Application of Advanced Pattern Recognition to Power Plant Condition Assessment

Marcus B. Caudill

Tri-State Generation & Transmission Assoc.

Ron D. Griebenow/Eric Hansen

Performance Consulting Services

Greg C. Alder

Halliburton NUS Corporation

Application of Advanced Pattern Recognition to Power Plant Condition Assessment

Marcus B. Caudill
Tri-State Generation & Transmission Association, Inc.
970-249-4501

Ron D. Griebenow, P.E.
Eric J. Hansen
Performance Consulting Services, Inc.
913-685-4432

Greg C. Alder
HALLIBURTON NUS Corporation
208-524-9350

Abstract

As electric utilities strive to become more competitive, efficient operation and maintenance of steam power cycles becomes a greater concern. As part of an aggressive thermal performance monitoring and improvement program, Tri-State Generation and Transmission Association has integrated advanced pattern recognition technology with more conventional heat balance analysis methods to monitor and diagnose plant cycle deficiencies.

Advanced pattern recognition (APR) technology is used for continuous monitoring of plant operation. It can quickly diagnose faulty instrumentation and supply accurate replacement values, providing confidence in the data used for daily operating and engineering decisions. Furthermore, it accurately identifies calibration drift, providing Tri-State with the information required to allocate manpower to only those instruments requiring maintenance. APR is also integrated into the data validation and analysis portions of annual thermal performance tests, providing increased confidence in test data and reducing data analysis time. This paper will present the advanced pattern recognition methodology applied at Tri-State and the benefits in optimizing plant operations and assessing power plant performance.

Introduction

In 1993, Tri-State's performance improvement group began investigating improved methods for validating performance test data and plant computer data. Previous data problems had increased the time required for precision performance tests, and incorrect plant computer data had caused major problems with our on-line performance monitoring program. In addition, comparisons of test grade instruments to plant instruments over a three year period yielded average deviations of 30 psi in main steam pressure, 5°F in main steam temperature and 5°F in hot reheat temperature. This translates to more than $800,000 annually in lost capacity and increased heat rate. This information led to the decision to pursue alternate methods of verifying instrument accuracy and the ultimate selection of advanced pattern recognition (APR) technology.

At the 1994 EPRI Heat Rate Improvement Conference in Baltimore, Maryland, we presented a paper outlining the installation of an advanced pattern recognition system for calibration monitoring at Tri-State's Craig Generating Station. Since that time, we have refined the calibration monitoring system installed on Craig Unit 3 for improved accuracy and expanded the use of APR to diagnosis of equipment deficiencies at our Nucla station.

Traditional Signal Validation

Signal validation consists of methodologies for distinguishing measurement failures from process faults and selecting which instrumentation signals to use in control and analysis functions. There are two legacy methods of signal validation which have been used extensively in process monitoring systems - comparisons among multiple values representing the same measurement and the use of reasonableness checking among multiple signals having a quantifiable relationship with the measurement being validated.

Comparison methods are based on the availability of at least two measurements (direct or derived) of a desired process state. These redundant measurements may then be used to make some judgment about the validity of the measurement signals. The simplest comparison methods involve the installation of two sensors at the same location for the same process state. These redundant sensors are compared with each other, and disagreement between them is considered indicative of failure of one of the measurements. However, if only two measurements are available, no decision can be made on which to accept. An unambiguous measurement quality determination requires more than two measurements. When at least three measurements are available for comparison, it is possible to make some logical choice of which to accept or reject, and to form a "best estimate" of the true value of the process state. When there are no physically redundant devices, it is possible to use analytically redundant measurements. An analytically redundant measurement can be found when there is a process model which can be used to derive a representative value of a directly measured state from measurements of other states. Once three or more representations of a particular measurement are available, there are methods which may be used to discriminate failed measurements and select the most representative "true" value of the measurement.

The other traditional method of signal validation, reasonableness checking, has been implemented much less frequently than direct comparison methods. This is a more qualitative or symptomatic method which is based on a comparison between the measurement value and certain reference values based on a qualitative estimate of what the measurement should be or on values associated with known failure modes of the measurement channel. In reasonableness checking, one takes advantage of the fact that deviations of measurements from expected values do not generally occur in an isolated fashion - that process faults usually result in the deviation of groups of measurements. Thus if a single measurement has an unexpected value or a value characteristic of measurement failure, it is most likely in error.

