Ensuring zero surprises and downtime with predictive analytics

Choosing the right software for asset performance analytics

Introduction

Maintaining critical equipment to ensure high levels of reliability, availability, and performance is a primary focus of process engineers in every plant today.

This effort requires frequent, accurate assessment of equipment operating conditions to judge whether equipment meets current production demands and minimizes operational risks of unacceptable schedule interruptions or maintenance costs. In a typical plant, making this assessment involves the collection and effective analysis of large amounts of data about the health of complex production system elements like compressors, turbines, pumps and fans.

The amount of data an engineer needs to analyze effectively is growing steadily as industry incorporates digital technology. Challenging plant economics frequently dictate relentlessly increasing operational demands on ever-aging critical equipment. At the same time, staff budgets are shrinking, experienced people are retiring, and new staff need experience to maintain asset health efficiently. Maintaining efficient operation of aging assets is increasingly costly, while plant economics are challenged by budgets, experienced staff retirements, and training new staff.

It is critical to select the best system for your plant. How do you make the right choice?

Equipment health and performance data come from periodic and real-time systems. Periodic methods for making measurements and analyzing equipment elements include handheld vibration spectral analysis, oil analysis and thermography, and boroscope inspection.

Such solutions provide a great deal of information about important equipment elements prone to functional failure, but they are time-consuming to analyze and intermittent by nature.

To obtain a timely understanding of equipment health for all the key resources in a large plant or fleet, engineers are turning to real-time, model-based solutions. Real-time systems can create actionable intelligence from numerous and diverse sources of data on many key pieces of equipment.

Such solutions can process, analyze, and detect problems and provide the basis for effective diagnosis and prioritization for many problems. Additionally, they can make periodic inspection and maintenance much more efficient.

Technology exists to facilitate prediction of asset failure, allowing engineers to target maintenance costs more effectively. Real-time systems focused on this area of equipment-health monitoring are frequently referred to as equipment condition monitoring (ECM) or predictive asset management (PAM) systems.

Real-time condition-monitoring and analysis tools need to be matched to established engineering processes. Many of these tools are employed at the plant in lean, demanding environments; others are deployed from central monitoring centers charged with concentrating scarce resources to efficiently support plants. Applications also must be flexible and simple to implement and use.

New tools can be very important to the future successes of plant operations, so these choices require a solid understanding of the problems to be solved and the advantages and trade-offs of potential solutions.

Engineers must evaluate several key technology elements to choose the best predictive-analytic solution:

- Core algorithm accuracy and robustness
- Core algorithm execution speed
- Simplicity of model design
- Simplicity and speed of model training
- Simplicity of model results
- Visualization and communication of model results

In addition to these technology elements, engineers must also carefully consider engineering elements that determine whether a solution will fit normal business practices, and thus, provide practical results for the business:

- Ease of use and training requirements
- Opportunity for automation
- Ease of maintenance and skill set requirements
- Flexibility across the equipment scope
- Adaptability across the organizational scope
- Ability to grow with organizational vision
- Modular design and reusability

Any potential solution for improving plant performance that does not consider these technology and engineering elements together can result in poor return on investment (ROI) and a lost opportunity to reach the required higher levels of performance at a reason able cost.

GE Digital's SmartSignal has become a leading predictive analytic solution for improving pl an t equipm ent performance. GE Digital's Asset Performance Management team made a series of carefully considered choices about the software's core algorithm technology, the product user interface that ensures its practicality in the user's engineering environment, and the product roadmap that ensures the investment made today will continue to look like a smart choice long into the future. This paper will describe why SmartSignal and its Similarity-based modeling provides a robust, complete solution for complex asset performance management needs.

The engineer also must carefully consider engineering elements that determine whether a solution will fit normal business practices and provide practical results for the business.

Some modeling basics

Engineers use mathematical models as representations of systems on which to base their understanding and make predictions. Yet these models often a simplified representation of reality.

Consider a first principles model of a simple trajectory. While it is useful for understanding the basic physics of a system and fast to execute, as more complex effects such as air resistance are added, measurement becomes more complex, and the inability to account for all important variables leads to prediction error.

Among modern data modeling methodologies are two groups: parametric and non-parametric methods. In each case, historical measurement data can be used as the inputs and response variables of a mathematical model. Parametric methods make some assumptions about the probability distribution from which a data sample is taken, and use the historic data to find the coefficients of the hypothesized model that best “fit” the data. In contrast, nonparametric methods do not fit to a distribution, but rather determine a model solely from the data. This freedom from assumptions about data makes non-parametric methods versatile, and they have become popular in the ML/AI discipline. It should be noted that a tradeoff of this freedom from assumptions is that non-parametric methods are often complicated to fit and require large amounts of training data, so their use in industrial monitoring applications requires careful consideration and planning. Empirical methods on which commercial condition-monitoring software applications have been developed include non-parametric methods such as linear and nonlinear regression, kernel methods, Kalman Filtering, ARMA; nonparametric methods include Support Vector Machines, Gaussian Processes, Artificial Neural Networks, and Tree-based models.

