CONTAC INGENIEROS LTDA.
INGENIERIA-SISTEMAS-AUTOMATIZACION

CONTAC Engineers Ltda.
Santiago, Chile

Pattern Analysis for Process Deviations Early Alert and Cause-Effect Relationships Evaluation

PRESENTATION OUTLINE

- BACKGROUND: NEEDS IN PROCESS SUPERVISION. THE PI-SCAN APPROACH
- OFF-LINE ANALYSIS
- ON-LINE APPLICATIONS
- SCAN PROJECT STATUS AND FUTURE DEVELOPMENTS

BACKGROUND

➢ The Needs

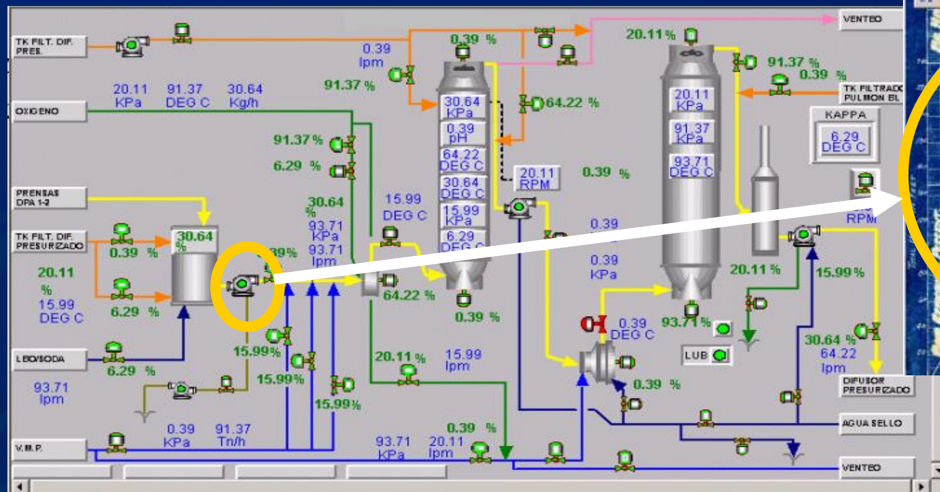
Two major PI users, a Copper Mining Operation and a Gas Plant:

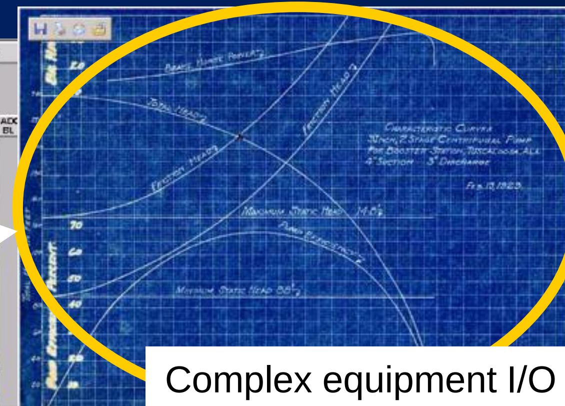
- Use of historical information for the characterization of plant behavior patterns and the analysis of cause-effect relationships
- Use of pattern parameters for:
  - Prediction of quality indicators and operational performance
  - Estimation of process variables in time (especially those which are hard to measure)
  - Early detection of abnormal operational conditions and equipment failures

BACKGROUND, THE NEED

Input Material Characteristics

Relationships between Processes





Complex equipment I/O relationships

Hard Links:
Physical Interconnections between equipment

Soft Links:
Process Control Loops and Operating procedures

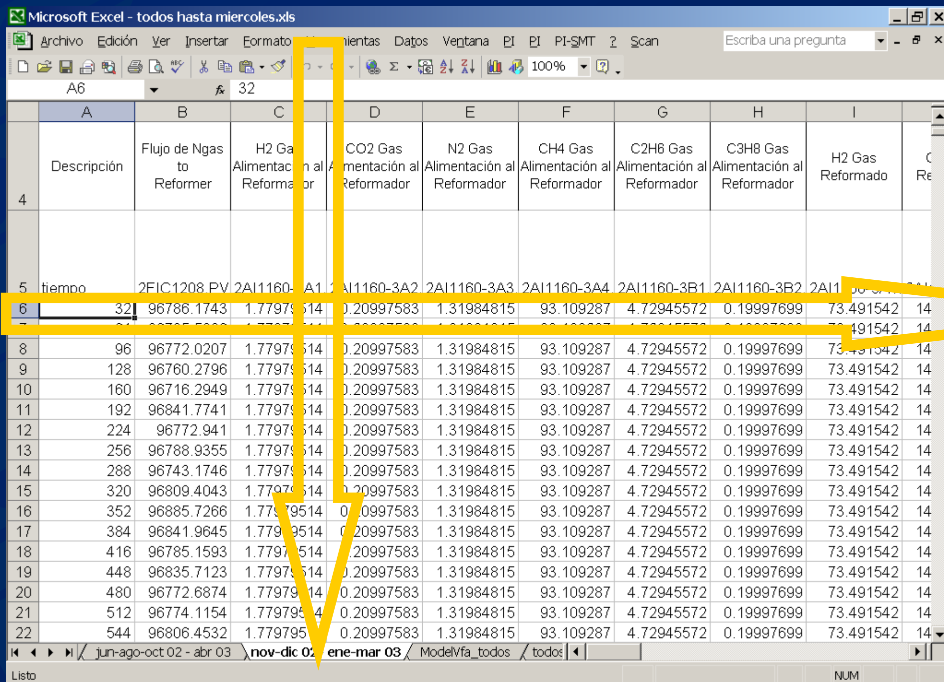
- Large number of process variables
- Complex cause-effect relationships
- High co-linearity and redundancy
- Different (time based) Operation Modes: drift, noise, start-up, set point changes, disturbances, etc.

GOAL: To characterize structural relationships between process variables and time series frames.

CONTAC

BACKGROUND, THE NEED

- Real Processes are both, Multivariable and Multi-stage:



	A	B	C	D	E	F	G	H	I	J
4	Descripción	Flujo de Ngas to Reformer	H2 Ga[illegible] Alimentaci[illegible]n al Reforma[illegible]or	CO2 Gas [illegible]mentación al [illegible]eformador	N2 Gas Alimentación al Reformador	CH4 Gas Alimentación al Reformador	C2H6 Gas Alimentación al Reformador	C3H8 Gas Alimentación al Reformador	H2 Gas Reformado	C[illegible] Re[illegible]
5	tiempo	2FIC1208.PV	2AI1160-[illegible]A1	[illegible]AI1160-3A2	2AI1160-3A3	2AI1160-3A4	2AI1160-3B1	2AI1160-3B2	2AI1[illegible]	[illegible]
6	32	96786.1743	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
7	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]491542	14
8	96	96772.0207	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73[illegible]	14
9	128	96760.2796	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
10	160	96716.2949	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
11	192	96841.7741	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
12	224	96772.941	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
13	256	96788.9355	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
14	288	96743.1746	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
15	320	96809.4043	1.77[illegible]514	[illegible]20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
16	352	96885.7266	1.77[illegible]9514	0[illegible]0997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
17	384	96841.9645	1.779[illegible]9514	[illegible]20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
18	416	96785.1593	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
19	448	96835.7123	1.7797[illegible]514	[illegible].20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
20	480	96772.6874	1.77979[illegible]14	0.20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
21	512	96774.1154	1.779795[illegible]4	0.20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14
22	544	96806.4532	1.779795[illegible]	0.20997583	1.31984815	93.109287	4.72945572	0.19997699	73.491542	14

Multivariable:

- Co-linearity
- Redundancy
- Cause/Effect

To be able to capture the Plant Structure

Multi-stage:

- Drift
- Noise
- Seasonal changes
- Control Settings, etc.