This form of reasonableness checking may also he applied to the derivative of a measurement. For most measurement situations, it is possible to estimate the upper bound of the derivative based on process knowledge or model studies, and then use that estimate as a limit for the allowable rate of change of the channel. This technique is particularly useful for measurement channels with relatively long time constants, such as temperature detectors in thermowells, or for measurements of process parameters which have long time constants themselves. A common form of applying limits on the derivative is to monitor two redundant measurements, immediately rejecting the one whose derivative assumes a high value.

The usefulness of the methods described above has been severely limited by the cost and complexity of their implementation. Installing redundant sensors is very expensive and tends to be limited to a small subset of the total number of sensors which might have significant impact on process reliability. Comparison methods also generally compare only two or three measurements to each other, or a measurement to a fixed limit. Consequently, this approach is not very robust and unlikely to cover a large subset of process measurements. In addition, comparison methods cannot be used for diagnosis of any type of process failure other than that of isolated sensors, and every comparison must be enumerated by the designer and appropriate limits assigned in advance.

There is, therefore, motivation for the implementation of more robust methodologies of signal validation. Methodologies which are able to use all measurement information and diagnose all sensors in the process, can be assembled and configured with minimal human intervention, and can adapt to changing process conditions encountered in normal operation, such as operation at various power levels.

Pattern Recognition Basics

Pattern recognition treats any process or system as a set of numeric data values. For example, a power plant is viewed as a list of pressure, temperature, and flow values rather than an assemblage of turbines, pipes, and heat exchangers. To understand how pattern recognition works with data from a power plant, consider a hypothetical system that has just two significant measurable parameters, P1 and P2. Data for this simple process can be collected through a series of tests and a graph can be constructed showing one parameter versus the other.

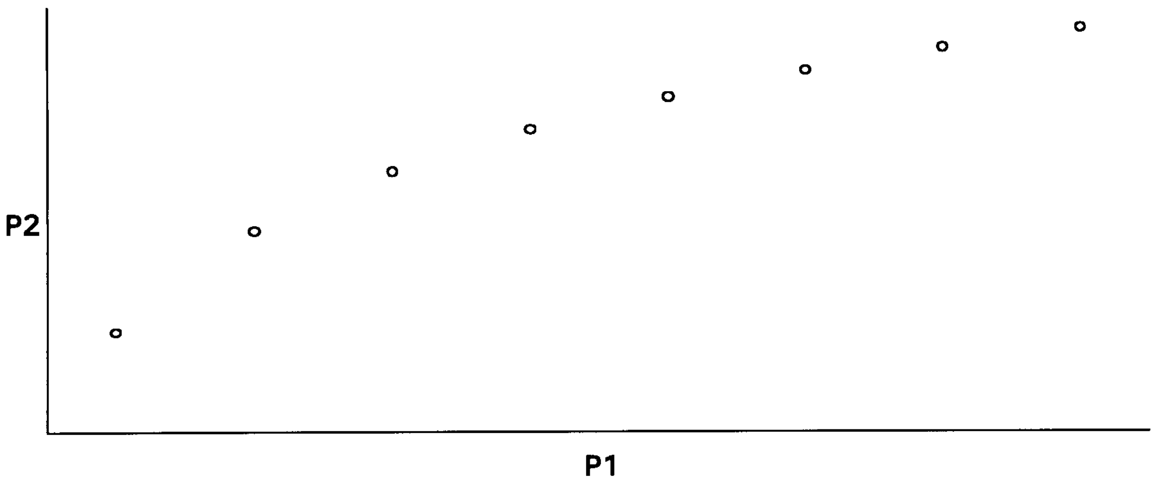

Figure 1. Plotting Process Parameters

The individual points in Figure 1 represent discrete operating states for a simple process. That is, each point represents a unique characteristic condition of the process. In order to accurately describe a single process state at any given time, the values of all the system parameters must be known (two in this case).

The process can then be generalized to estimate behavior at operating states for which no discrete data values have been collected by placing a continuous curve through the plotted points (Figure 2). This curve represents known normal operation and can be referred to as the "system performance curve". It should be noted that, when using pattern recognition techniques, these points are considered to be interrelated and are not separated into input and output (independent and dependent) categories.

