# Think You've Got a Lot of Data?  Think Again!

This article explores the data requirements that allow plant-wide monitoring and diagnostic technologies to accurately characterize changes in performance.  In particular it examines the requirements of key performance indices (KPIs) common to control loop performance monitoring (CLPM) technologies.  These KPIs are essential to identifying issues and isolating the associated root-causes that undermine both production and asset reliability.

Manufacturers are looking increasingly to technology to make sense of their facility's abundant process data. Indeed, a surge in the number of available diagnostic technologies provides the ability to detect changes in all manner of performance from production output to OEE and asset reliability.  For anyone who has leafed through an industry publication or visited an automation blog the persistent uptick in demand for diagnostic technologies isn't surprising.  Still, for all of the headlines about the impact of Big Data and IIoT on the manufacturing industry, there's only limited content that points out a simple reality: More often than not the masses of data being collected is simply not enough.

**Distinguishing Information from Analytics**
While used interchangeably the words 'information' and 'analytics' aren't synonyms. The data for general information purposes such as performance trending is limited. Whether used to indicate performance relative to safety limits or production goals, a single data sample every minute is often all that's displayed on control room HMIs. One reason is that the human brain can only process a limited amount of data. According to researcher Fermín Moscoso del Prado Martín of University of California, Santa Barbara that limit is just 60 bits of data per second. That's 'bits' with a very small 'b'. Another and more common reason is the production facility's limited data storage capacity. With lower cost sensor and higher capacity storage technologies, the steady influx of diagnostic technologies is pushing manufacturers past their traditional use of information and into the more actionable realm of analytics. Even so the data required to fuel these advanced technologies continues to lag.

More often than not HMI trends showcase data that is sampled once every minute.

The word 'trend' itself describes a general direction in which something is heading whether that direction indicates Up vs. Down or Good vs. Bad. Such information is not intended to provide a detailed picture of a system's health or a process' performance. Rather, trend information is generally displayed in relation to historical benchmarks or physical constraints. Only when a threshold is exceeded is a detailed investigation or corrective action considered. That's precisely what distinguishes analytics from information and why analytics requires so much more data. Whereas information trends need only limited data to indicate a general direction, analytics require as much data as is available to proactively uncover hidden insights that will impact performance.

**An Inconvenient Truth about Data**
Increasingly analytics capabilities are being developed to proactively optimize process control. According to a study produced by Future Market Insights the market for analytics is projected to exceed USD 13 Billion by 2022. Just as predictive analytics continue to gain traction in asset reliability, so too control loop performance monitoring (CLPM) solutions have become a staple across the process industries for continuous improvement. CLPM is software that analyzes a production facility's process. It focuses on the performance of a facility's core regulatory control systems, the 100s or 1000s of PID control loops that regulate production plant-wide. Those regulatory controllers serve as the foundation on which advanced control systems operate, and their efficient operation has been shown to be instrumental in achieving and maintaining profitable production.

Failing to sample data sufficiently fast relative to the Time Constant of the process can result in a condition called Aliasing. While the necessary sample rate will vary from process to process advanced analytic technologies depend on access to high
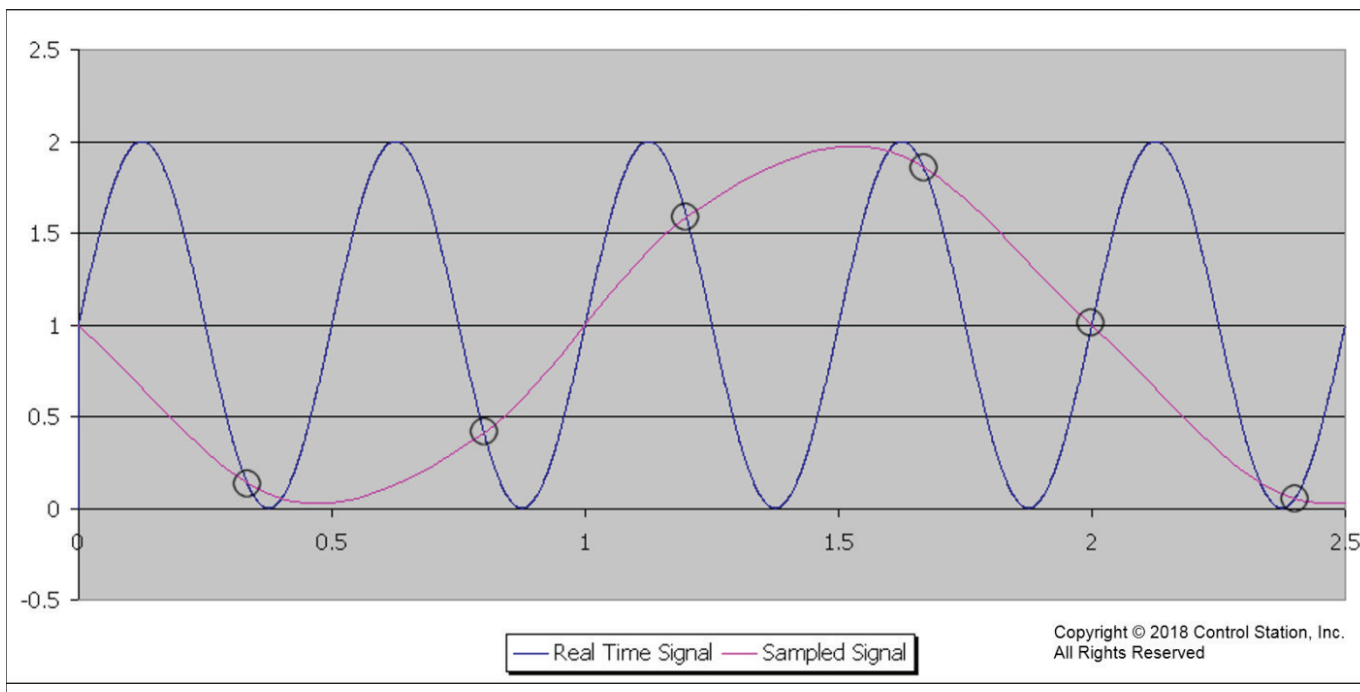
*Failing to sample data sufficiently fast relative to the Time Constant of the process can result in a condition called Aliasing. While the necessary sample rate will vary from process to process advanced analytic technologies depend on access to high resolution data in order to fully capture relevant details.*