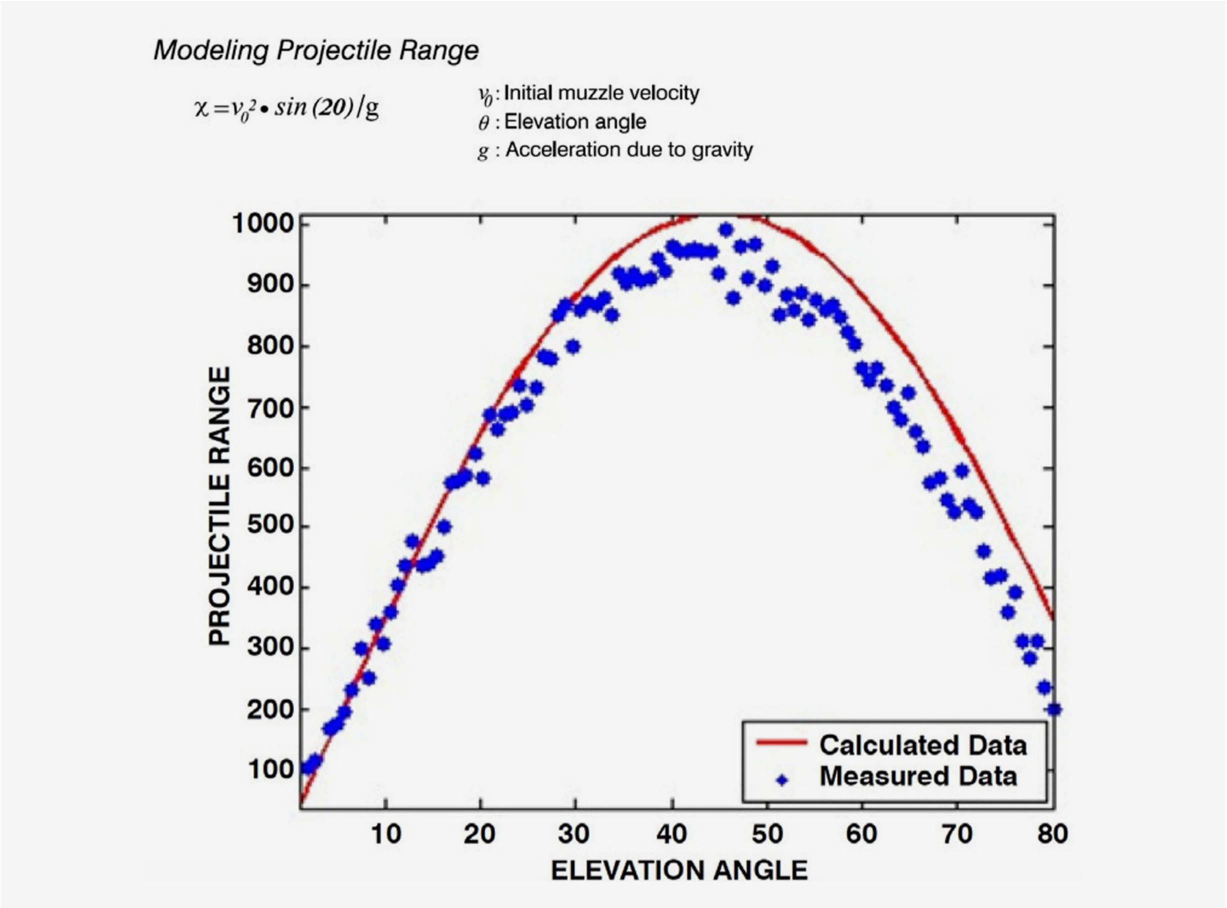


Figure 1: This example of a simple first principles model of a projectile trajectory illustrates the effect of variables not accounted for by the model.

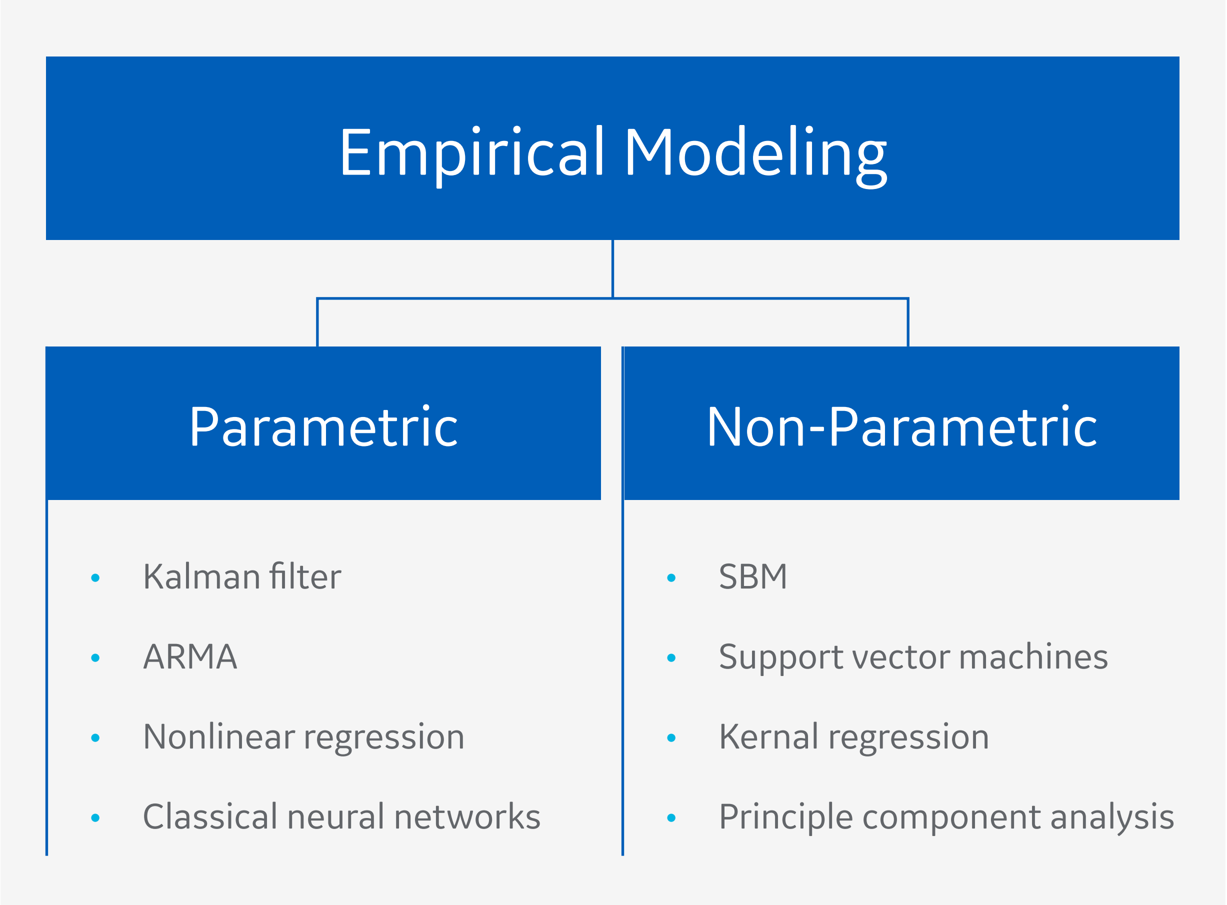


Figure 2: Examples of empirical modeling methods

Key technology elements of a good modeling solution

It was noted above that the choices made in the selection of new tools for improving plant performance can be very important to the future success of plant operations. These choices require a solid understanding of the problems to be solved and the advantages and trade-offs of potential solutions.