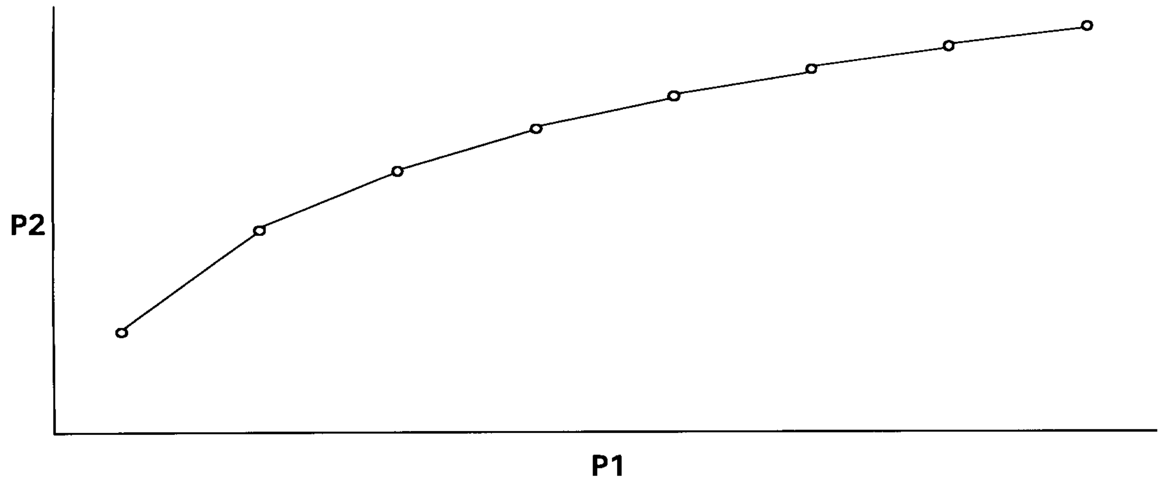

Figure 2. Generalizing Process Behavior

This modest procedure can lead to some very useful results. For example, if at a later time the value of one of the parameters is known, the value for the other can be readily estimated (Figure 3). In addition, the generalized plot could be used to determine whether or not the process is currently operating in a manner consistent with past operation (Figure 4).

