

[6] Rhinehart, R. R, S. Gebreyohannes, U. Manimegalai-Sridhar, A. Patrachari, and Md S. Rahaman, "A Power-Law Approach to Orifice Flow Rate Calibration", ISA Transactions: The Journal of Automation, Vol. 50, No. 2, pp 329-341, 2011