How FALCONEER Handles Process and Sensor Noise, Missing Data and Poorly Defined Models

Daniel Chester, PhD, Assoc. Chair, Computer and Information Sciences, University of Delaware
Richard Fickelscherer, PhD, PE, Principal, FALCONEER Technologies

Operations understands plant data is sometimes imperfect and contains noise but that operation within a normal range is usually good enough to provide required performance – from a safety, cost, regulatory or production basis. Process operation based on this imperfect information is still better than operating blindly and this same information can be used to further troubleshoot and optimize plant performance. Poorly controlled (large process noise) operations or inaccurate (large sensor noise) or mis-calibrated or improperly installed sensors (large offsets) make it difficult, but not impossible to find problems, faults, failures, and other abnormal conditions – it just may take longer or more effort than a well controlled process with properly calibrated sensors. THE PROCESS CONSEQUENTLY ALMOST ALWAYS TELLS ON ITSELF.

Process data cannot be considered an asset unless correctly presented, analyzed, and converted into information[1]. Using models to do this analysis and conversion and then perform sensor validation and proactive fault analysis (SV&PFA) based on those results in real time proves the assertion that "Models are the means by which data can be converted to meaningful information"[2]. These models are based on a fundamental understanding of normal operating behavior of the given process system. They thus generate an unimpeachable source of knowledge for logically inferring conclusions about the process being modeled. Automatically performing this inference after each update of process sensor data allows such fault analyzers to perform "intelligent supervision" of the daily operations of their associated process systems. These models of normal process operation are thus paramount in the process knowledge for describing normal behavior of the modeled system. These are the types of models FALCONEER uses for its real-time evaluation of current process operation. In summary, we just require such models alone to be able to perform competent SV&PFA, converting the very complicated problem of fault analysis into a simpler problem of process modeling.

Computer Aided Process Engineering (CAPE) design or simulation tools like Unisim, Hysys, Aspen, ICAS, etc. use typical or recent or average value of plant conditions and parameters to evaluate snapshots of current operation compared to design or target. While simulations and

---

[1] Kennedy, J. P., "Data Treatment," Foundations of Computer-Aided Process Operations II, ed. by D. W .T. Rippin, J. C. Hale and J. F. Davis, Austin, TX, CACHE, Inc., 1994, pp. 21-44.

[2] Kramer, M. A. and R. S. H. Mah, "Model-Based Monitoring," Foundations of Computer-Aided Process Operations II, ed. by D. W .T. Rippin, J. C. Hale and J. F. Davis, Austin, TX, CACHE, Inc., 1994, pp. 45-68.

FALCONEER are both based on first principle engineering equations and theory, FALCONEER uses the equations and models in a different, more robust way for real-time evaluation.

Our solution has been designed and developed to handle processes and systems with noisy data, unmeasured variables, and few and/or poorly defined available Primary models (Primary models are defined below). Our patented methodology provides unique, useful and innovative sensor validation and proactive fault analysis (SV&PFA) in these situations (United States Patent No.: US 7,451,003 B2, "METHOD AND SYSTEM OF MONITORING, SENSOR VALIDATION AND PROACTIVE FAULT ANALYSIS").

Black box data analysis and even advanced statistical methods applied to mass and energy balance models cannot do this with both good and noisy or bad data. Those methods require the noisy or bad data to first be scrubbed or reconciled. Our approach concurrently validates and estimates the real-time data that are linear terms in the models as well as using it for monitoring and detecting faults, failures and abnormal conditions.

All Primary Models will be written in the following format:

0 = Input Terms - Output Terms - Accumulation Terms

As many unique Primary linearly independent process models as possible should be derived for describing the normal operating behavior of the target process system. These models should be based upon the most fundamental understanding of normal process behavior known and should be limited only by the specific type and frequency of process data being collected. The resulting set of well-formulated models should constitute a highly detailed description of the target process system's normal operating behavior.

All Primary models describing normal process operation can be characterized as functions of the following quantities:

$$0 = f_i \text{ (modeling assumption variables of model } \mathbf{i}\text{)}$$

where modeling assumption variables are specific sensor measurements ***or*** standard or extreme values of specific parameters.