A number of elements to be considered in making selection choices were identified and separated into "technology" elements—judged for suitability to the technical challenge of the application—and "engineering" elements that determine whether a potential solution is a good fit for a particular organization. Here is a brief review of the technology challenges:

Core algorithm accuracy and robustness is a fundamental requirement for analyzing complex plant systems. The goal is to provide the highly sensitive ability to detect impending problems at their earliest manifestations in the data. The wide variation in operating conditions of complex plant equipment such as compressors, turbines, pumps, and fans, present challenges for collection of high quality, real-time operating state, performance, and health data. Any modeling algorithm must, therefore, provide sufficient accuracy in early detection of equipment problems while being robust against diminished data quality.

Core algorithm execution speed is required for analyzing complex systems at high sample rates in real-time, particularly when a large number of assets are monitored, such as in monitoring and diagnostics (M&D) centers or decision support centers (DSC). In these environments, instrument counts can be in the tens of thousands, with sample rates as fast as five minutes The chosen method must be fast enough for the current monitoring scale requirements, as well as those anticipated in the future. Slower methods may be useful for offline analysis and comparison.

Simplicity of model design is a critical requirement for engineers integrating real-time monitoring and analytics into their workflow. Models should be transparently designed to easily detect and diagnose specific failure modes. This requirement ensures that the implementation and maintenance of models does not require highly specialized knowledge that would limit their scaling and adoption.

Simplicity and speed of model training is important during initial model building/implementation, as well as subsequent model retraining, system maintenance, and equipment overhauls. Identifying the data for an empirical model that best represents full operating range variation must be an intuitive, streamlined process for the methodology to fit into an already complex, busy process of operating a plant or DSC. Choosing a model whose training process can be automated is highly valuable.

Simplicity of model results means that a plant engineer can use the model results directly to quickly diagnose and prioritize problems, without having to seek specialist resources to gain the proper level of understanding for rapid, certain convergence to the right diagnosis and action plan.

Visualization and communication of model results is a critical element of the successful application of predictive analytics to maintain complex industrial plant systems. A clear presentation of a large amount of data and diagnoses can be communicated through effective visualization.

Simplicity of model design is a critical requirement for engineers integrating real-time monitoring and analytics into their work processes.

Key engineering elements of a good modeling solution

Failure to make good choices about the technical suitability of a methodology for improving plant performance can mean that the capability to provide analytic and diagnostic results useful to plant engineers fails to meet expectations, so the solution fails. Just as important is that the potential solution be capable of dropping into the work process of an organization without unconstructive or undue amounts of adjustment by the organization. Failure to pay attention to such engineering details can result in failure of the new solution.

The following is a brief review of some key engineering elements of a good solution for plant analytics:

Ease of use is a factor that must be considered equally with any other if the solution is going to successfully insert into the successful work process the organization is already employing. Ease of use can include many important factors, including some of those mentioned above as “technology elements.” Ease of use certainly includes ease of training.

The point here is that a solution must not only automatically detect equipment and performance problems early, but must save engineers time deriving diagnostics and prognostics from the diverse and complex data streams that provide real-time evidence of equipment health.

In Figure 3, some of the key steps, from detection of a problem through return to service, are identified. The logic of analytical results from real- time monitoring should augment the roles that the plant engineer already knows and should not require specialized statistical or software knowledge or training. This means that the inevitable organizational adjustment doesn’t eliminate the ability of the organization to utilize the solution.