To be able to recognize deterministic & stochastic frames and events

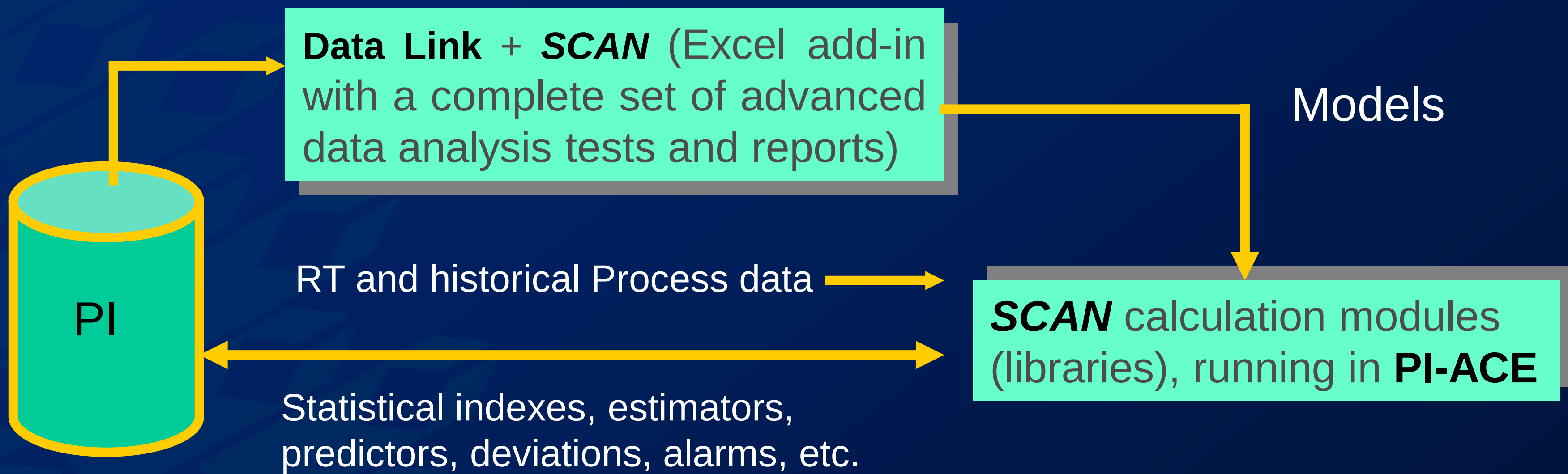
The (SCAN-PI) Goal

+

Off-line analysis &
On-line deployment

BACKGROUND

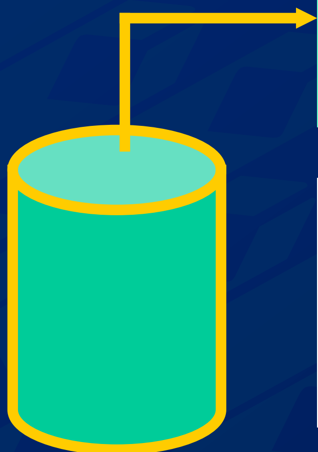
➢ The PI-SCAN Approach:



OFF LINE	**ON LINE**
Use of historical information for the analysis of cause-effect relationships and plant behavior patterns	Use of pattern parameters for early detection of abnormal Operational Conditions, quality predictors and Estimators for process variables.

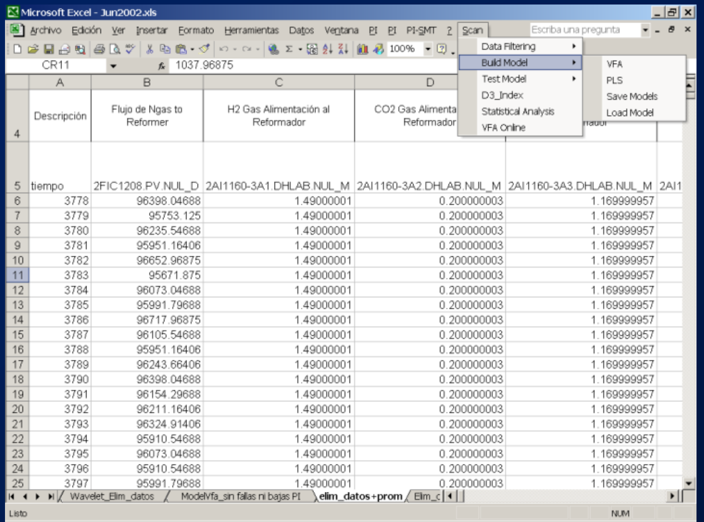
SCAN-PI, Off-line Analysis

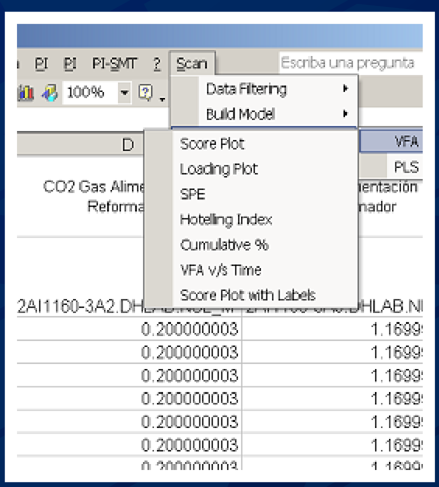
Process Data Analysis: Off Line Analysis

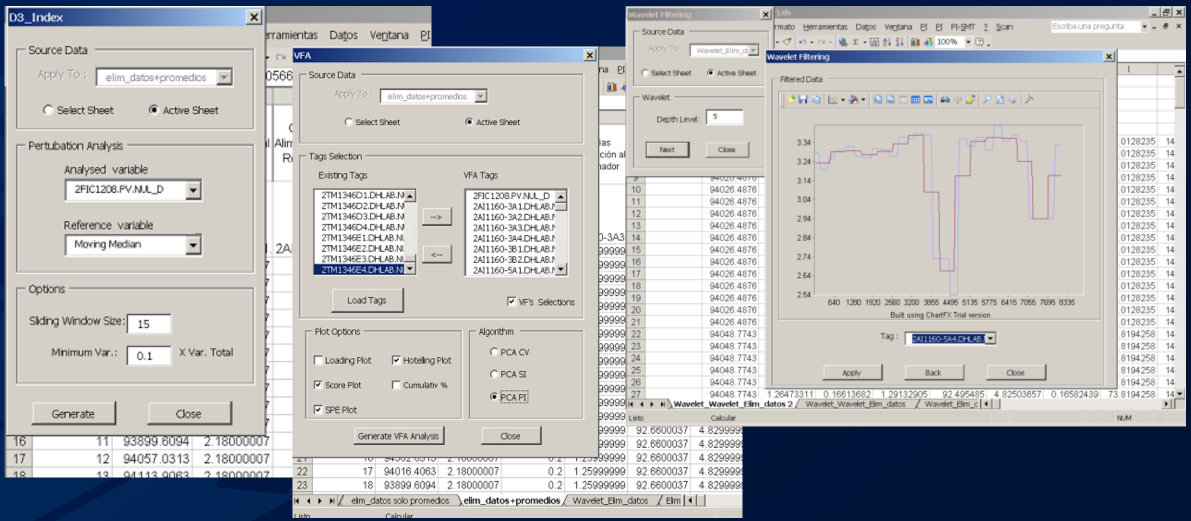


Data Link+*SCAN* (Excel add-in with a complete set of advanced data analysis tests)

Use of historical information for the analysis of cause-effect relationships and plant behavior patterns







Process Data Analysis: Off-Line Analysis

Data Set

Acquirement of data time series and related information: Representing normal operation, failures, seasonal changes, operation procedures, etc.

Data Set Conditioning

Checking for out of range, missing data

Filtering, Averaging

Generation of new (calculated) variables:
- Incorporation of Delayed Variables into the analysis
- Incorporation of phenomenological knowledge

Data Set Selection

Data Clustering, based on the operation characteristics which are being studied

Select reference models and patterns

Test application, definition of test sequences, selection of model parameters and variables, training refining

- Model and pattern definition (or “model training”) is an iterative process
- Data Link + Excel + SCAN: the training environment

ANALYTICAL TESTS (some of them...)