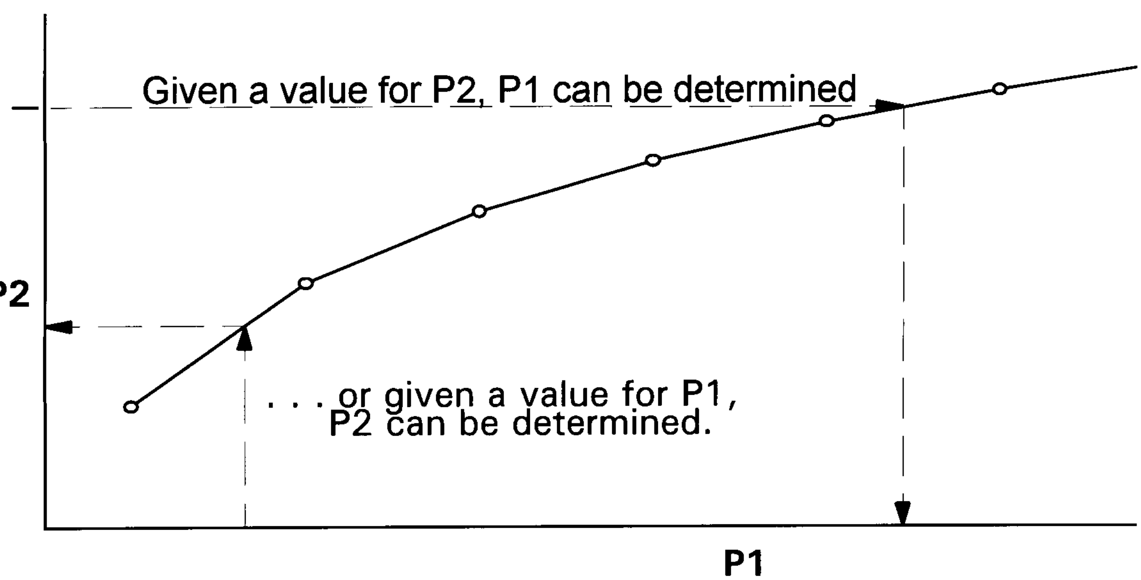

Figure 3. Determining Unknowns

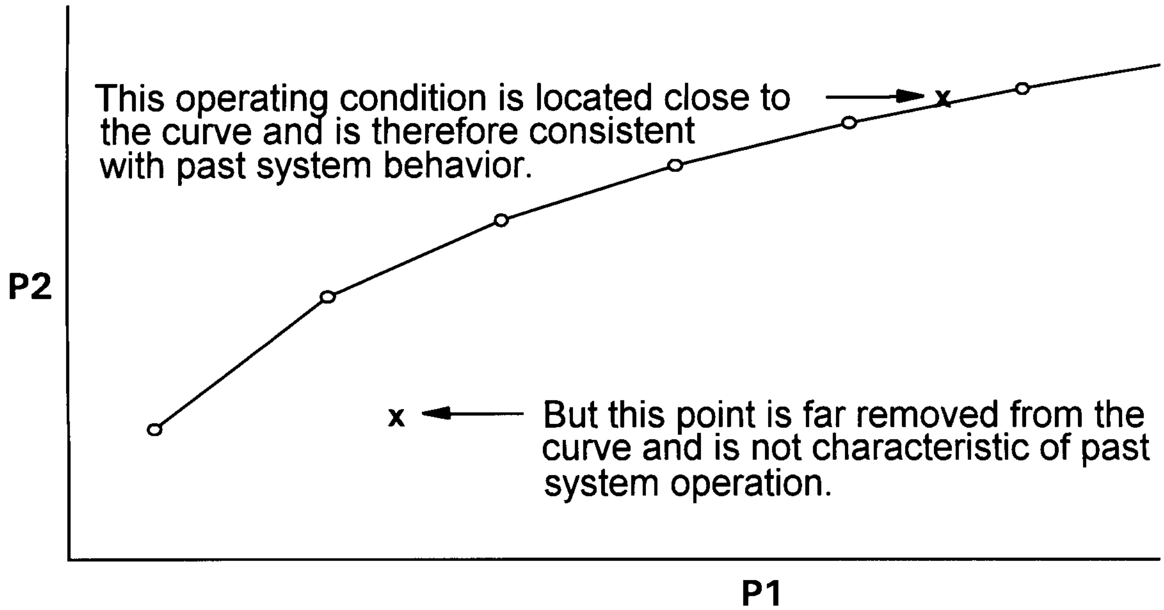

Figure 4. Assessing Process Operation

Although the examples above are presented graphically, similar results can be obtained numerically. However, most processes cannot be characterized sufficiently with just two parameters. For example, a complex process such as a power plant requires the measurement of many different parameters to describe plant conditions completely. Furthermore, most of the parameters associated with a complex process are interrelated and cannot be assessed simply as pairs of variables isolated from the rest of the system. So with most real-world processes it is necessary to examine many different parameters simultaneously rather than pairing them off in sets of two.

Instead of the simple case of P1 vs. P2 for two parameters, we now have a more complicated situation of P1 vs. P2 vs. P3 vs. P4 vs. . . . P*n* for *n* different parameters associated with a complex process. Although a graph showing all parameters plotted simultaneously for a complex process is not feasible, numerical methods analogous to the concepts presented graphically above can be employed. In other words, numerical techniques can be applied that utilize historical data values collected from a complex process to form the numerical equivalent of a "system performance surface".

Advanced Pattern Recognition Methodology

Advance Pattern Recognition (APR) works with system data that is captured and arranged in arrays called "data records". A data record is simply a snapshot of the system data for a single instant in time. The individual items which comprise the data record are called "points" and the values associated with these points are called "point values". Table 1 illustrates these concepts for a system consisting of five measured points.

Point ID	Point Description	Point Value
1	Main Steam Temperature	1005.2
2	Main Steam Pressure	1925.5
3	Condenser Pressure	2.05
4	Feedwater Flow	2716.7
5	Generator Power	435.8

Table 1

APR analysis is performed by quantifying the "similarity" between any two data records that are being compared for purposes of creating modeled estimates and other essential functions. Computed similarities are scalar values that range between zero and one. A similarity value of one indicates the two plant conditions being compared are identical (e.g., each temperature, pressure, and flow value is exactly the same in both data records). A similarity value of zero indicates that the two conditions are completely different from each other (e.g., plant conditions at full power compared to plant conditions at zero power).

