

Honeywell Process Solutions



The Early Event Detection Toolkit

Don Morrison
Wendy Foslien
Ward MacArthur
Peter Jofriet, P.Eng.

Table of Contents

Abnormal Situation Management and Early Event Detection	4
The Value of Early Detection	4
The Technology of Early Event Detection	4
Univariate vs. Multivariate Statistics	4
Process State Estimation	6
Early Event Detection Using PCA Models	6
Offline Design of Early Event Detection Applications	8
Online Interpretation of PCA Statistics	9
Fuzzy Logic Tuning	10
Run Time Environment and User Interface	11
The Engineer's Interface	11
The Operator's Interface	11
Results that Deliver	13

Table of Figures

Figure 1:	Univariate control limits	5
Figure 2:	Multivariate data plotted against Hotelling's Ellipse	5
Figure 3:	Combined visualization of Univariate and Multivariate statistical analysis	5
Figure 4:	not currently captioned	6
Figure 5:	A simplified representation of PCA.....	7
Figure 6:	Profit Design Studio provides a comprehensive environment for empirical modeling.....	8
Figure 7:	Staged approach to online event detection.....	9
Figure 8:	Fuzzy logic configuration for EED applications.....	10
Figure 9:	Engineer's Interface for EED.....	11
Figure 10:	Example deployment of an operator's interface augmented with EED visualization components	12

Abnormal Situation Management and Early Event Detection

Manufacturers recognize the value of an integrated abnormal situation management[®] program as crucial to minimizing the number and impact of costly and potentially hazardous upsets. An abnormal situation is simply defined as a process upset that requires human intervention. This definition may include events ranging from a planned startup or shutdown procedure to an unexpected and potentially serious incident. It has been estimated that the inability to diagnose and control abnormal situations has an economic impact of at least \$10 billion annually in the U.S. petrochemical industry alone, based on research from the Honeywell-led Abnormal Situation Management (ASM) Consortium.

Recent work by the ASM Consortium has shown that a key part of abnormal situation management is the detection of events or process and equipment problems before they become catastrophic or fatal. Operators must daily manage a real-time, highly complex, highly correlated, dynamic environment. Any assistance to receive warnings of pending events is very valuable.

Alarm systems, if well maintained, can provide the operator with guidance during normal operation, and possibly during an upset. Providing some indication of an upset before it generates an alarm is the intent of the ASM Consortium's Early Event Detection (EED) research. More precisely, the goal of early event detection is to identify the presence of abnormal conditions prior to crossing alarm thresholds.

The Abnormal Situation Management Consortium (ASM) was informally started in 1992 and formally chartered in 1994 to enable and empower operating teams to proactively manage their plants, maximize safety and minimize environmental impact while allowing the processes to be pushed to their optimal limits. Original ASM Consortium members included Honeywell, Chevron, Exxon, Shell, BP, Mobil, Nova Chemicals and Texaco.

The Value of Early Detection

Enhancing industrial process performance means improving how well its operators can recognize an occurring or pending undesirable change. This is difficult given the amount of process data that the typical operator must digest in his or her daily routine. An additional consideration is how well the operator can determine what is causing that change.

Greater and sustained process performance is almost continually threatened by process and equipment degradation. These threats are often hidden because they are not directly visible in normal process variable trends, but rather in the underlying relationships between the variables. The process can appear to be fine, but hidden forces are constantly driving the process quietly towards the outer envelope of good performance.

The causes behind these changing variables are interrelated and often subtle. Consequently, the process behavior an operator must monitor, investigate, and redirect quickly exceeds the human capacity to detect and take effective preemptive action. Honeywell's EED solutions help reduce this barrier to improved process performance.

The Technology of Early Event Detection

Univariate vs. Multivariate Statistics

The traditional methods for process monitoring using statistical process control (SPC) methods such as CUSUM charts, Shewhart Charts and/or EWMA charts are not effective on the volume of data that exists in a process plant. Although these methods are simple to implement, univariate methods do not accurately portray what is happening in the large mass of highly correlated data often experienced in today's processes.