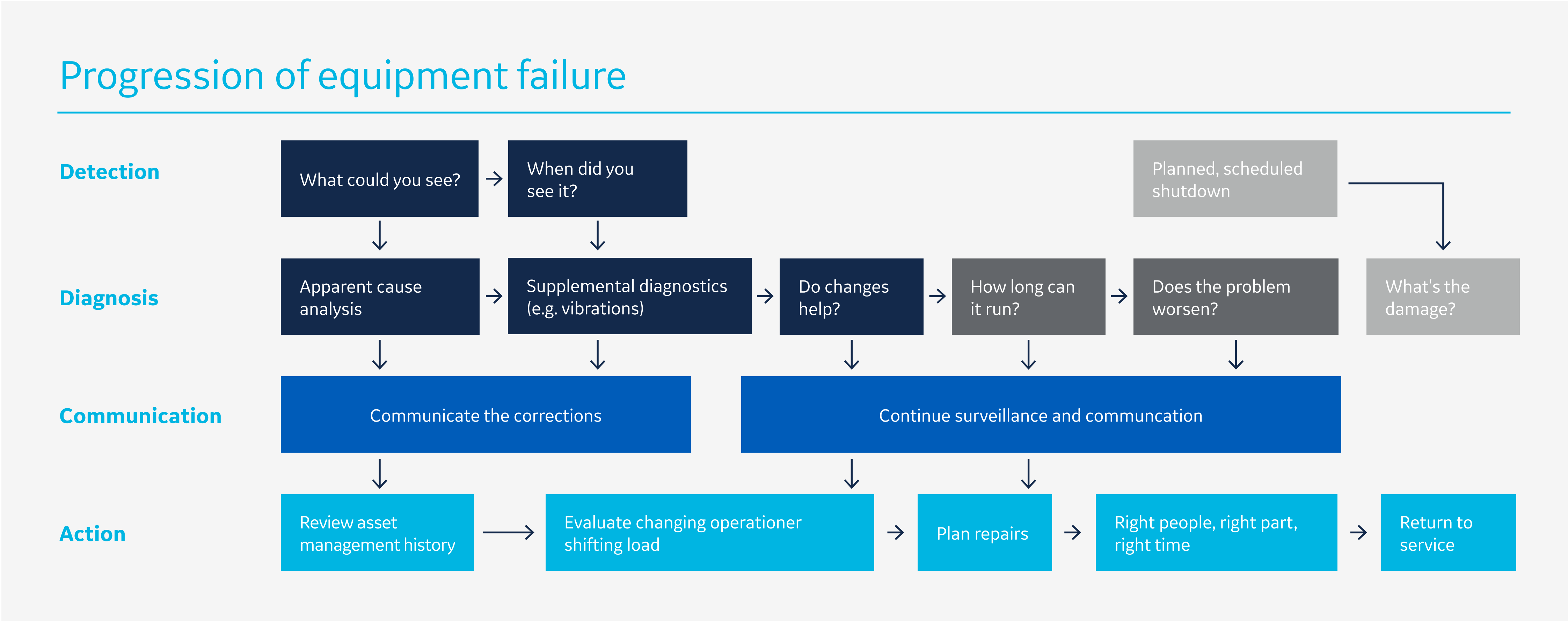


Figure 3: Progression of equipment failure

Ease of maintenance is critical to the long-term success of any modeling solution. Maintenance will have to be done to every piece of plant equipment at some time; when it is done, the analytical model for that piece of equipment will likely need maintenance. Maintenance is normally required for one of the following reasons:

- To adapt models for new behavior (automatic and manual methods)
- To account for sensor availability
- To modify and create new alarms
- To make changes to the model design

A model that is quick and simple to adapt or retrain has an obvious advantage. The skill set requirements for maintaining models needs to be in the range possessed by the average engineer in the organization; maintenance should not require specialized statistical or software knowledge or training. Yet, powerful modeling capabilities will not be maintained by trivial software capabilities, so model maintenance typically would involve either a central monitoring center or a subscription service with the software vendor, which would provide expert support. In either case, interaction with plant engineers having the proper subject matter expertise is critical to proper maintenance.

Flexibility across the equipment scope ensures both ease of implementation and that the number of monitoring solutions is kept to a minimum. A key problem the ECM solution needs to solve is the simplification of many diverse operating data streams into an understanding of current equipment health. This means effective modeling is required for a mix of temperature, pressure, load, vibration, flow, valve position, and other sensors across the broad range of operating conditions that represent normal operation. Monitored equipment is liable to range from the very old to the very new, implying that reference to a standard operating curve may be variable for different instances of the same equipment type because of past duty cycle and maintenance history. And it is likely that one or more of the signals providing operating condition data may have failed or may be failing. The modeling solution chosen must be able to cope with all these conditions to be effective. It must be easily implementable across a single plant or a large enterprise. A single solution that can do so provides a better ROI than multiple solutions needed to cover the full scope of important assets.

Adaptability across the organizational scope is an issue for large organizations. For example, in the power industry, a generating company may have coal-fired, gas-fired, nuclear, hydro, and wind generation capabilities. In the oil and gas industry, the equipment base may include production, refining, and distribution assets. In these cases, some centralization of ECM is likely to be used as a method to drive business transformation, business integration and standardization, and collaboration. Business transformation frequently looks to digital technology to enable the more efficient use of skilled personnel and to facilitate best practices, given that some monitored sites may be very remote and leanly staffed. So, it is critical that any solution for monitoring has the ability to integrate any equipment—despite vendor or lifecycle stage—any operating condition, and any operational culture. It should also help the organization leverage its existing tools used for periodic monitoring.