Prior to analyzing a system, a number of snapshots of plant data are collected and stored in a file. These "reference data records" generally cover a range of system operating conditions and act as a knowledge base which is used to define the characteristics of the system.

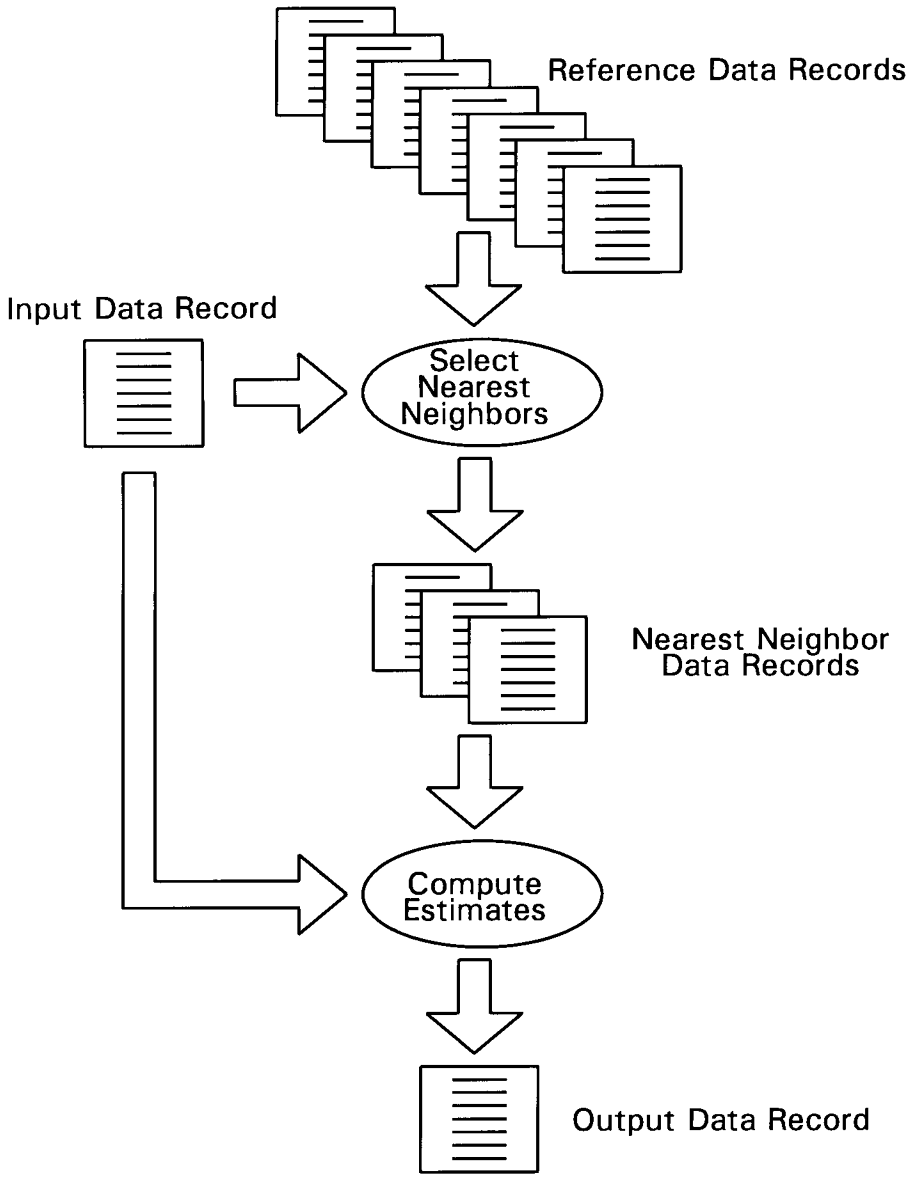

Figure 5. Schematic of the APR Process

After the reference data records have been collected and stored, a new snapshot of system data called the "input data record" is obtained for assessment. The input data record is first compared with each of the reference data records using the APR similarity operation. Several data records are selected from the reference set that have the highest similarity to the input record, while striving to bound each of the input point values. The data records selected from the larger reference set in this manner are referred to as "nearest neighbor" records. Once the nearest neighbor records have been selected, the similarity between all pairs of nearest neighbor records is computed, forming a "recognition matrix" of similarity values. This recognition matrix is then utilized in conjunction with the input data record to compute values for each of the monitored points. That is, a computed point value is calculated for each of the input points. The computed point values are collectively known as the "output data record". Figure 5 summarizes this procedure.