Univariate vs. Multivariate Statistics Example

Using standard univariate statistics, control limits are calculated based on normal run data. Univariate Shewhart control charts measure the difference of the current point from the variable mean. Thus, Shewhart charts focus on measuring the amount of variation in a single variable. A point is considered out of control if its value is outside lower and upper control limits (shown in Figure 1). Note that the focus is on a single variable. Interactions between variables are not considered.

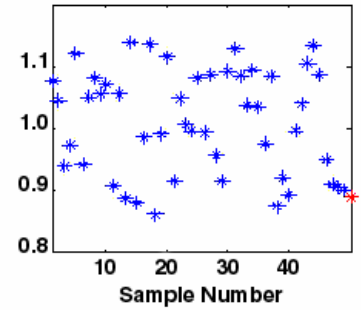


Figure 1 – Univariate control limits

In the multivariate approach, a distance from the center of the normal run data multivariate mean is observed using Hotelling's ellipse¹ (Figure 2). Data that is in control must be inside the confidence limits for normal operation defined by the ellipse.

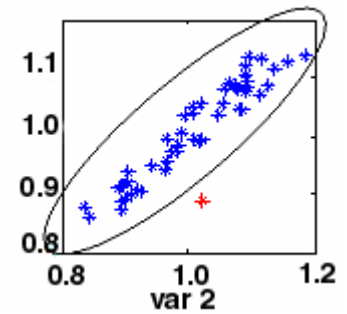


Figure 2 – Multivariate data plotted against Hotelling's Ellipse

If both forms of analysis are plotted together for a common data set (Figure 3) we can make the following observations:

1. Data that is in control (*) for univariate statistical analysis can be shown to be out of control (*) when examined using multivariate statistics.
2. In other cases (not shown) data that is out of control for univariate statistical analysis can be shown to be in control when examined using multivariate statistics.
3. Consequently, the multivariate case provides a more complete indication of the state of the process.

To summarize, the use of univariate control charts can lead to misleading conclusions, or even worse, to false corrective actions. The following sections of this paper explore how use of multivariate statistical analysis can be used for process monitoring in the most effective manner.

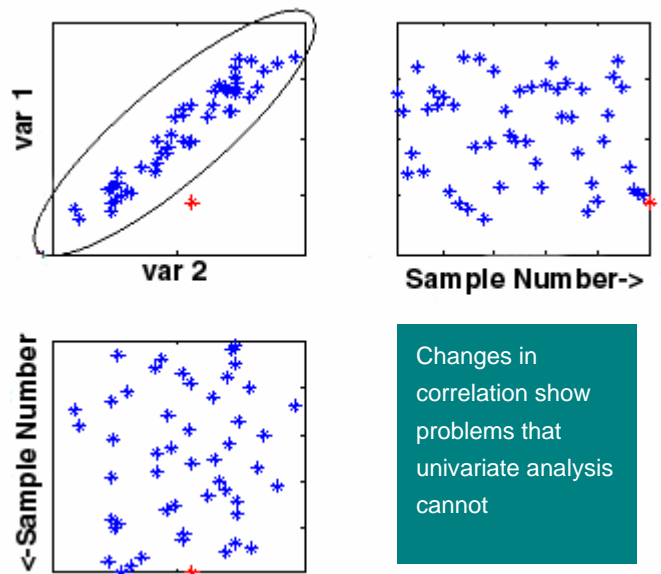


Figure 3 – Combined visualization of univariate and multivariate statistical analysis. Figure is reprinted from paper [1] – T. Kourti

¹ Hotelling's ellipse is also known as the T^2 parameter. T^2 is a multivariate analogy to Student's t distribution statistics (used for calculation of critical values/control limits in univariate statistics).

Process State Estimation

It is crucial to detect a process change of state before it becomes a fully developed abnormal situation. This early knowledge of an emerging abnormality allows for preemptive action which prevents more serious consequences. Given the high turnover rate in today's workforce, it is likely that a company will have significant variation in the experience of their operations personnel. It is likely that the most experienced personnel will be able to detect developing abnormalities based on experience and intuition. However, less experienced personnel may detect an abnormality too late or possibly miss the precursors completely because they are buried in the data. Data-driven technologies that help operators of any experience level focus on early detection of abnormal events are a potential answer to these evolving experience gaps and a means to better overall plant operations.