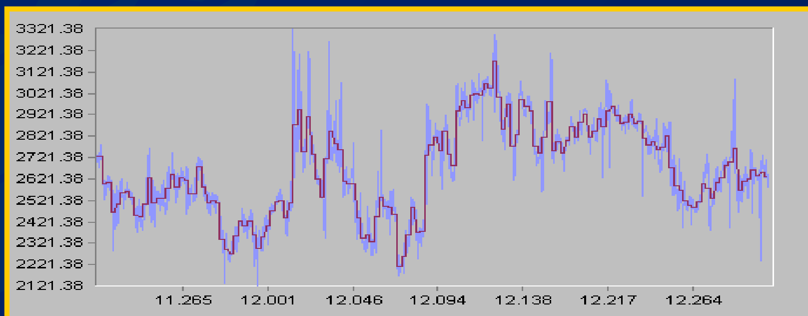
Wavelet Analysis

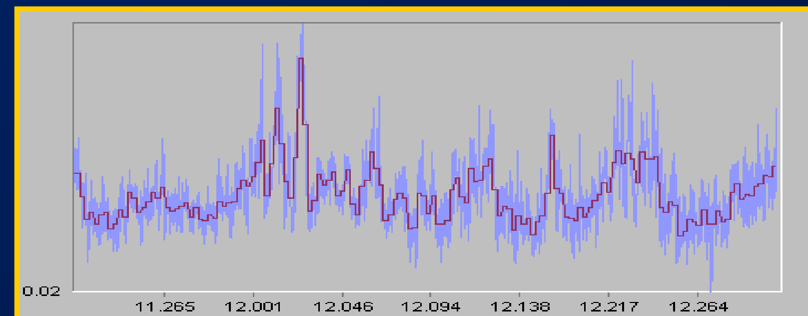
Linear transformation which can be used to obtain the decomposition of given signal into different **time-based scales** (or "frames", or "shapes").

- Uses:
  - Time series data reduction (it can compress redundant data)
  - Outlier detection in sampled data
  - Missing data reconstruction
  - Detail-Based Signal Analysis: Selective filtering and analysis of time series shapes
    - \- Noise Filtering
    - \- Detection of seasonal shapes
    - \- Detection of signal drifts (weariness?, need of calibration?, other?)

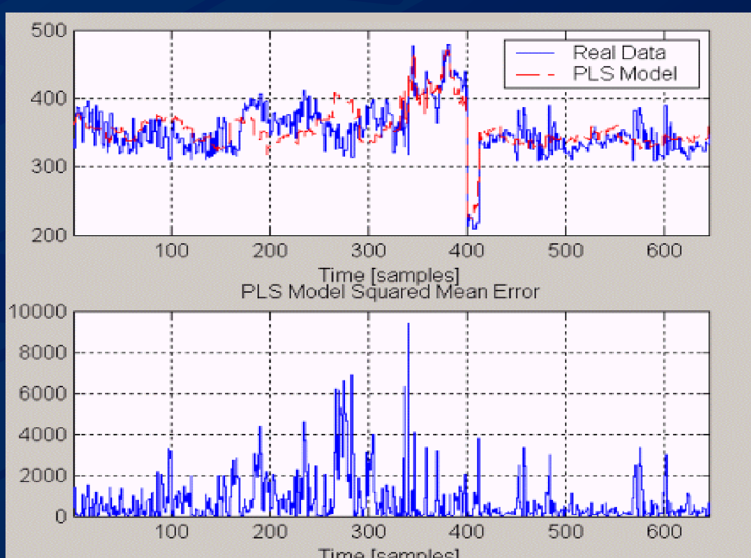
WAVELETS: EXAMPLES AND DEMONSTRATIONS

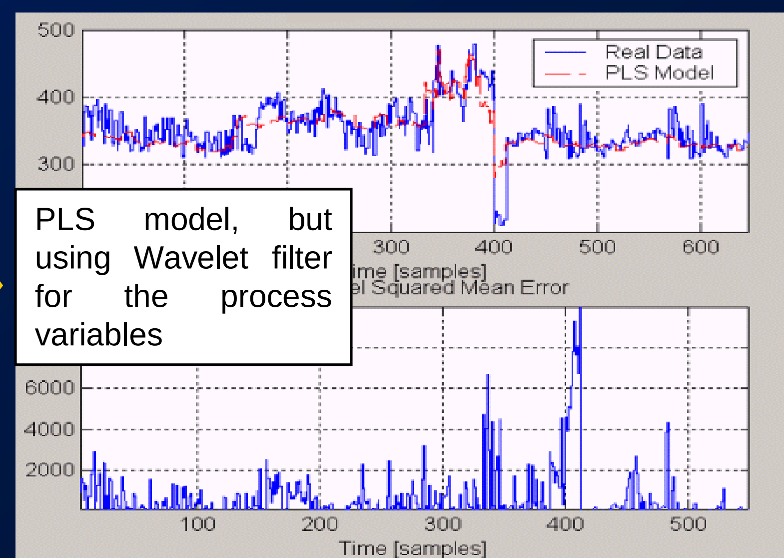
- Noise filtering and Outlier elimination as a prior tool before any multivariate analysis. Signal trends are represented in a better way and erroneous variability sources are avoided.





- Model identification improvements can be obtained (better RMS error):





PLS model, but using Wavelet filter for the process variables

ANALYTICAL TESTS (some of them...)

D3 INDEX

Dynamic Disturbance Detection Index: Statistical Univariate tool which is able to detect abnormal dynamic behavior, by analyzing the residual between a process signal and its reference over a sliding time-window.

Two possible analysis modes:

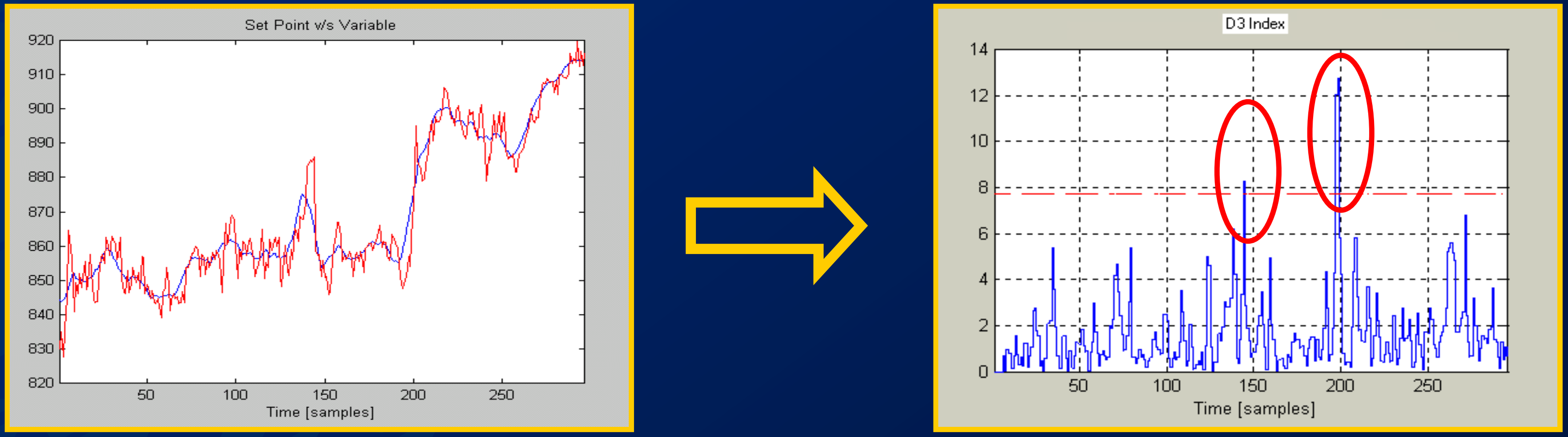
- **Univariate Analysis**: Reference is the signal moving mean or the signal moving median.
- **Model Based Analysis**: Reference is a Control loop set point, another process variable or a desired dynamic behavior (1^{st} or 2^{nd} order transfer function, other)