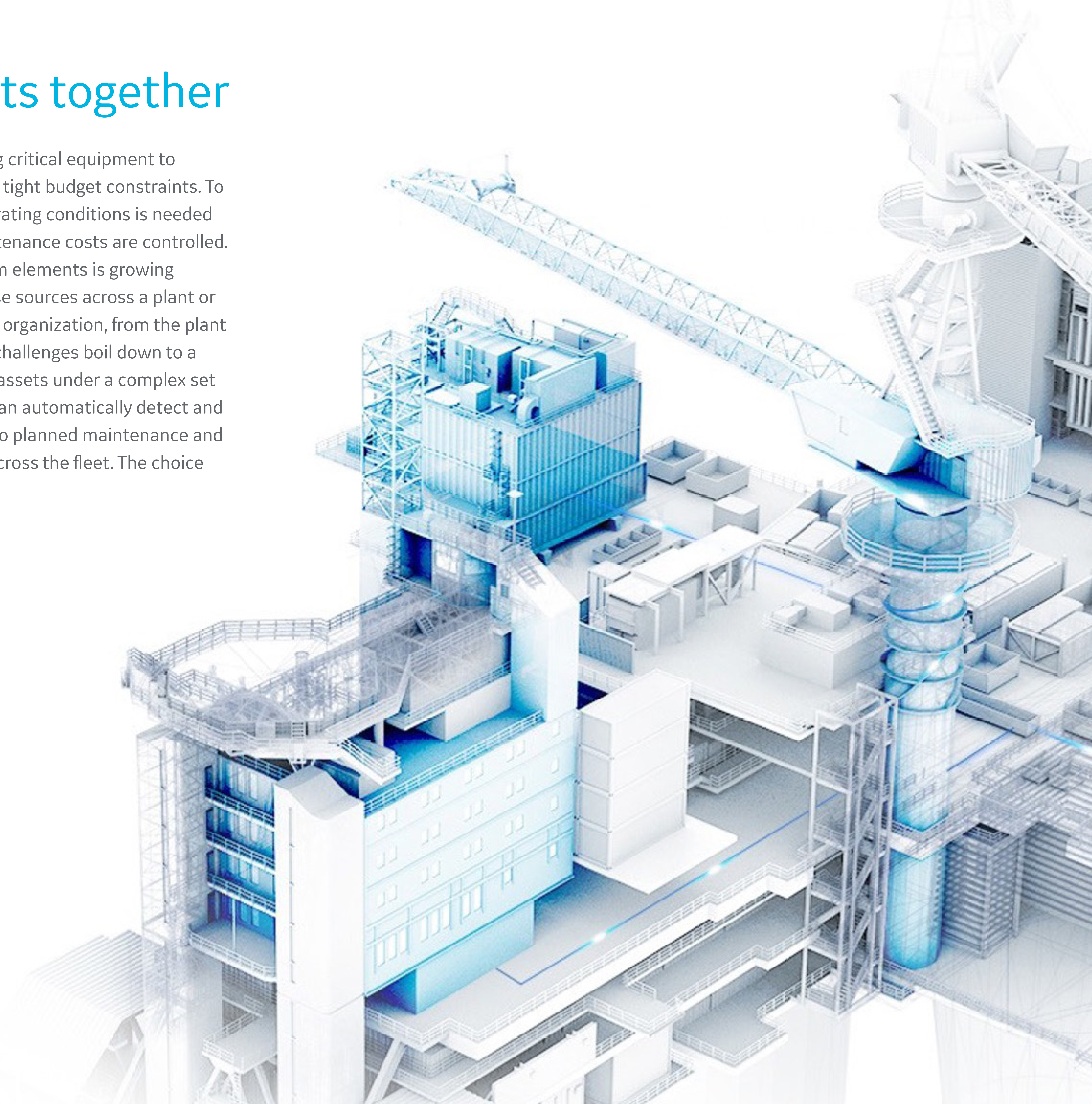
Ability to grow with organizational vision is one of the last and often ignored criteria for selecting the right ECM solution. It is tempting to limit immediate costs when selecting a solution. But, given the complexity of the performance and engineering requirements discussed here, and given that the lifetime of the asset base is in the tens of years, the real opportunity to drive success—from the asset level across the broad scope of the organization—is to select a solution that facilitates the continuous need to produce higher performance and efficiency over many years. Considered from this perspective, the chosen solution needs to be selected on the basis of not only the cost and level of technology, but also on the basis of the people and processes that can be brought to bear by the vendor to solve your problems.A strong product-development vision and a history of past execution are good indicators of an ability to facilitate your organizational vision.

The logic of analytical results from real-time monitoring should augment the roles that the plant engineer already knows and should not require specialized statistical or software knowledge or training.

Putting all the key elements together

Engineers are faced with a demanding responsibility of maintaining critical equipment to ensure high levels of reliability, availability, and performance under tight budget constraints. To avoid operating surprises, accurate assessment of equipment operating conditions is needed to judge whether production demands can be satisfied while maintenance costs are controlled. The amount of data about the health of complex production system elements is growing steadily. Pulling together large amounts of current data from diverse sources across a plant or an enterprise to create actionable intelligence is a challenge to the organization, from the plant engineer to the CIO. The performance challenges and engineering challenges boil down to a need to improve understanding of immediate and future health of assets under a complex set of demands. Here, the focus is the choice of an ECM solution that can automatically detect and diagnose problems early enough to shift unplanned maintenance to planned maintenance and can be easily implemented and maintained across all equipment across the fleet. The choice needs to apply now and as a foundation for future innovation.

Benefits to users include improved understanding of asset readiness, improved maintenance costs, and improved resource utilization—all leading to improved business performance in highly competitive industries.



The best modeling solution: SBM for predictive analytics

We carefully considered the foregoing requirements for a good modeling solution and chose similarity-based modeling (SBM) as the technology foundation for development of a predictive analytics solution to a broad spectrum of real-time modeling needs. Other analytical methods failed to satisfy the key technology and engineering requirements outlined above—therefore, they were rejected. Validation of this conclusion may be inherent in the fact that no other analytic method today has been applied effectively to ECM on such a broad scale as SBM.

Product development has focused on providing and improving the solution requirements for application in the power, oil and gas, and other process industries. Application in these industries requires demonstration of cost-effective value from the single-asset level to fleets of complex assets across several divisions of global companies. In these industries, user benefits include improved understanding of asset readiness, reduced maintenance costs and improved resource utilization—all leading to improved business performance in highly competitive industries.

SBM is a particular form of nonparametric regression. SBM was built as a predictive- modeling solution to the need for actionable intelligence from large amounts and diverse sources of current data on equipment like compressors, turbines, pumps, and fans brought together in today's complex production systems. SBM is quickly built from an asset's historical data, with a structure that mimics a natural engineering design. This produces a result that is quickly and efficiently implemented, from a single asset to the largest corporate scope. Using a sample of the data collected from a complex plant system such as a compressor, a set of "normal" operating conditions can be defined that can be used to reconstruct normal operational behavior in real-time and exclude or flag abnormal behavior.

SBM provides the essential fidelity of the natural system, and it has the advantage of (a) utilizing simple selection guidelines for reference data and (b) requiring no model parameterization. Analytic computation can be done very rapidly, so predictive analytics can be applied as effectively to an entire enterprise as to a single asset in a plant. Model design can follow familiar engineering principles, which facilitates interpretation of results and post-processing operations, like application of diagnostic logic.

Some of the practical technology benefits of SBM that are not likely to be found in other methods used for equipment monitoring include:

Fast and easy setup and execution

- Few model design decisions required
- Models based on engineering logic, not arcane statistical concepts
- Simple guidelines for reference data selection
- Computationally expensive modeling processes done off-line and stored

A clear estimate of normal behavior

- Works for all equipment, all operating modes
- Easy-to-interpret results
- Supports automated diagnostics

Very robust to typical data problems

- Bad data does not disrupt model
- Very tolerant of multiple sensor losses

Brief review of SBM for predictive analytics

SBM is a kernel-based, pattern-reconstruction technique using multidimensional interpolation that is designed to exactly fit training data. SBM produces very stable estimates by using a non-linear and nonparametric kernel (similarity operator) to compare new measurements to a set of reference states (state matrix D)— without making stringent requirements on the smoothness and statistical distribution of the data.