It has been generally found that the detection of performance changes in a process can be accomplished using multivariate statistical analysis methods originally used in the study of economics, psychology and sociology. The commonality between these areas of study and plant processes is the volume of data, number of variables and the underlying structure found between process variables.

These multivariate methods first measure the interactions between variables under normal operating conditions. Once established, these mathematical models can subsequently be used to monitor a process (grouping of equipment) or specific equipment for changes in these interactions between the variables. Consequently, changes in performance or operation can be found that traditionally would have gone undetected by monitoring individual process variables.

As discussed in the earlier example, multivariate statistics (such as Hotelling's method) can be applied to complex processes to provide better indication of problems than univariate statistics. However, when applied to raw data, Hotelling's approach does have potential drawbacks. First, if variables are dependent (co-linear), calculations leveraging Hotelling's method fail due to matrix singularity problems. Second, Hotelling's statistic detects distances away from the process mean, but not unusual behavior near the process mean. Thus, a frozen sensor may not be detected in the raw Hotelling statistic if it lies near the process mean.

A recent and viable approach to detecting changes in process performance is to use principal component analysis (PCA) and partial least squares (PLS) regression. These multivariate statistical methods are the product of the study of chemometrics (the use of mathematical and statistical methods to relate measurements made on a chemical system to the state of the system). These methods take advantage of dependent variables and thus avoid the issues with Hotelling's approach for correlated data. Studies by the ASM Consortium have found that PCA can provide value by effectively detecting abnormal situations before operators. Consequently, the basis of Honeywell's process state estimation efforts and the EED Toolkit are based on this technology.

Early Event Detection Using PCA Models

PCA is an accepted method to manage highly correlated process variables. Process variables are often correlated because of redundancy of measurements and more importantly, physical laws such as mass and energy balances, reaction kinetics and separation thermodynamics. An illustrative example of variable interactions is the dependence of temperatures on consecutive trays on a distillation column.

Using PCA, the analysis of a large number of process variables from an area or sub-process is reduced to a subset of linear combinations. These linear combinations of process variables are also known as latent variables. This subset is sometimes referred to as a subspace or a mapping of the original set of inputs. The original inputs can be thought of as projecting to a subspace by means of a particular transformation. Unlike the raw inputs, the latent variables are guaranteed to be independent.

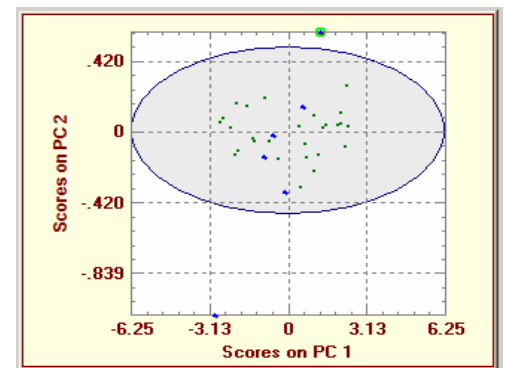


Figure 4 – Score Plot of Principal Components 1 vs. 2 with confidence ellipse

Consider a simple example of PCA analysis, where we reduce a set of four process measurements to two latent variables (or “scores”). These latent variables capture key relationships between process measurements. We can now visualize four process measurements using only the two latent variables. By plotting these two latent variables against one another, the user has a summary view of the process that originally would have required separate trends (see Figure 4).

The principal components capture the significant variation in the data. The first principal component indicates the direction of maximum variation. The second principal component is orthogonal to the first component and captures the maximum remaining variation in the data. The following figure shows the original data and resulting principal components. Notice how the key underlying sinusoidal trend is captured in the first principal component. The second principal component captures the step near the end of the data.