- Uses: Control Loop Assessment
  Process Variable Upset detection
  Changes in Noise/Signal ratio
  External disturbance detection

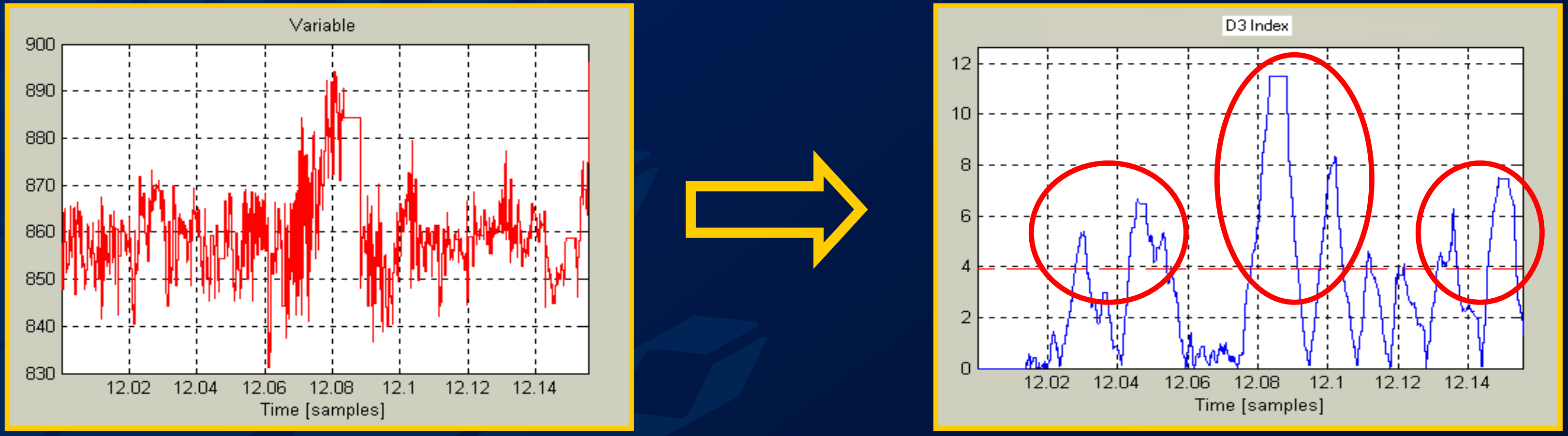
SCAN -PI

D3: Examples and Demonstrations

1) Control Loop Assessment:



2) Detection of disturbances in process variables, regardless noisy behavior and also updating changes in their dynamics.



ANALYTICAL TESTS (some of them...)

VFA

Variability Factor Analysis: Process Structure Representation

Being "**X**" the process variables (X_1, X_2, X_3,...., X_n), **VFA** determines pseudo variables "**V**" (V_1,V_2.....V_m, **m<<n**), such that $\mathbf{V_i}=f_i(\mathbf{X})$, where "**V**" represents the "directions" of **Maximum Variance** in Process Data.

- The resulting Pseudo-Variables (**V**) are:
  - Non correlated
  - Non redundant
  - They contain Independent information

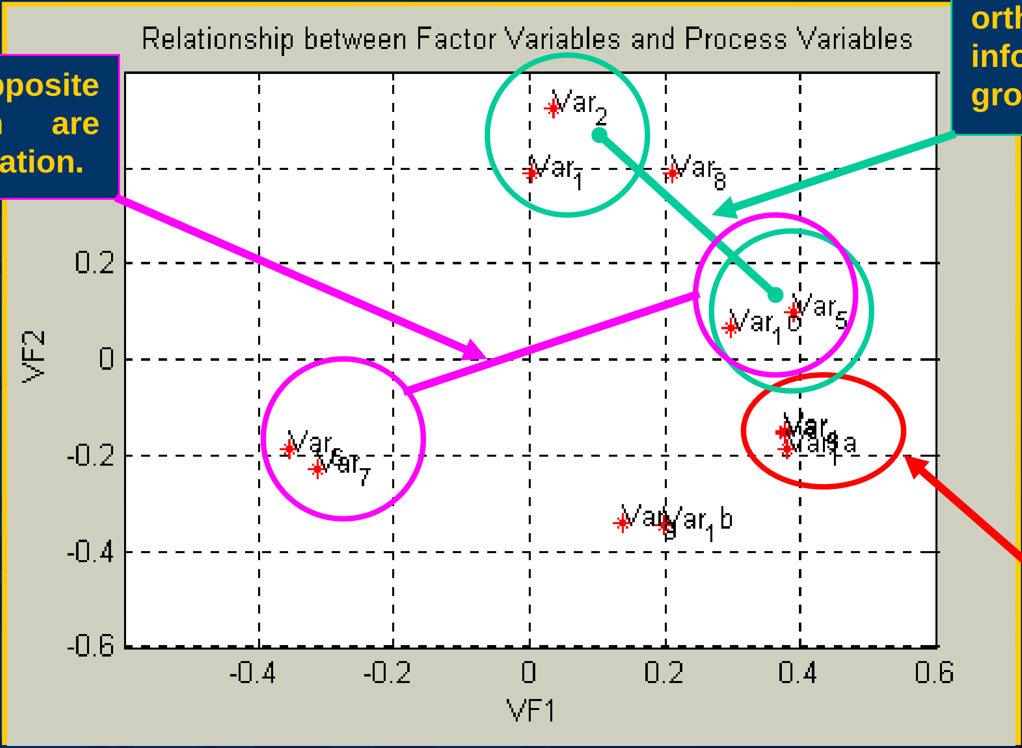
- VF can be characterized, then: process behavior can be also characterized
- Pattern (**structure**) can be associated with certain Process Behavior
- Structure changes are (often) detected before individual changes in process variables, allowing early detection of process changes: upsets, quality alerts, outside disturbances, etc.

VFA: EXAMPLES AND DEMONSTRATIONS

1) The Loading Plot is able to show simultaneously information about correlation, signal independence and influence in process variability for every process variable included in the analysis.

When two groups of process variables appear to be closer to orthogonal VF axes, then the information contained in each group is complementary.

Two groups located in opposite sides of the origin are indicating negative correlation.