The SBM approach measures the input vector's closeness (similarity) to the observation vectors (states) in the D matrix to generate the estimate for that input vector. This has the effect of deriving a current value estimate from the contents of the training space, normalized to the conditions of the current observation. Weight coefficients are computed by solving a system of equations formed using selected reference data points.

Several studies have shown that SBM technology outperforms other candidate technologies in detecting faults. While other nonparametric techniques can produce estimates with similar accuracy metric (a measure of how closely the estimates follow the actual), SBM outperforms them in robustness (the likelihood that the estimates will over-fit a fault) and "spillover" (the influence a fault in one variable has in the estimates for the other variables).

SBM was designed specifically for the problems of data analysis and diagnostics encountered in real-life equipment. It has been proven through modeling of tens of thousands of assets in power generation, oil and gas, mining, aviation, transportation, and other applications. It can provide accurate estimates for any number of sensors, of any type, over any load range. Even if a quarter of these sensors fails over the course of operations, SBM still can generate accurate estimates for the remaining sensors without following the faulted signals, making it a very robust methodology7. Analytical methods such as PCA and clustering do not have the same level of accuracy and robustness as modeled signals fail.

Model training is an area in which SBM is strong. Training data can be assembled based on subject- matter expertise from empirical data collected in the plant historian. Because the training data can easily be collected over the range of any independent variables of the system, the effect of these variables can easily be normalized. Automated algorithms can be applied for quickly selecting a set of model conditions that represents the full operating range of the data very well.

SBM training is a non-iterative, single-pass operation that involves a single matrix multiplication and inversion. A model matrix, D, represents the entire dynamic range of the reference behavior—selected from historical data, personalized to every piece of equipment. An automated SmartSignal proprietary vector selection method is used to build D. The selection algorithm can rapidly sort through tens of thousands of observation vectors to construct it. SBM, with the SmartSignal training vector selection algorithm, exhibits a very consistent modeling behavior.

SBM allows computationally expensive processes involved in building the model to be done off-line and stored. This allows SBM to operate easily, in real time, for large-scale fleets of assets. The computationally expensive, automated process of building the model matrix D, self-similarity matrix G and its inverse, is done off-line and stored. Because of its unique pattern-recognition approach, interpolation technique, and use of non-linear similarity operator, it can readily model any non-linear function to any desired degree of accuracy.

As with the initial model-training process, model retraining is a quick and simple process, taking several minutes to complete. The retraining for a new operating condition involves a simple inclusion of new training vectors into the D matrix from the new operating range via the same vector-selection algorithm used to create the initial model.

In real time, a multi-step calculation computes the similarity vector A, the weight vector w0, the normalized weights w and, lastly, the estimates for every signal in the model. there is no limit on the number of model variables or the number of models. SBM implementations of 70,000 tags sampled at ten-minute intervals have been made using a single server.

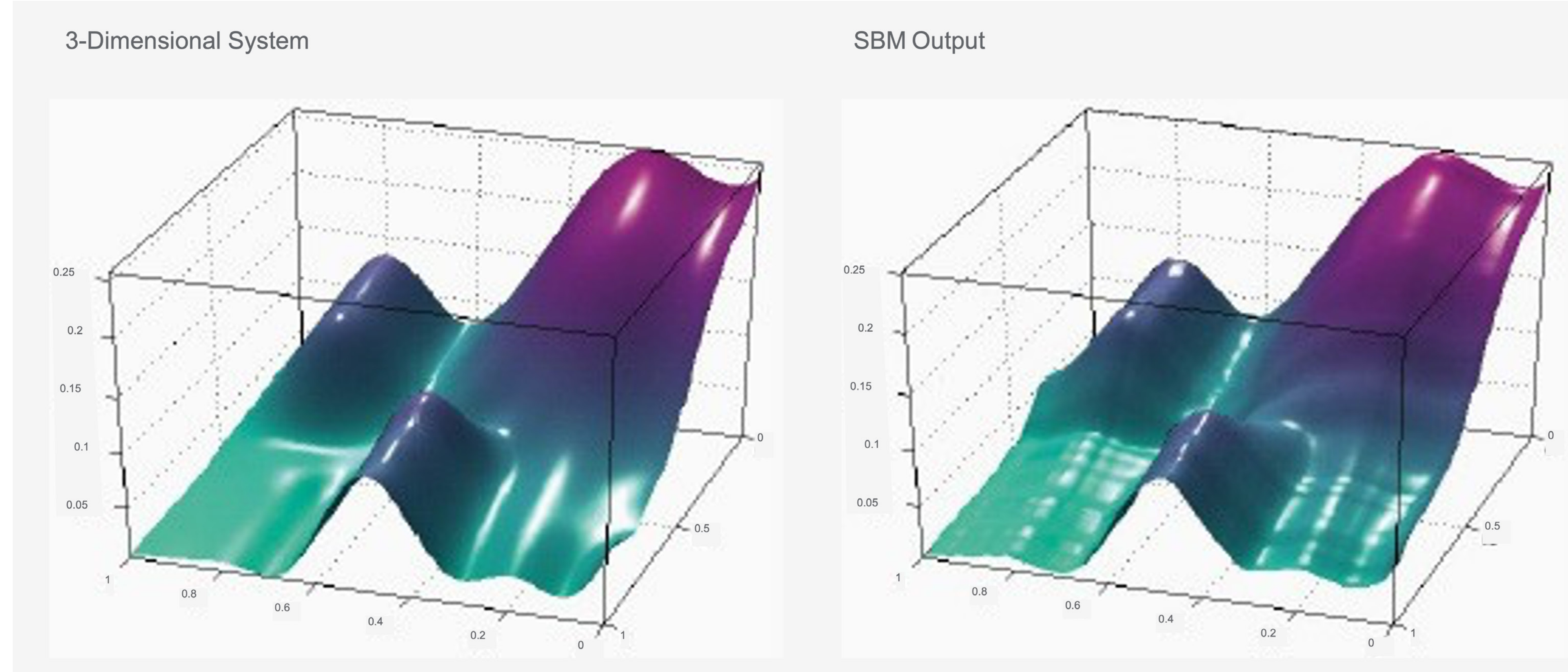


Figure 5: shows the SBM reconstruction of a complex 3-D surface produced by a parametric function, using a sample of data across the surface.