These models may be simple or complex, steady state or dynamic, linear or non-linear. When the models are evaluated with actual sensor data and standard or extreme values of unmeasured parameters, a residual results (i.e., $\varepsilon(\mathbf{i})$ below) which is just a function of process and sensor noise and any currently occurring modeling assumption variable deviations, e.g.,

$$\varepsilon(\mathbf{i}) = f_i \text{ (noisy and deviated values of modeling assumption variables of model } \mathbf{i}\text{)}$$

FALCONEER Technologies LLC
4455 Transit Road Suite 3C, Williamsville, NY 14221
Phone: 716 204-0273 Fax: 716 558-7250
Email: falconeer@falconeertech.com

The accuracy of a process model is determined by evaluating that model with process data that is collected over the entire range of normal process operating conditions in which the model is expected to be satisfied (its residual is zero or close to zero and within its normal standard deviation range). This allows a representative distribution of that model's residual to be derived as a function of the normal or acceptable process state (e.g., production level, etc.), process noise and sensor noise. This accuracy directly establishes the diagnostic sensitivity of the fault analyzer for each of the possible fault situations. Note that a possible fault situation could be that one of the sensor readings used in that model is inaccurate or bad.

If the **residual** (i.e., $\varepsilon(\mathbf{i})$) of a model is **significantly high or low** (significantly higher or lower than its normal value, for example more than 3 standard deviations for a Gaussian distribution), then it can be inferred that at least one or more of the possible modeling assumption variables is deviating from normal or expected behavior and is likely indicating that a fault, failure, or abnormal condition currently exists in the process operation. The model standard deviation therefore incorporates the normal process & sensor noise automatically in its analysis.

If the Primary model as formulated is an inaccurate representation of normal process operation (perhaps due to high process and sensor noise levels or inaccurate values of sensors, unmeasured variables or other parameters, etc.) then a skewed normal model residual will occur with normal operation data. Our method accounts for these situations by always adjusting residuals by subtracting the historical average normal residual from the current model residual each time it is calculated to center the adjusted residual on zero. This adjusted residual is used for all subsequent SV&PFA. This is thus how we handle inaccurate Primary models.

The fault analyzer's diagnostic sensitivity can be further improved by putting a first order lag filter on the various Primary model residuals used by the program. This directly helps to mitigate the associated sensor and process noise included in the calculation. It however induces a time lag into the analysis as the filtered residual is flushed with current data.

To continuously perform real-time fault analysis, the system evaluates the set of all of the linearly independent Primary models' adjusted residuals, evaluates all residuals resulting from exhaustively combining pairs of those Primary models to eliminate all linear terms common to each pair of Primary models (the so-called set of all linearly dependent Secondary models), determines the diagnostic significance of all these different residuals' Certainty Factors (i.e., low, satisfied, or high), and then interprets the resulting patterns of violated and satisfied models using our patented FUZZY Logic diagnostic rules (for Single and Multiple Fault situations) for each analysis period. Therefore, our patented intelligent fuzzy logic-based diagnostic approach allows this data/information, however noisy or not, to uncover the more subtle or difficult to spot issues using the combination of first principle primary models (mass & energy balance, other acceptable plant models/equations) along with additional secondary models that are formed via combinations of the primary models.

Concurrently for each analysis period, the sensitivity of a Primary model with respect to each of its sensors or assumption variables that are formulated as a linear term in the model is also determined. This sensitivity analysis is done by evaluating the partial derivatives of a Primary model with respect to each of those linear sensors or associated assumption variables.

We provide dashboard windows to monitor just how FALCONEER is "thinking" (i.e. we display the results of these calculations for every analysis interval). All of these linear sensor and unmeasured variable sensitivities are continuously calculated and updated and can be displayed and monitored by FALCONEER to assist with the tuning and adjustments for specific models and/or specific sensors and parameters. Likewise, the resultant certainties for validation of each sensor and/or the certainties for identifying a fault or abnormal condition are displayed and monitored by FALCONEER to assist with the tuning and adjustments for specific models and specific sensors and parameters. This also allows a continuously updated display of slowly developing incipient fault situations.

The figure below summarizes our patented model-based SV&PFA solution.

SENSOR VALIDATION & PROACTIVE FAULT ANALYSIS (SVPFA) MODULE