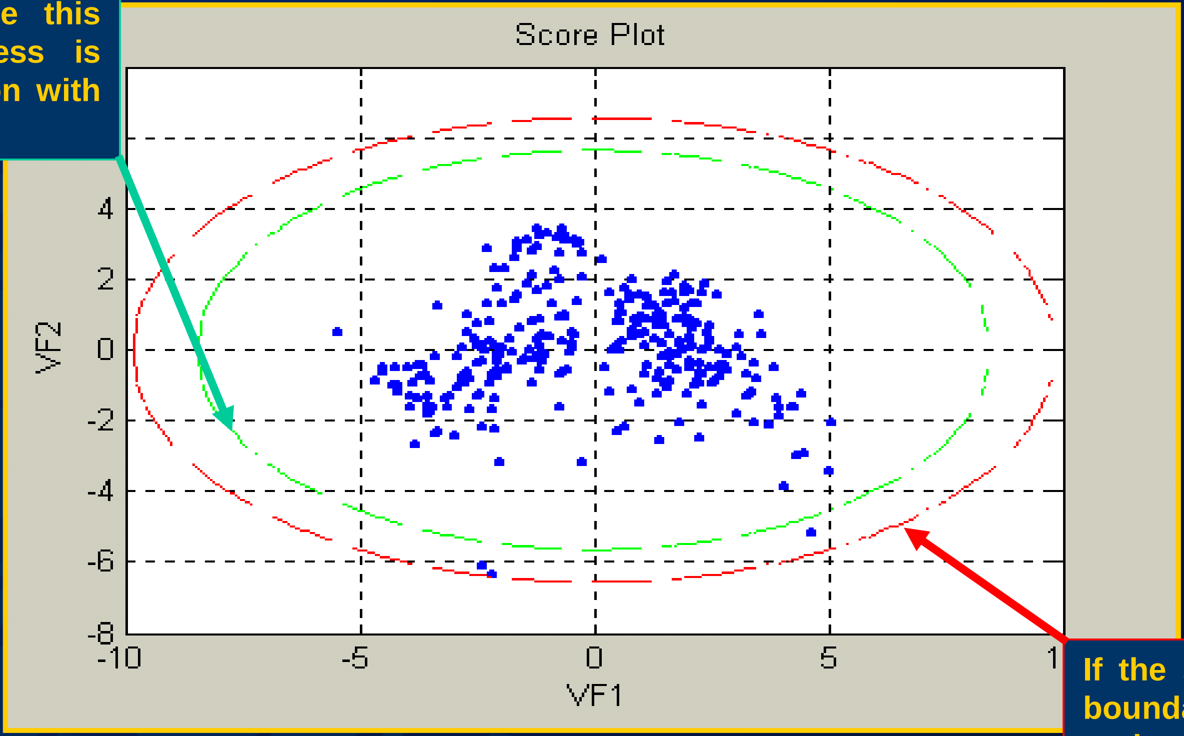


When process variables appear to be together, a positive correlation it is been indicated.

VFA: Examples and Demonstrations

2) The Score Plot defines the statistical boundaries for desired (or normal) operation of the entire process. Thus, it is possible to define the membership of the present behavior to any desired operational condition.

If the score is inside this boundary, then process is under desired operation with a 95% confidence

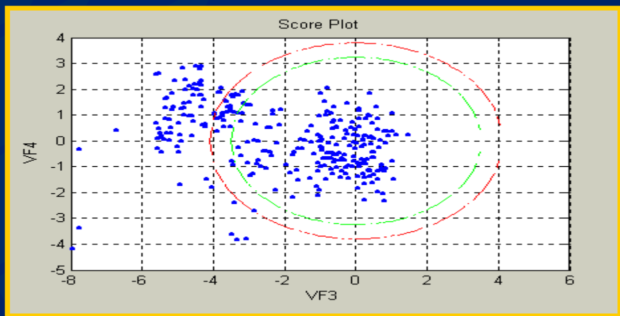


If the score is inside this boundary, then process is under desired operation with a 99% confidence

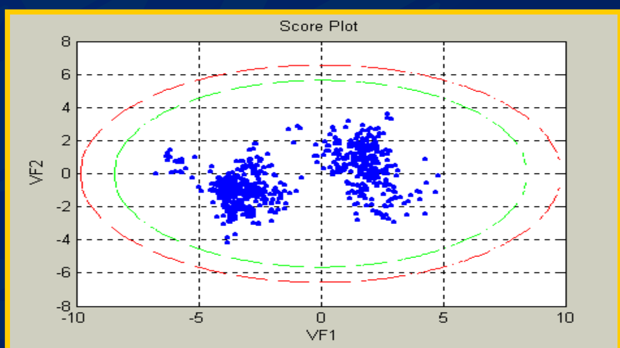
SCAN-PI

CONTAC

VFA: Examples and Demonstrations

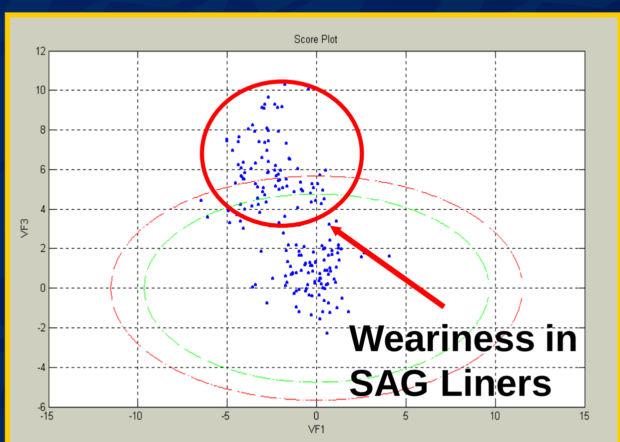


3) Therefore, it is also very easy to detect an abnormal condition, when the data scores begin to appear outside the confidence elliptical boundaries.



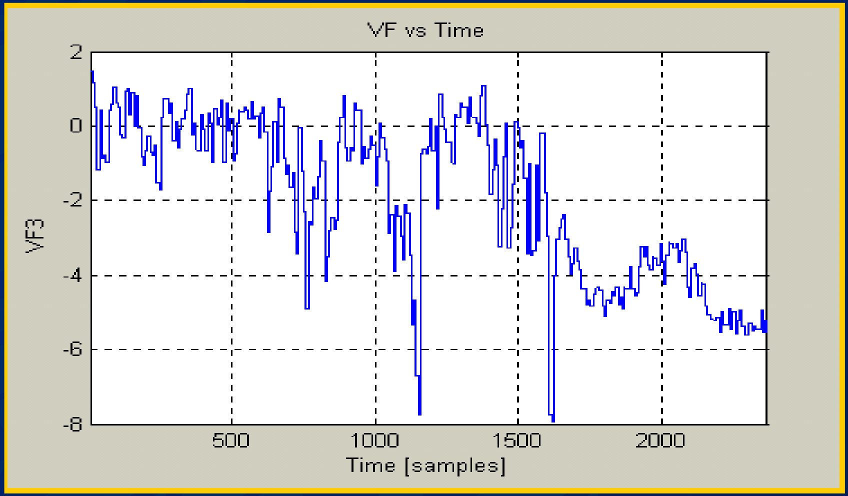
4) Process operation can vary according to different factors, but not necessarily that variation is considered as a "fault".

VFA allows to define different valid operation ("clusters") points inside the ellipsis, and to characterize them according to input sources, etc.

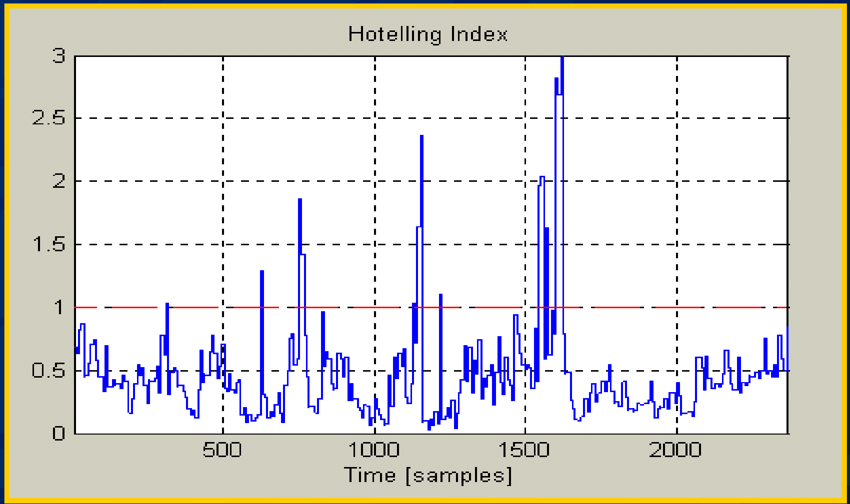


5) "Excursions" (changes) in Operation Points can be also characterized, and used for latter identification of "causes".

VFA: EXAMPLES AND DEMONSTRATIONS



6) The trend of a Variability Factor (VF) aggregates the information of several process variables. It also represents an independent, non correlated piece of information.



7) T^2 (or “Hottelling”) Index aggregates the variability information for the whole plant in just one index.

Since Structures are more sensitive than individual variables, T^2 can be used for process abnormalities early alert.

ANALYTICAL TESTS (some of them...)