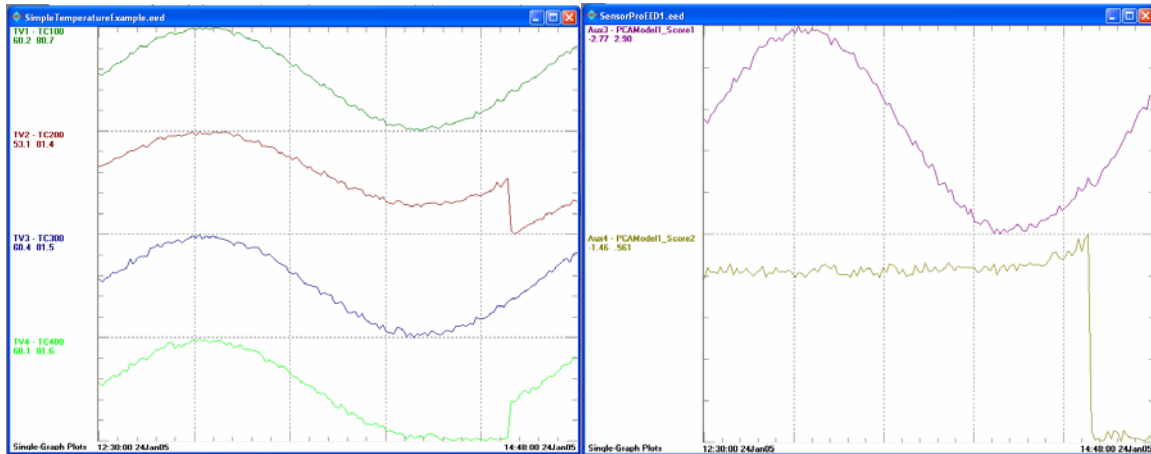


Figure 5: Principal Component Analysis consolidates key process variation from multiple process measurements into a smaller subset .

Once this plane of normal operation is established, there is a benchmark from which to judge future process states. The plane or mapping of the four variables to the two partial components in this example becomes the model. Confidence limits are established around this plane or model to determine the boundaries of the subspace.

Fault detection or process monitoring is done by periodically taking new values of the input variables that represent a new process condition or state. These are mapped onto the established plane using our modeled transformation. Calculated statistics then indicate the proximity of this new state to the previously modeled state.

The question at this juncture is simple. Do the new values project past the boundaries of the model? To answer this question, two statistics are calculated and presented to the user.

- **Q statistic:** Defined as the sum of the squares of the residual errors of the PCA transformation. This statistic indicates how much the correlations between variables are different from normal operation.
- **T² statistic:** Hotelling’s statistic calculated using latent variables shows variation consistent with training data and is mathematically expressed as the distance from the centroid. This represents the amount of drift from normal ranges, as well as how different the variation in current data is from historical data.

Both abnormality indicators (Q and T²) can react at the same time. Abnormalities in the process may be indicated in the data as an unusually large variation, or as a variation in a direction inconsistent with normal historical data, or both.

By calculating the Q and T^2 statistics for a modeled process in real time, early detection of changes in the process can be detected as the statistics cross the boundaries of the plane. The challenge is now multi-faceted.

- How does one develop the models from plant data?
- How are the model results interpreted online?
- How are the model outputs displayed in a way that is meaningful and usable to the process operator?

These issues will be discussed, beginning with the tools to develop EED applications.

Offline Design of Early Event Detection Applications

Design of EED applications is managed with Honeywell's Profit® Design Studio application. Profit Design Studio is the central hub for Honeywell's Profit Suite product line, providing a consistent user experience for all process data modeling needs. Common functionality shared between all applications includes:

- Extensive data import/export capabilities
- Advanced methods of handling time-series data
- A rich tool set for data visualization and manipulation
- A wide variety of data modeling options including dynamic models (multivariable control), ordinary least squares, weighted least squares, partial least squares and principal component analysis
- Integrated online configuration and deployment capabilities for the supported model types