The output record is an extremely accurate representation of how the system should be behaving based both on past performance and on current operation. The calculated values are highly fault-tolerant because defective plant parameters in the input record do not markedly bias or affect the accuracy of the computations. In addition, for situations where one or more plant parameters are completely missing from the input record, APR will provide very accurate predictions for these parameters as well. Once the output record has been computed, it may be compared with the input record for further data manipulations. Again, it should be noted that the APR approach calculates an output point value for each and every input point (Figure 6). The significance of any differences between the input and output values is generally viewed in the context of the specific application (e.g., indication of a signal failure, calibration drift, abnormal system operation, etc.).

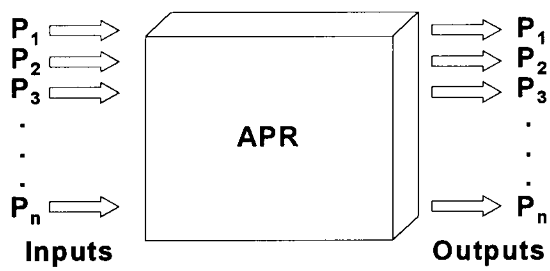

Figure 6. An Output for Every Input

Advanced Pattern Recognition Characteristics

The APR technology has several important characteristics:

Fault Tolerance: APR minimizes the effects of incorrect or missing plant measurements. This is due to the fact that it treats all monitored parameters as being interrelated. APR uses all parameters simultaneously to determine all of the predicted values so the impact of individual defects is minimized.

Localized Modeling: APR forms a local model for every new input data record. Because it does not operate using a single set of coefficients, complex non-linear systems can be modeled more accurately and with fewer examples than are required with other techniques. In addition, localized modeling automatically takes into account changes in a system as it progresses through different operating conditions.

No Iterative Training Phase: APR does not require an iterative training phase. Because no iterative training phase is involved, the reference data can be changed as often as desired allowing immediate adaptation to new information.

High Dimensionality: APR is unrestricted in the number of plant parameters that it can handle in a single model, however, extremely large systems should often be broken down into multiple, smaller models based on point relationships identified using APR tools. These smaller models enhance model accuracy and fault tolerance. In practical applications, it has been successfully applied to systems containing as many as 400 different parameters.

Models All Parameters: APR is unusual in that it models all variables simultaneously. This is essential for detecting faults in individual plant parameters and for assessing abnormalities in overall system operation. Most other techniques must separate system parameters into dependent and independent sets which often precludes the ability to pinpoint specific parameters that are deviating from normal operating behavior.

Repeatable: APR is 100% repeatable. There are no random settings of any initial conditions. If the analysis is executed twice using the same reference data set, it will produce identical results.

Deterministic: The APR algorithm is a well-behaved deterministic relationship. There are no iteration schemes requiring convergence criteria. The algorithm has no failure modes and will always produce an optimal solution given the available information. Further, it is unique in that it enables the determination of uncertainties associated with the final estimated values.

Tri-State Advanced Pattern Recognition Applications

At Tri-State, we have successfully applied APR to three different areas of our performance improvement program. These three areas are continuous calibration monitoring, plant health monitoring and diagnostics, and precision test data validation. Following are discussions of each application.

Continuous Calibration Monitoring

An APR calibration monitoring system was installed on Craig Unit 3 during late 1993 and early 1994. The system monitors 159 plant process points, including turbine, boiler, feeedwater and circulating water system data. These 159 points are the primary inputs to the Craig 3 performance monitoring system.

The initially collected data is the basis of the reference data set, and spans the range of plant operating conditions, from start-up to full load operation. Realize that in a fossil plant at a given load, there can often be many unique combinations of equipment in service and acceptable operating configurations, and accurate modeling of system conditions requires data for those unique conditions be bound by the reference data set.