PLS

Projections

Cause-Effect (Structural) Representation Models

Being "**X**" the process variables (X_1, X_2, X_3, ...X_n) selected as "cause" or process drivers

Being "**Y**" the "quality" variables (Y_1, Y_2,.........) selected as the effect variables

PLS Tools

Reduction to underlying structure

Reduction to underlying structure

Causes that most accurately describes the **effect** in the (variability of the) quality variables

SCAN-PI

PLS

Projections

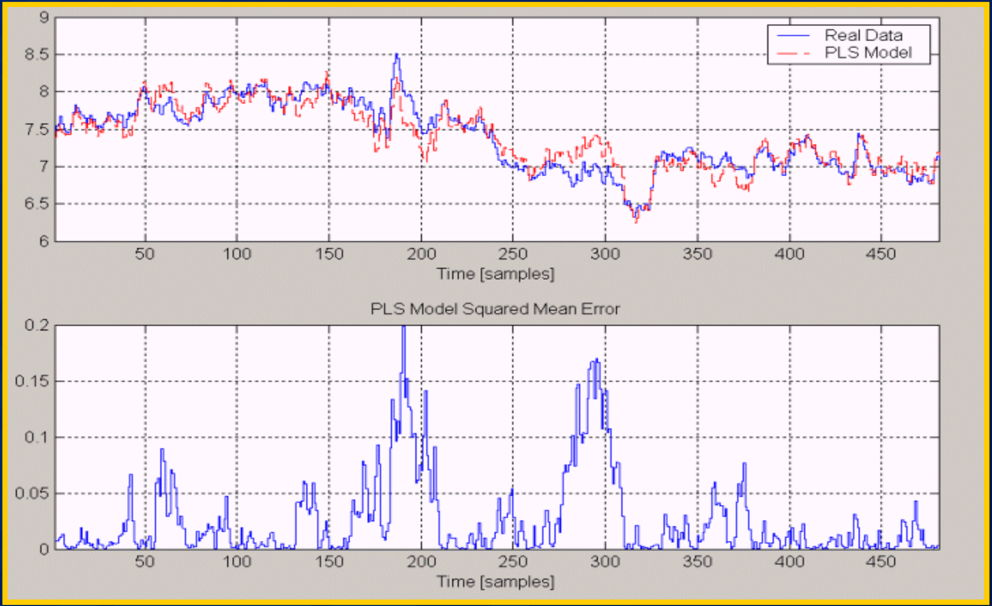
Cause-Effect (Structural) Representation Models

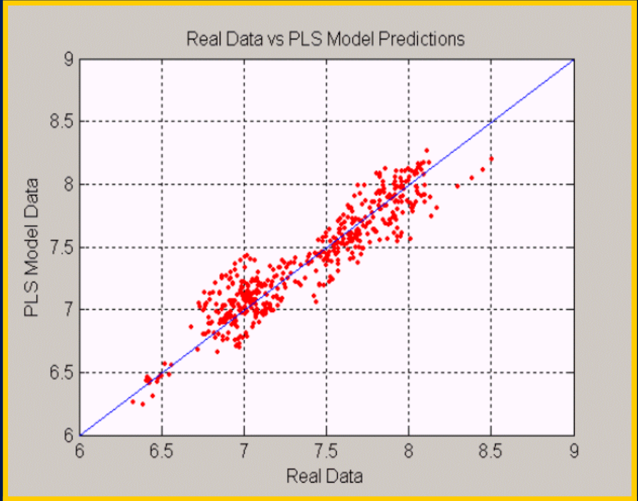
Some examples of practical applications:

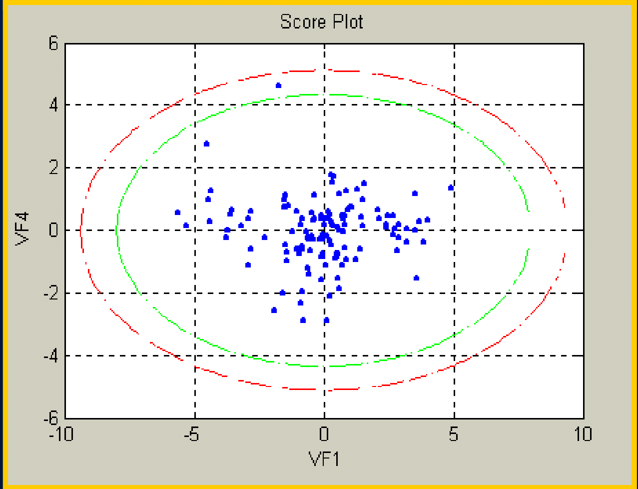
- **Process Variable Estimation**:
  - *Variables which are difficult to measure, ex.: “weariness”, “sheet break risk alert”, etc.*
- **Soft Sensors**:
  - *Variable values between Laboratory Tests*
- **Predictors:**
  - *EOB, End of Batch quality predictor*
- **Cause-Effect scenario analysis**
  - *Expected quality variability under different process variables values*

PLS: Examples and Demonstrations

1) Efficient Model Structure Identification: The method not only gives information about the structure of the model, but also about the RMS Prediction error in time.

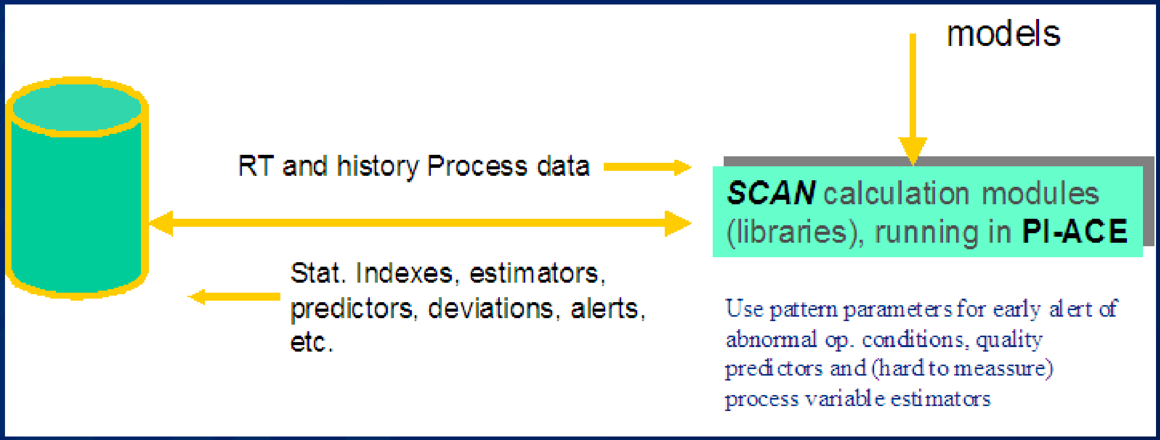






2) Additionally, it is possible to supervise the prediction ability of the model which it is been used in On-Line operation. The data score plot give that information by setting confidence limits for the model.

SCAN-PI, On-line Applications



ACE Calculation

Calculation execution trigger based on:

- *Time (clock)*
- *TAG Value*

- Tests are managed as ACE calculations
- Test inputs are “PI TAG´s”
- Test outputs are “PI TAG´s”
- Test parameters are maintained in PI MDB modules

This inherently modular architecture allows for:

Test enable/disable:

- Run Test$_1$ whenever ......, or TAG Value is ...... or...... GT than ......
- Run a Test$_2$ every …... [min]

Multi-test Linking

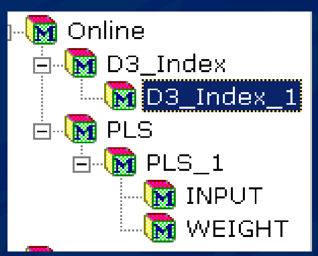
- Test$_1$(input) equals Test$_2$(output)