resolution data in order to fully capture relevant details.

While the data requirements of key performance indices (KPIs) vary, there are minimums. What's widely understood is that most metrics offer insight when they're provided with higher resolution data. Indeed, in the world of process control the sample speed for data is dictated by its Time Constant. The Nyquist-Shannon theorem stipulates that such data should be collected at a speed ten times faster than the process Time Constant. That's a necessity for fully capturing the relevant detail within a process. Whereas data from typical Temperature, pH and Concentration processes can be sampled at rates between 5- 30 seconds, faster processes such as Level, Flow and Pressure usually require data sampled at 1-2 seconds. This underscores the challenge with the 60-second standard. If the documented findings from Nyquist and Shannon are to be believed, then data collection practices are woefully inadequate on an industry-wide basis. Manufacturers looking to leverage CLPM solutions as a means of improving process performance must first confront the limitations of their data sampling and storage practices. Fortunately for those concerned

about the storage requirement there is good news: Most CLPM solutions only require the data to be stored for a limited time.

**Empowering Advanced Diagnostics**
There is a degree of irony in the data requirements of basic vs. advanced CLPM KPIs – a fact that warrants explanation. It is true that basic KPIs do not require access to high resolution data whereas the more advanced metrics included in CLPM solutions are dependent upon access to it. Essentially, basic CLPM metrics are used to calculate fundamental attributes such as Average, Variance and Time in Automatic. Such information provides relative insight into process performance, and it can be accurately calculated using the 60-second data that is commonly captured. These KPIs are generally used to track how well a controller maintains Set Point and operates in its assigned mode. Additionally, the output of basic KPIs is often used in the calculation of composite metrics such as a process' overall health. Even so basic KPIs are not the ones that assess the variability associated with complex process dynamics. It is their advanced counterparts that deliver those assessments along with actionable insights.

High speed data is needed to conduct the advanced CLPM diagnostics associated with process and system dynamics, and it is needed to perform detailed interaction analysis. When given access to high resolution 1-second data CLPM solutions can apply these KPIs to offer significant insight. Such data allows CLPM solutions to thoroughly assess valve characteristics in a fashion similar to predictive analytic tools. Specifically, it supports the identification of final control element behavior that either inhibits efficient production or leads to catastrophic failure. Similarly, high speed data permits CLPM solutions to isolate underperforming PID controllers and calculate alternative tuning coefficients that are capable of reestablishing optimal regulatory control. So too cross correlation and spectral analysis used in isolating root-causes are made possible with access to high resolution data. Conversely data that is too slow according to the Nyquist-Shannon theorem undermines any CLPM solution's ability to effectively perform these advanced diagnostics.