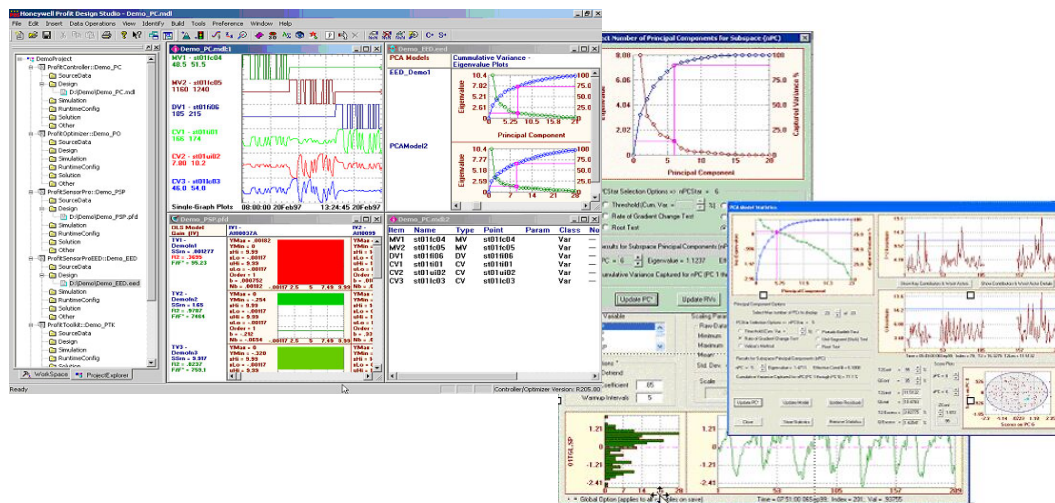


Figure 6 – Profit Design Studio provides a comprehensive environment for empirical modeling.

The Profit Design Studio/EED offline analysis capability can be used as a forensic tool to analyze past incidents and is a key to the development of models to execute online against live data. Principal component analysis is a particularly effective tool for forensic analysis because many variables are summarized into two key statistics. The offline design process requires a set of historical data,

and provides tools to develop and compare multiple models of the same data set. The underlying calculations for the models are automated and transparent to the user, providing an environment that facilitates rapid development and comparison of models.²

Some of the key decisions in developing a PCA model for process monitoring involve selection of normal operational data for training and selection of the number of principal components. Interpretation of the model involves evaluating previously known and newly discovered excursions in the high-level statistics and understanding how this compares to observed process conditions. The EED Toolkit design environment allows the user to select a common underlying data set, tailor the input data for a specific model and compare key components of multiple models at a glance. The detailed interpretation of a single model uses yoked displays to give the user a unified view of key outputs from the model at each time step.

Online Interpretation of PCA Statistics

PCA statistics alone can be problematic when used for detecting early events in live operations. Using raw statistics alone may result in a perception of frequent false alarms by process operators. This results because the raw output from a PCA model is driven by the characteristics of the data used to develop the model. Thus, the sensitivity of the PCA model is tied to the data set used in training. Any statistically unusual observation will be flagged by the model as such. However, in process applications, it is important to note that statistically unusual does not always imply operationally unusual. When an operator is repeatedly alerted to statistically unusual conditions that do not have operational consequences, they are not likely to accept the model as a useful tool.

A second issue with the PCA model is the interpretation of results. The principal components and Q/T^2 statistics are not tied specifically to process functions and thus they need to be interpreted to provide appropriate operational context. These issues are overcome through the use of fuzzy logic post-processing of the PCA model and through proper user interface definition. We will discuss these issues in the following sections.

Honeywell's online EED technology uses a staged approach to detecting faults that applies different technologies along an ultimate solution path. The detection process starts with the PCA model which is attenuated by the use of fuzzy logic. The combination of these two technologies forms the EED Model. The results of this combination (and the underlying PCA statistics) are then displayed to the operator to assist in fault localization. Additional post-processing of these outputs (by way of integration with Honeywell's other Asset Management and Operator Effectiveness solutions) can be applied on an as-needed basis.