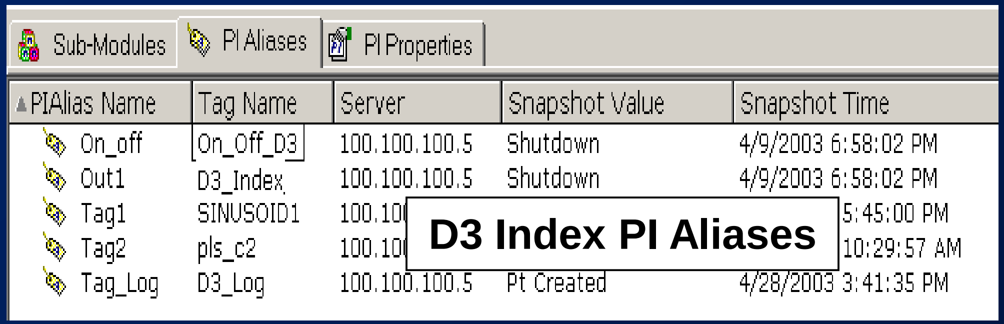
SCAN-PI, On-line Application

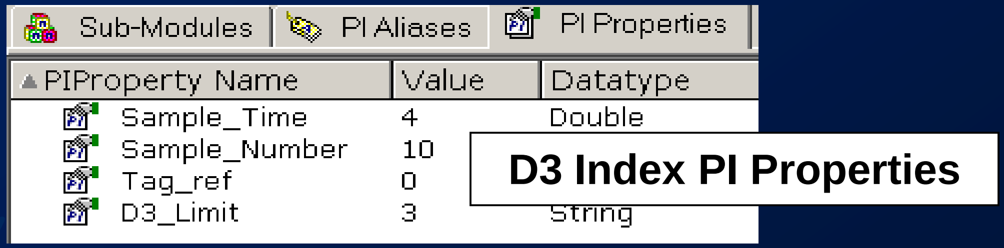
CONTAC

PI ACE (MDB) Model, ex.: D3 Index

MDB Structure





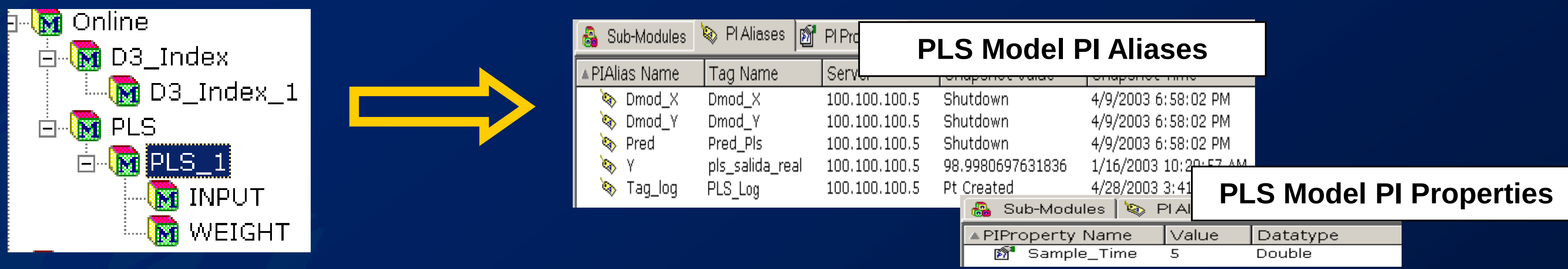


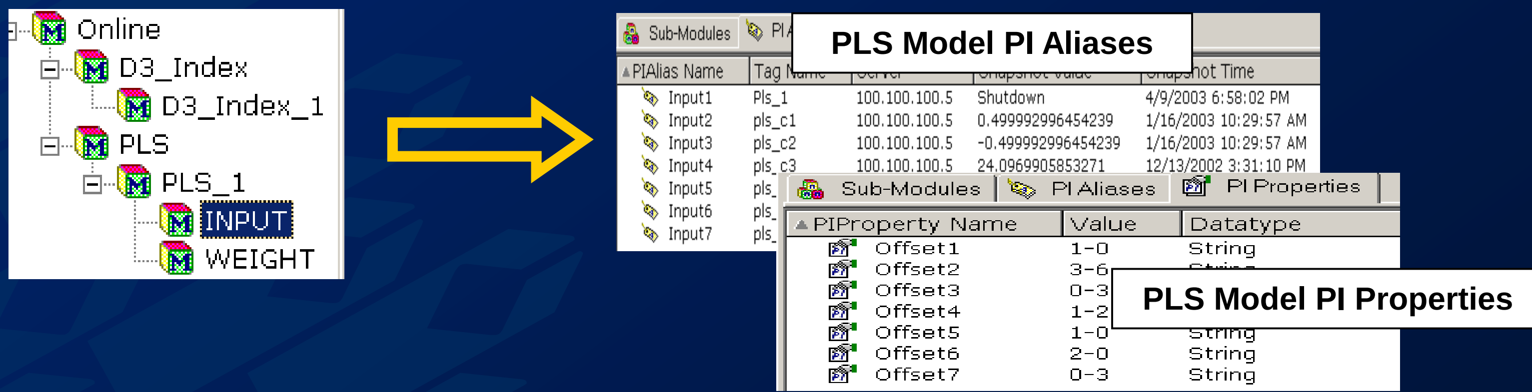
- Many instances can be defined for each test
- Specific parameters can be defined for each instance
- Specific I/O Tag references can be defined for each test
- Trigger can be a Tag reference signal or a fixed time period