Off-line (training)

1. Select patterns from reference data (Y) to form a state matrix (D)
2. Calculate and invert the interpolation matrix (G)

On-line (estimate generation)

3. Calculate similarity of input to patterns in the state matrix (D)
4. Transform similarities (a) into weights (w0) – then normalize (w)
5. Generate an estimate by linearly combining patterns and weights

1) $Y \Rightarrow D$ 2) $G^{-1}=(D^t \otimes D)^{-1}$ 3) $a=D^t \otimes X_{in}$ 4) $w_o=G^{-1} \cdot a$, $w=w_o/\sum w_o$ 5) $\hat{X}_{in}=D \cdot w$

Figure 5: A multi-step calculation of SBM identifies off-line and on-line parts of the process.

A distinct advantage of SBM is that model designs normally reflect the most sensible structural elements of the asset being modeled. a design for an SBM model is typically as straightforward as designing a physical model for first principles methods. The available instruments for an asset can be partitioned into sub-systems that have physical meaning to the engineer. In the example of a steam turbine, shown in Figure 7, the available sensors are grouped such that oil temperatures (OT), metal temperatures (MT), and vibrations (V) are collected to form a mechanical model. temperatures (T), pressures (P), flows (F), and valve positions (VP) are collected to form individual models of the high-pressure turbine, intermediate-pressure turbine, and low-pressure turbine.

Because SBM was built from the ground up with the detection and analysis of complex plant systems and in mind—like compressors, turbines, pumps, fans, and heat exchangers, it produces results that can be easily interpreted using the subject-matter expertise of the equipment expert, rather than requiring the subject expertise of a statistician or vibration expert. The simplicity of SBM results derives from the simple model structure—this simple structure is not characteristic of statistical methods that have been adapted to real-world modeling problems. Statistical methods such as clustering-based, neural-network-based, or principle components analysis-based, model systems all can suffer from the problem of design complexity unless this is well handled by the application.

Model-design simplicity, the accurate and robust production of estimated normal conditions for each modeled signal and the way this leads to straightforward interpretation of model results, creates another advantage for SBM. The product of the modeling is a set of estimates that mimics the actual data under normal conditions, and it easily shows the trend and magnitude of any differences from normal conditions using a chart format familiar to any plant engineer. This leads to a visualization that complements normal engineering structure, based on familiar failure-mode and failure-analysis representation. In the example of a tube leak in a heat exchanger, Figure 8, the pattern of the failure easily stands out from the normal condition of the equipment.

The simplicity of the SBM model facilitates development and usage of automated expert rules logic to distinguish between normal operating condition and a faulted condition. It can be used to identify a new operating condition, facilitating automated adaptation of models. Expert rules logic also can be extended to provide fault diagnostics in important cases, like the complex plant systems listed above, where "fingerprints" of failure modes are known from development of subject-matter expertise and knowledge capture.

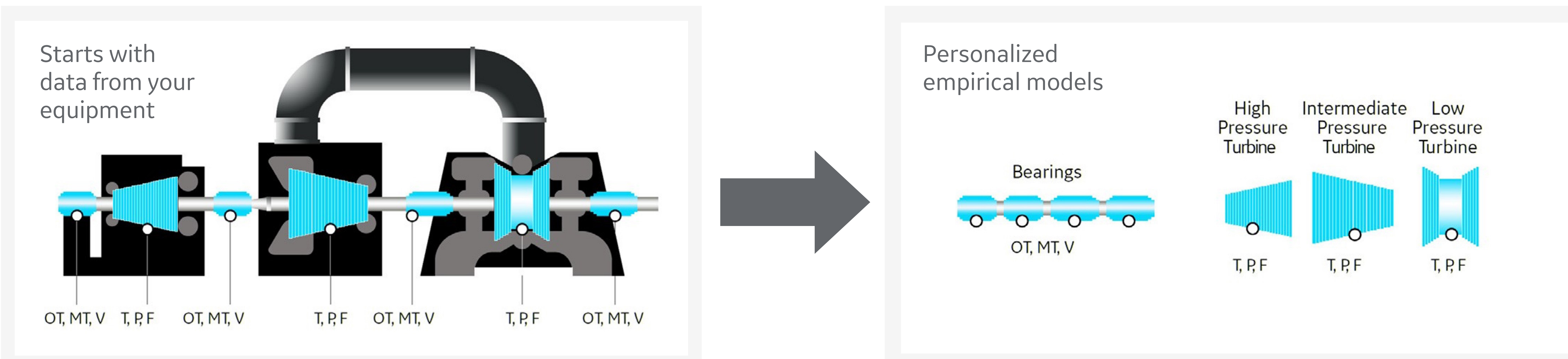


Figure 6: Instruments, like this steam engine, are partitioned into models based on natural engineering components. Use of empirical data to build models results in "personalized" models normalized to natural process variation.

Conclusion

Engineers are faced with the demanding responsibility of maintaining critical equipment to have high levels of reliability, availability, and performance under tight budget constraints. To avoid operating surprises, accurate assessment of equipment operating conditions is needed to judge whether production demands can be satisfied while maintenance costs are controlled. Large volumes of data about the health of complex production system elements are generally available, and the amount of data is growing steadily. Pulling together large amounts of current data from diverse sources across a plant or an enterprise to create actionable intelligence is a challenge to the organization, from the plant engineer to the CIO.