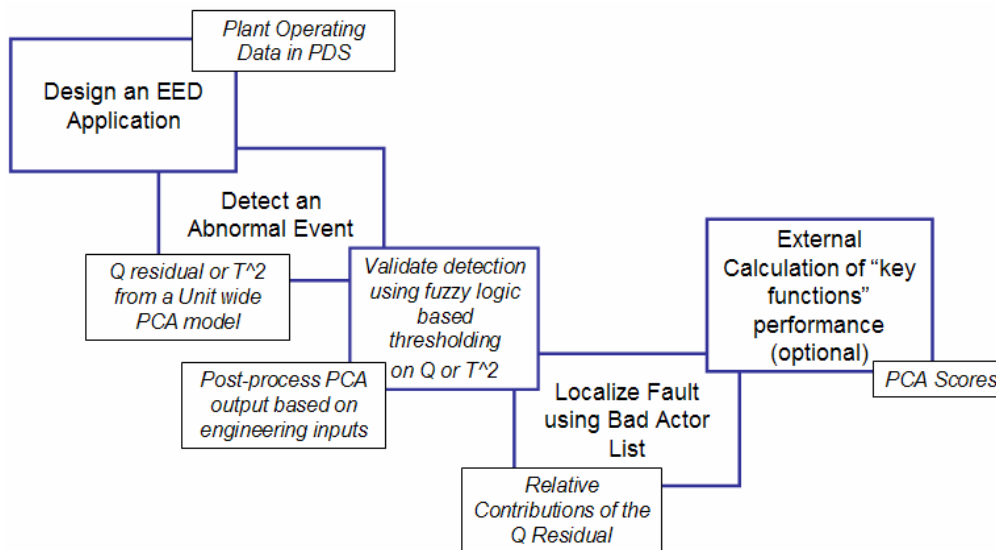


Figure 7 – Staged approach to online event detection.

² Detailed algorithm adjustments are available for the advanced user.

Fuzzy Logic Tuning

Fuzzy logic is leveraged as part of the EED solution to provide an algorithmic means of post-processing PCA data. There are a variety of settings that the user, typically an engineer, can set to attenuate the response of the PCA model to normal process drift and noise, as well as to choose an appropriate balance between model sensitivity and robustness. The decision to apply fuzzy logic as post-processing technology was deliberate. The ASM Consortium has found fuzzy filtering to be a simple, effective approach for reducing statistically unusual outputs that do not have operational significance. Fuzzy filtering is preferred over more complex techniques because:

1. The approach requires minimal tuning by the developer.
2. It provides a consistent filtering approach independent of the modeling process.

The end result of the fuzzy processing is a more robust model that provides the engineer the capability of attenuating detection to minimize false detections, thus improving usability and long-term viability of the solution.