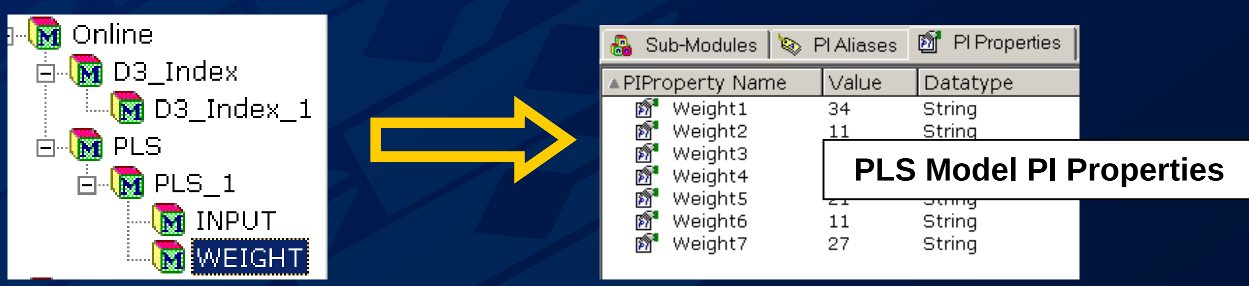
SCAN-PI, On-line Application

PI ACE (MDB) Model, ex.: PLS Model

MDB Structure



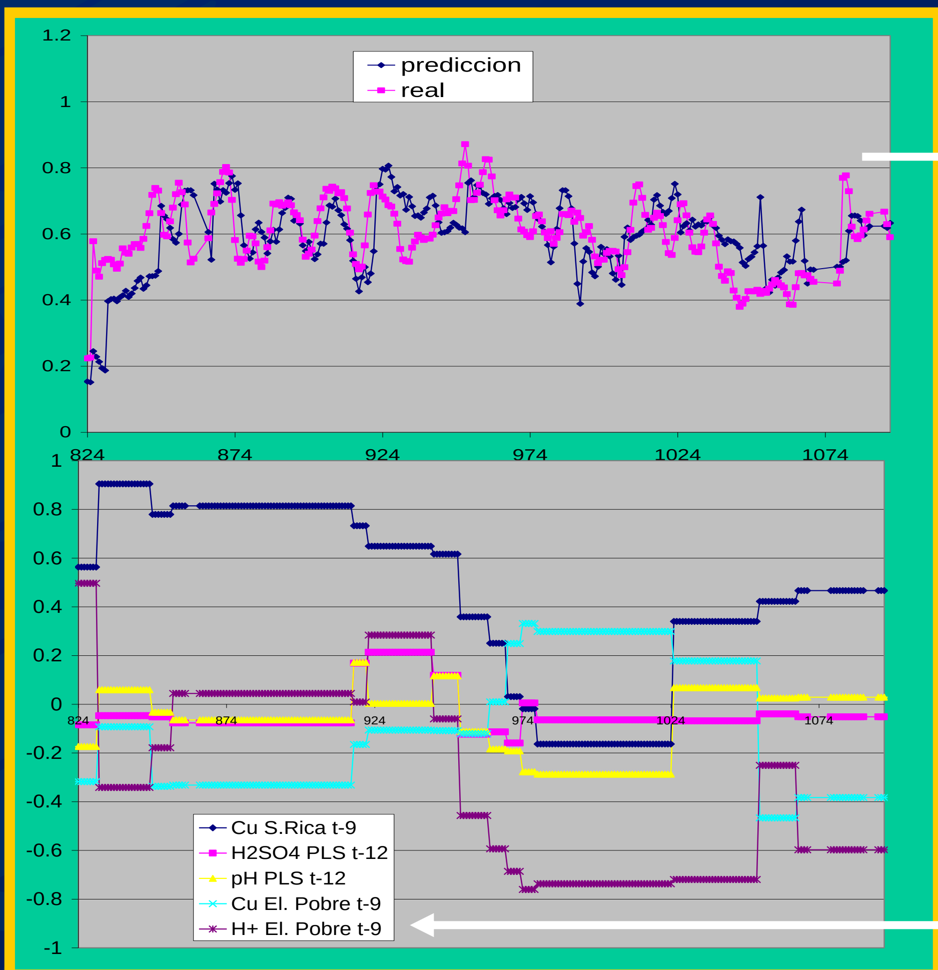


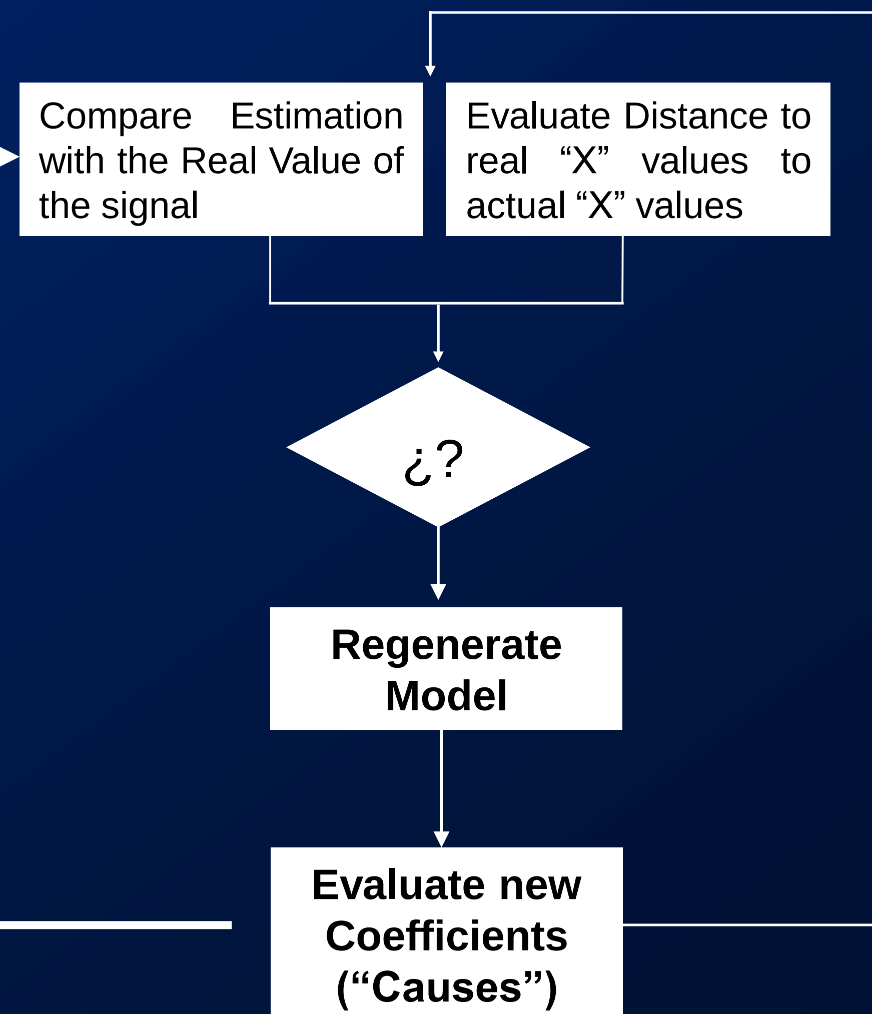


ON-LINE APPLICATIONS

EXAMPLE

The relative influence of every process variable in a certain output can be supervised On-Line by using an inteligent PLS Model actualization:





SCAN-PI, FAQ

- How are the model parameters saved?

  They are saved as Excel Worksheets. Therefore it is very simple to identify them and use them for further analysis. The same Worksheets can be used as input for an On-line application.

- Is it possible to model the process dynamics?

  Yes, by using delayed values of the sampled process variables.

- What about Batch Processes?

  A common technique called "unfolding" allows Batch Process analysis, by adding a new dimension of the data set (TAG-Time Matrix) related to the the Batch Number (Batch Run).

- How is it possible to model a non-linear process?

  When linear models are mentioned, we are talking about linear-in-the-parameters models. The use of calculated variables as model inputs (ex: $(\text{Pressure})^2/\text{Temperature}$) allows the modeling of non linear relationships.

- What about graphic representation of the on-line analysis?

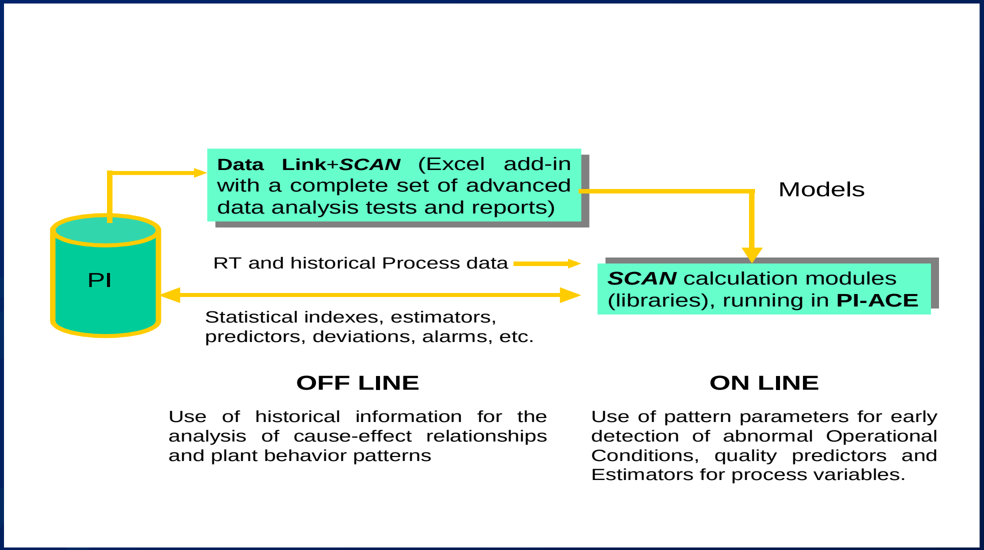
  Test´s outputs are sent back to PI, model parameters are managed in PI-MDB. Thus: all the PI power can be used.

- What about on-line deployment in DCS´s or PLC´s

  Model equations from the Off Line analysis can be implemented using calculation capabilities of the Control Systems, otherwise, ACE calculations can be sent back through PI.

- More questions?: please e-mail us

FINAL REMARKS



- PI Infrastructure provides an integrated environment for off-line analysis and on-line deployment of advanced process analysis.
- Infrastructure allows for a continuous improvement of SCAN capabilities, since new TEST´s can be similarly added to the libraries.
- Openness of the working space: both Off-line and On-line analysis structures allow the combination of Tests, Models and ad-hoc programming

CURRENT STATUS

OFF Line: V1.0

On Line (ACE libraries): Beta

- *D3 Test* (stability analysis, control loop assessment)
- *PLS projection model,* estimators, soft-sensors, predictors

FUTURE DEVELOPMENTS

- *Addition of new ACE Test´s: Wavelet, Specific Test´s for **failure early detection**, Batch models.*

lyacher@contac.cl