The multi-dimensional system performance surface is generated by the advanced pattern recognition software. For the Craig system, this could be a 159-dimensional performance surface, but to improve the accuracy of all predictions, the model was divided into five sub-systems of closely coupled points. These sub-systems range in size from 18 to 39 points. Note that all five sub-systems are modeled using the same reference data set.

Case Studies

During the initial data collection and model set-up at Craig Station, APR identified problems with high economizer O2 level, a modulating valve on the steam coil air heater, a modulating high pressure feedwater heater level controller, and problems with the main condenser pressure calibration. These problems had apparently existed for some time, but were readily identified using APR technology.

Four months after the installation of the APR calibration monitoring system on Craig 3, we were performing upgrades to the Craig 3 performance monitoring system. Because of the sensitivity of heat balance calculations to data errors, we were using the APR system to screen and verify plant data. We noticed that auxiliary power, calculated as gross generator power less net, was unusually low. Historically, generation metering problems on this unit were associated with the gross generation metering, and our first instinct was to write a work request to have the gross generation transducers calibrated. However, we decided to first look at the APR predictions for all metered generation parameters. We discovered that APR-predicted gross generation matched very well with measured gross generation, but actual net generation was higher than the APR-predicted value. Further investigation revealed that auxiliary power was approximately 5 Mw lower than normal because several large compressors were out of service. In this instance, the APR system helped diagnose a problem before a work request was issued, avoiding wasted manpower for the recalibration of valid instruments.

The "C" feedwater heater on Craig 3 is the third highest pressure feedwater heater, with its extraction located at the intermediate pressure (IP) turbine outlet. Because the station extraction pressure transmitters are located at the feedwater heaters, when an extraction is out of service there is no pressure indication. Since this pressure is used to determine the IP turbine outlet condition and the low pressure (LP) turbine inlet condition, a missing value of "C" heater extraction pressure will cause a failure in turbine cycle heat balance calculations. When this condition occurred, APR provided an accurate replacement value for "C" extraction pressure. If integrated with the performance monitoring system, this would have prevented failure of the turbine cycle heat balance calculation.

During recent performance tests, a plant information computer input card failed, sending no data to our archival system for several process points. One of these points was net generation. The problem was discovered and corrected, but provided an opportunity for some additional validation of the APR technology. Figure 7 depicts the measured net generation and the APR predicted net generation. Note the excellent agreement between the measured and predicted values both before and after the input card failure. This excellent agreement increases confidence that the predicted value during the card failure is a realistic replacement value

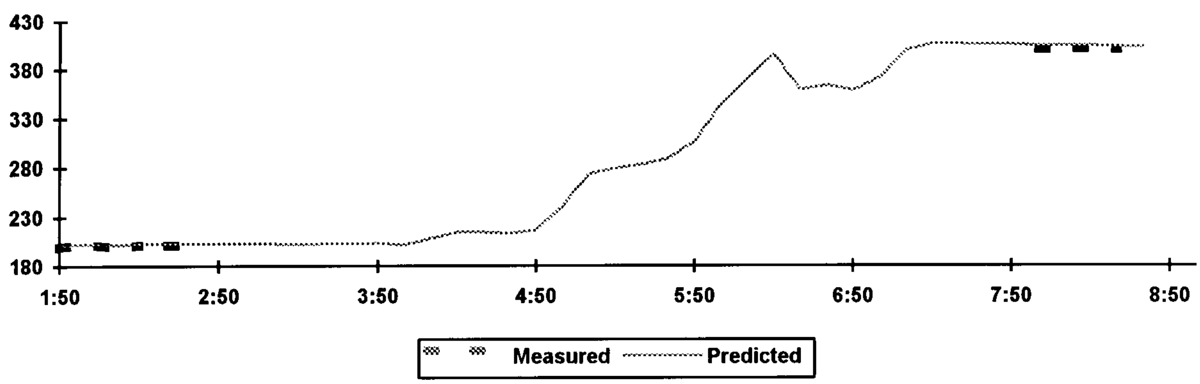

Figure 7. APR Net Generation Prediction

Tri-State has recently begun exploring alternative chemical treatments for improved cycle chemistry. One of these methods, oxygenated treatment, requires feedwater heater vents to be normally closed and then opened as needed to minimize build up of non-condensable gas in the feedwater heater. In preparation for an oxygenated treatment trial, we successfully validated the feedwater heater instrumentation using APR technology, and issued work requests on instruments that were out of compliance.