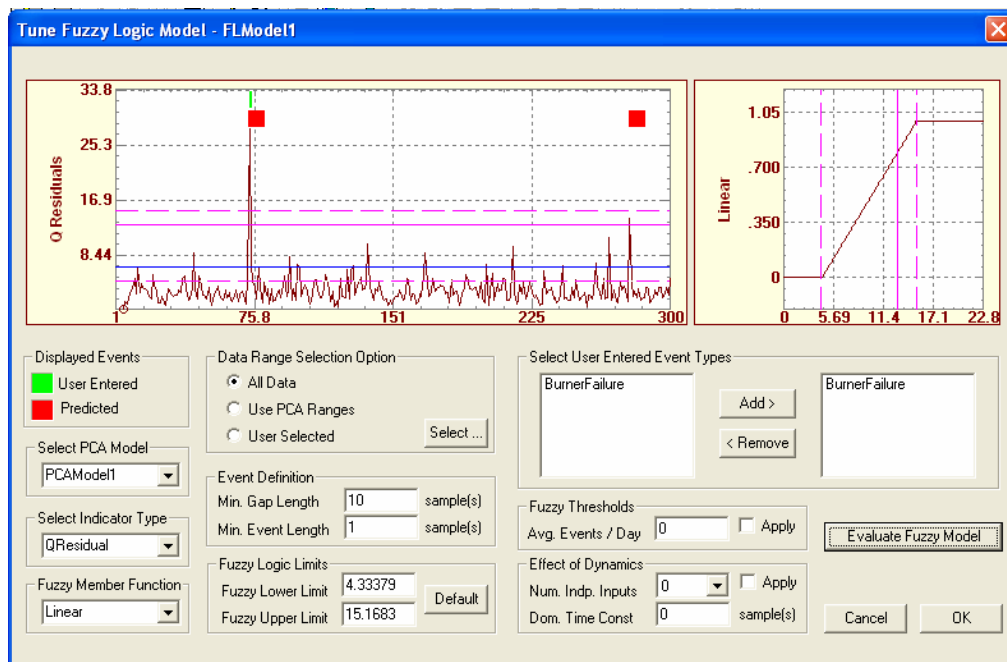


Figure 8 – Fuzzy logic configuration window for EED applications

Run Time Environment and User Interface

Perhaps the most critical aspect of any solution for Abnormal Situation Management is that the user interface for EED defines the success of the application, regardless of what level of rigor is provided in the modeling phase. As such, the EED visualization is designed to address both the needs of the engineer and those of the ultimate end-user, the process operator.

The Engineer's Interface

The EED engineer's interface is designed to be a quick-deployment visualization environment that needs extremely limited configuration. Additionally, the design was largely driven to have common visualization elements with the EED's offline design environment, Profit Design Studio.

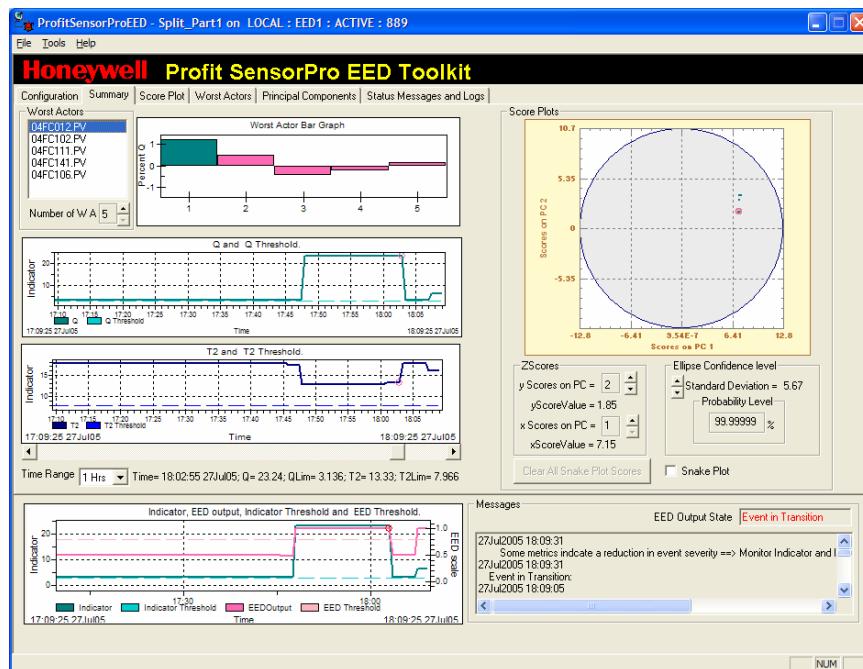


Figure 9 – Engineer's Interface for EED.

As the primary engineering interface, it provides all the necessary information to deploy, validate and commission an EED application. Each tab of the interface provides various levels of detail regarding the application chosen, ranging from a summary page to detailed information on the worst actors and principal components. The information is presented in a variety of formats including tabular, time-series trends, and bar charts. In addition, all control functions are managed from this interface, including activation of the applications, historization of data/results and management of bad or dropped input information.

The Operator's Interface

Each operating facility has different standards for design of the operator's interface to the distributed control system. In addition, each area of a processing plant may have different needs for visualization that are unique to the personnel and process that are involved. Driven by this and guided by ASM Consortium user interface guidelines, the EED operator's interface was designed to allow maximum flexibility with respect to integration with the existing operator workflow, as well as compatibility with a large variety of distributed control systems.