The ECM software product described in this article is focused on providing predictive analytics solutions for control application in power, oil and gas, and other process industries. Application in these industries requires demonstration of cost-effective value from the single asset level to a level of inclusive fleets of complex assets across several divisions of global companies. Benefits to users include improved understanding of asset readiness, decreased maintenance costs, and improved resource utilization, all of which lead to improved business performance in these highly competitive industries. Both technology and engineering challenges must be addressed to provide a comprehensive solution. The technology challenges involve selecting a solution that has the accuracy and robustness to provide early warning of failure under all operating conditions. It must have the speed for real-time application across hundreds or thousands of assets across an entire enterprise. However, technology challenges include more than success at automated detection. The technology must also facilitate diagnostics and prognostics of problems by nature of providing simple connection of equipment design, model design, and interpretation of results. Visual presentation of results that fosters communication and understanding is a critical requirement in order for such software to add value.

In addition to these technology elements, application choice must carefully consider engineering challenges to determine whether a solution will fit normal business practices, and thus provide practical results for the business. To satisfy engineering requirements, the solution must be easy to use and have reasonable training demands. The solution must be capable of managing a wide variety of equipment types, ages, and operating conditions in any plant of an enterprise. It must be flexible in its equipments, adaptable to many different workflows, and low in implementation and maintenance cost in order to demonstrate positive ROI.

Finally, it is critical that any solution chosen must possess the ability to grow with organizational vision—even to help lead it in the case of advanced technology solutions. This requires attention to not only the technology, but also to the people and processes supporting its implementation and integration into your organization.

The foregoing requirements for a good modeling solution were carefully considered, and SBM was chosen as the real-time modeling technology foundation for development the Smart Signal predictive analytic solution. This software solution was developed to meet the engineering needs of the power, oil and gas, and other process industries. Successful application in these industries requires demonstration of cost-effective value from the single-asset level to global enterprise level.

In these industries, benefits to users include improved understanding of asset readiness, decreased maintenance costs and improved resource utilization. In these industries, a predictive analytics solution must improve understanding of immediate and future health of assets under complex sets of operational and organizational demands and provide the analytics support that helps grow a company's vision.

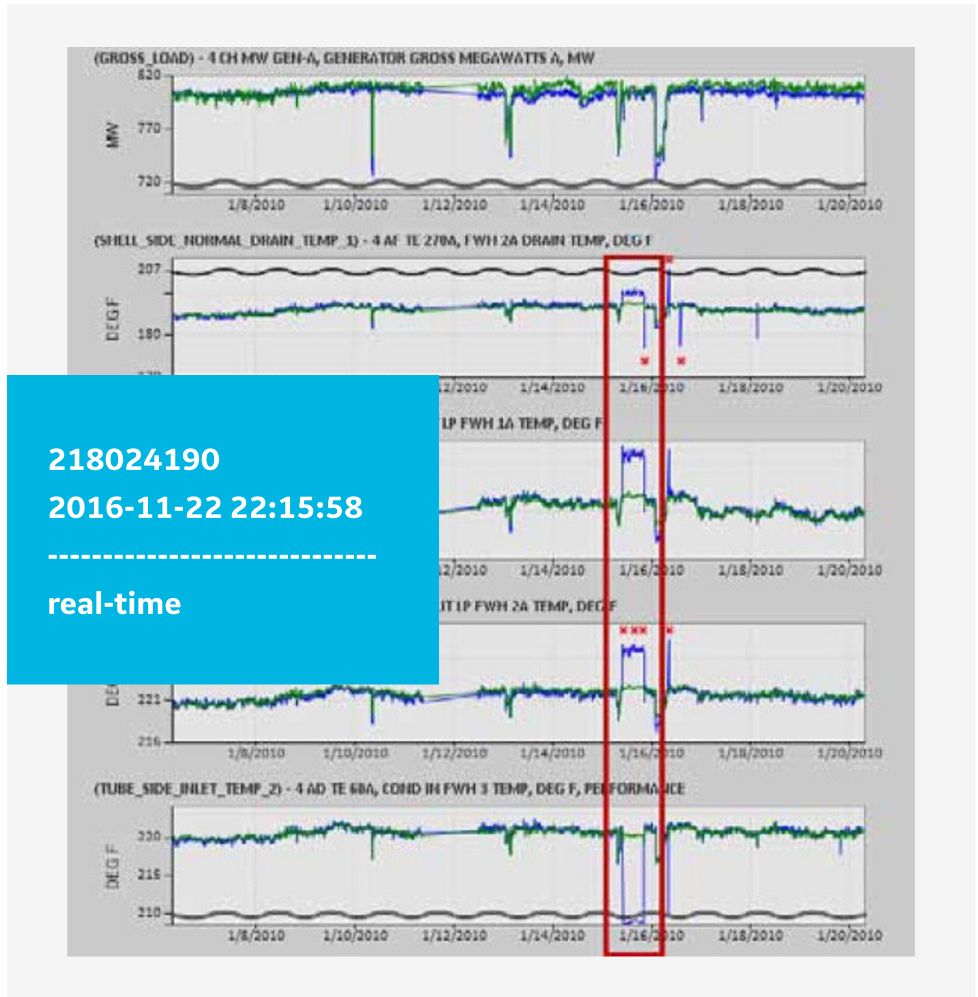


Figure 7: In this example of a tube leak in a heat exchanger, the pattern of the failure easily stands out from the normal condition of the equipment. Blue indicates actual instrument data and green indicates the SBM estimate of normal operation.

About GE

GE (NYSE: GE) is the world's Digital Industrial Company, transforming industry with software-defined machines and solutions that are connected, responsive and predictive. GE is organized around a global exchange of knowledge, the "GE Store," through which each business shares and accesses the same technology, markets, structure and intellect. Each invention further fuels innovation and application across our industrial sectors. With people, services, technology and scale, GE delivers better outcomes for customers by speaking the language of industry.

Learn more about predictive analytics with SmartSignal from GE Digital

LEARN MORE

Contact Information

Americas: 1-855-YOUR1GE (1-855-968-7143)
gedigital@ge.com

www.ge.com/digital