Plant Health Monitoring And Diagnosis

Nucla Unit 4 is a 110 Mw circulating fluidized bed (CFB) unit located in Nucla, Colorado. Twice annually, Tri-State is required to perform a Uniform Rating of Generating Equipment (URGE) test to demonstrate the output of the unit. During the Spring 1995 test we discovered that our fans were limiting our load. Since this was not usually the limiting factor, we began investigating the situation.

Plant data had been collected during previous URGE tests and during boiler acceptance tests and saved in an ASCII format. Unfortunately, APR had not been utilized previously at Nucla station, which necessitated development of an APR system model. Based on previous experience at Craig station, an initial list 168 data points was reduced to 143 by removing those which would have minimal correlation to the rest of the system (e.g. barometric pressure, ambient temperature, etc.) and those points which were known to be faulted at the time of the initial data collection. The remaining 143 points consisted of pressures, temperatures, flows, and generation measurements associated with the boiler and the four steam turbines.

Data collected during January and February 1994 were combined and used as a base reference library, then data from March 1994 was used for model verification. The APR predictions were in excellent agreement with the March 1994 measured data. Additional data from May 1994 was verified and samples were added to the reference data set to account for seasonal changes in unit performance. This provided a robust reference data set that spanned the full range of operating and ambient conditions. This reference data set was then used as the basis of the analysis on the Spring 1995 data. It should be noted that development and validation of the Nucla APR model required less than eight hours to complete.

Based on the APR analysis, several potential problems were eliminated. Generation metering, main steam flow, and fuel flow, in addition to a number of turbine cycle boundary conditions, were verified as being correct. However, primary air, secondary air and induced draft fan results all indicated increased air flow through the system. This narrowed the problem down from a large number of instruments to a more manageable number. After verifying instrumentation around the fans, we confirmed that air and gas flow were indeed higher than before and identified air in leakage as a problem. During a subsequent outage we found a number of leaking tubes in the tube and shell air heater and a damaged seal on the induced draft fan.

Precision Test Data Validation

During the winter of 1995-1996 we began integrating advanced pattern recognition into the Tri-State precision performance test program. Tri-State's test program is based on the intent of ASME PTC-6.1. Approximately 40 pressures, temperatures and flows that are critical to the accuracy of the test results are measured with test grade instrumentation and a stand-alone data acquisition system. The balance of the data are collected via a serial link from the plant information computer. This yields a data record approaching 300 process points, and another challenging data validation problem. Since test preparation and set-up is a two to three week process and the testing and analysis usually occupies an additional two weeks, we need to be sure that the information collected is valid. Previously, we relied solely on more traditional methods of data validation, such as instrument redundancy, test redundancy, and manual comparison of test results to historical test data. While these traditional methods are still employed, we have added APR validation to provide additional confidence to our test data.

The APR test data validation is performed in the same manner as it is for continuous calibration monitoring. A reference data set is compiled from previous test data, then current data is analyzed using APR. Any deviations are resolved before the initial tests are run. Data can be readily validated before tests each day, or at any point during the test. This provides almost real time capabilities for test data validation and minimizes "bad" tests.

Conclusion

Advanced pattern recognition technology offers a more accurate and reliable method for validating plant data and detecting calibration drift that traditional validation techniques. In addition, it does not require complex, high-maintenance process models to relate process parameters. Our analysis indicates that APR-based data validation will result in greatly increased data confidence, providing for improved operating and maintenance efficiencies. At Tri-State, advanced pattern recognition applications have provided replacement values for failed process measurements, helped in the validation of performance test data and aided in the diagnosis of shifts in unit operating conditions. Based on these successes, APR will continue to be applied to data validation and plant health monitoring at Tri-State.

References

1. "Applied Pattern Recognition for Plant Monitoring and Data Validation"; R.D. Griebenow, E.J. Hansen, A.L. Sudduth; 6th Annual EPRI-ISA POWID; La Jolla, CA; 1995.

2. "Similarity Based Regression: Applied Advanced Pattern Recognition for Power Plant Analysis"; E.J. Hansen, M.B. Caudill; 1994 EPRI-ASME Heat Rate Improvement Conference; Baltimore, Maryland; 1994.