The EED operator's interface actually consists of a library of ActiveX objects. Each of these objects has commonality with pieces of the engineer's display, but can be inserted into standard operating schematics³ to augment rather than replace or add to existing operator visualizations. Prototype implementations have shown that this flexibility is necessary to the success of the installed applications.

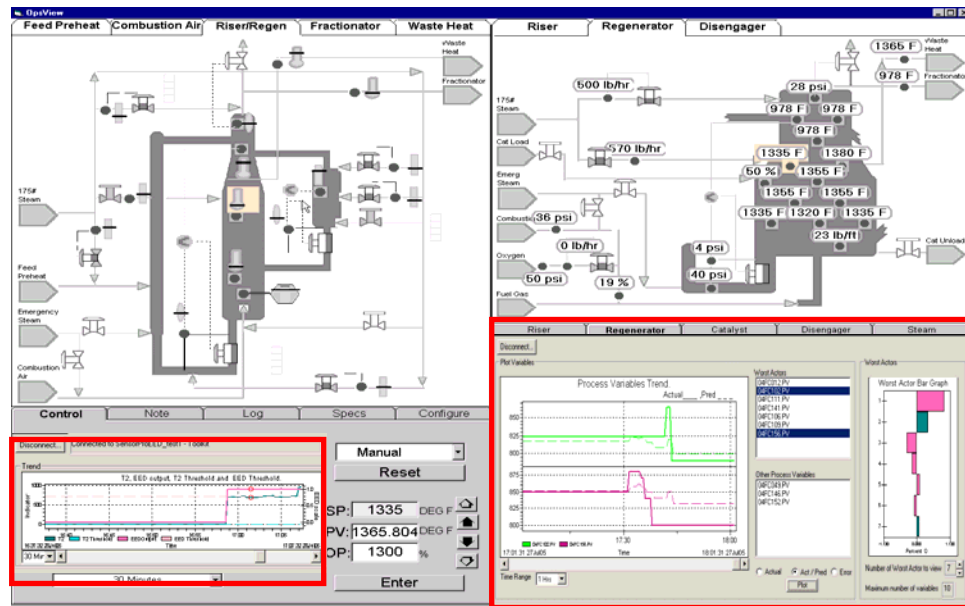


Figure 10 – Example deployment of an operator's interface augmented with EED visualization components

³ Requires ActiveX-compatible human machine interface on non-Honeywell DCS

Results that Deliver

The following results have been demonstrated during prototype ASM Consortium implementations:

- Detection of hydrate formation in a refrigerated overhead system on an ethylene plant resulted in incident avoidance benefits for a single event in excess of US \$100,000⁴
- Detection of failed sensors (reading, but significantly inaccurate) during design phase of project provided increased fidelity of control to the process
- Online multivariate sensor validation of key sensors provides efficient, high-fidelity detection of analyzer drift and helps prioritize maintenance of instrumentation.

The EED application results in a truly multivariate SPC solution that enables:

- Early detection of abnormal situations
- A reduction of the frequency and magnitude of plant incidents
- Reducing variation in both operations and product properties
- More consistent operation closer to optimum plant conditions
- Six Sigma analysis of process data
- Effective forensic analysis by facilitating a detailed understanding of normal and abnormal operating conditions from plant historical data
- Online applications to provide early detection of process abnormalities
- A more consistent engagement of process operators in proactive monitoring of the process.

⁴ Early Event Detection – Results From A Prototype Implementation. Michael B. Bell, Wendy K. Foslien, AICHE Spring Meeting – Ethylene Producer's Conference

For More Information

To learn more about how Honeywell's Early Event Detection Toolkit, visit www.honeywell.com/ps or contact your Honeywell account manager.

Automation & Control Solutions

Process Solutions

Honeywell

2500 W. Union Hills Dr.

Phoenix, AZ 85027

Tel: 877.466.3993 or 602.313.6665

www.honeywell.com/ps

WP-06-01-ENG
January 2006
© 2006 Honeywell International Inc.

Honeywell