Select CLPM solutions utilize high resolution data to automatically isolate process changes, model the associated dynamics,
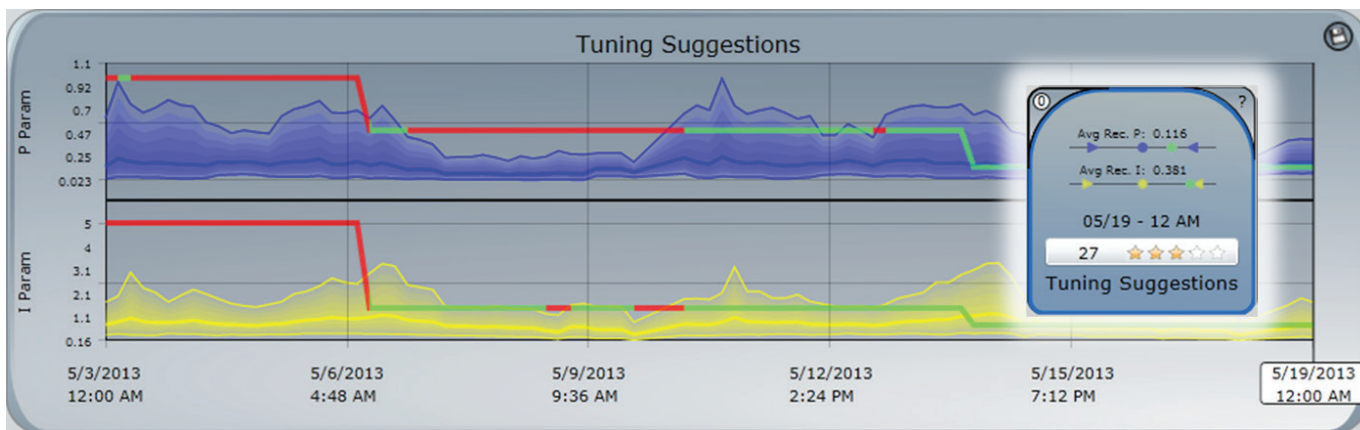
*Select CLPM solutions utilize high resolution data to automatically isolate process changes, model the associated dynamics, and calculate optimal PID tuning coefficients. The image from PlantESP's TuneVue™ utility indicates when existing coefficients fail to satisfy the control objective (Red) and how recommended coefficients would satisfy the same objective (Blue/Yellow).*

and calculate optimal PID tuning coefficients. The image from PlantESP's TuneVue™ utility indicates when existing coefficients fail to satisfy the control objective (Red) and how recommended coefficients would satisfy the same objective (Blue/Yellow).

**A Short-Term Requirement with Benefits**
It is common practice to configure a historian to capture and store information for extended periods of time. In most cases data storage procedures adhere to industry best-practice. In other cases those procedures are dictated by regulatory requirements. The Good Automated Manufacturing Practice guidelines are among the regulations applied to the pharmaceutical industry that assure production facilities maintain auditable production information. CLPM solutions are not as demanding in their need for long-term storage and integrity. Their requirement for storage is limited to the short-term. Indeed, once CLPM solutions perform their analysis the resulting calculations are stored in a separate database and the original process data is purged. As such the need for storing data in a facility's process historian is generally limited to a matter of weeks. That requirement offers a unique benefit during the initial configuration of select CLPM solutions, and it addresses an increasingly common IT issue.

Select CLPM solutions possess the ability to backfill data both during initial solution configuration and whenever access to the historian is lost. The ability to backfill during configuration is unique to a select subset of

CLPM solutions. It significantly reduces the time involved with setup as initial KPI values can be automatically established using a facility's historical process data. These same solutions allow benchmarks to automatically update as the performance of individual control loops improves, thereby supporting the general goal of continuous plant-wide improvement. The ability of select CLPM solutions to backfill data also addresses common IT disruptions. Such disruptions result from routine updates to a plant's IT infrastructure, network or switch upgrades or other communication failures. The potential for disruptions such as these increases in IT environments where multiple historians are utilized whether within a single plant location or across an enterprise. With the ability to backfill for any data gaps resulting from IT disruptions, specialized CLPM solutions assure the integrity of their analytics. Since most IT disruptions are of limited duration the need for long-term storage of high resolution data isn't necessary.

**An Industry in Motion**
Industry continues to move towards plant-wide optimization through advanced analytics. This shift beyond simple information trending tools is evident in the growing number of analytical applications. The common challenge that manufacturers face when implementing these advanced solutions is the shared need for high resolution data. Current data sampling practices fall well short of what's needed to gain meaningful insights and realize the full value of advanced technologies.

As a subset of the advanced analytical market CLPM solutions require access to high resolution data to accurately assess factors that negatively impact plant-wide process performance. The benefits of CLPM have been proven to increase production throughput and quality in addition to decreasing energy consumption and the risks of asset failure. Although they require high resolution data when performing calculations, their benefit is more than worthwhile and the need is only limited in duration. This short-term requirement enables faster solution configuration and it assures the integrity of the solution's analysis considering occasional IT disruptions.

In the end individual manufacturers need to choose between information and analytics. Given the steady growth of the advanced analytics market, it seems clear which of the two, manufacturers will chose.◢

**Author Details**

**Damien Munroe**
General Manager
Control Station EMEA, Roscrea, Ireland
E: damien.munroe@controlstation.com