

PETRONAS TECHNICAL STANDARDS

DESIGN AND ENGINEERING PRACTICE

TECHNICAL, PROCEDURAL AND SPECIFICATIONS

BUILDING AND MONITORING STATISTICAL INFERENCE

PTS 20.00.10.30

MARCH 2010



© 2010 PETROLIAM NASIONAL BERHAD (PETRONAS)

All rights reserved. No part of this document may be reproduced, stored in a retrieval system or transmitted in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the permission of the copyright owner.

PTS Circular 2008 – 20.00.20.39		PETRONAS 
PTS No:	20.00.10.30	
Publication Title:	Building and Monitoring Statistical Inferential	
Base PTS Version:	Release 0	

This new PTS **20.00.10.30 – Building and Monitoring Statistical Inferential** has been developed to provide a standard guideline for building and monitoring performance of statistical inferential properties. It is intended for use by PETRONAS.

This PTS also includes PETRONAS Lessons Learnt and PETRONAS Best Practice for the subject matter Building and Monitoring Statistical Inferential

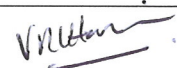
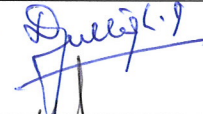
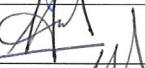
The custodian of this PTS is:

Name: Ir. Fauziah Ali

Tel. No: 03-27836449

Please direct any questions regarding this PTS to the above-named.

Document Approval

	Name	Designation	Date	Signature
Prepared	Fauziah Ali	Principal Advanced Process Control - GTS	30/3/10	
Reviewed		SKG16.9 TPs, APC CoP members	18 Nov 2009, 4 Feb 2010	
Approved	V. R. Harindran	Technical Authority	1/4/10	
Verified	Dzulkifli A. Razak.	Senior Manager, Process Automation & Optimisation, Engineering - GTS	2/4/2010	
Verified	Pau Kiew Huai	General Manager, Engineering - GTS	2/4/2010	
Endorsed	Pau Kiew Huai	Acting Senior General Manager - GTS	2/4/2010	

Revision History

Date	Version	Description of Updates	Author
December 2009	0	Original	Fauziah Ali

PREFACE

PETRONAS Technical Standards (PTS) publications reflect the views, at the time of publication, of PETRONAS OPU/Divisions.

They are based on the experience acquired during the involvement with the design, construction, operation and maintenance of processing units and facilities. Where appropriate they are based on, or reference is made to, national and international standards and codes of practice.

The objective is to set the recommended standard for good technical practice to be applied by PETRONAS' OPU in oil and gas production facilities, refineries, gas processing plants, chemical plants, marketing facilities or any other such facility, and thereby to achieve maximum technical and economic benefit from standardisation.

The information set forth in these publications is provided to users for their consideration and decision to implement. This is of particular importance where PTS may not cover every requirement or diversity of condition at each locality. The system of PTS is expected to be sufficiently flexible to allow individual operating units to adapt the information set forth in PTS to their own environment and requirements.

When Contractors or Manufacturers/Suppliers use PTS they shall be solely responsible for the quality of work and the attainment of the required design and engineering standards. In particular, for those requirements not specifically covered, it is expected of them to follow those design and engineering practices which will achieve the same level of integrity as reflected in the PTS. If in doubt, the Contractor or Manufacturer/Supplier shall, without detracting from his own responsibility, consult the owner.

The right to use PTS rests with three categories of users:

- 1) PETRONAS and its affiliates.
- 2) Other parties who are authorised to use PTS subject to appropriate contractual arrangements.
- 3) Contractors/subcontractors and Manufacturers/Suppliers under a contract with users referred to under 1) and 2) which requires that tenders for projects, materials supplied or - generally - work performed on behalf of the said users comply with the relevant standards.

Subject to any particular terms and conditions as may be set forth in specific agreements with users, PETRONAS disclaims any liability of whatsoever nature for any damage (including injury or death) suffered by any company or person whomsoever as a result of or in connection with the use, application or implementation of any PTS, combination of PTS or any part thereof. The benefit of this disclaimer shall inure in all respects to PETRONAS and/or any company affiliated to PETRONAS that may issue PTS or require the use of PTS.

Without prejudice to any specific terms in respect of confidentiality under relevant contractual arrangements, PTS shall not, without the prior written consent of PETRONAS, be disclosed by users to any company or person whomsoever and the PTS shall be used exclusively for the purpose they have been provided to the user. They shall be returned after use, including any copies which shall only be made by users with the express prior written consent of PETRONAS.

The copyright of PTS vests in PETRONAS. Users shall arrange for PTS to be held in safe custody and PETRONAS may at any time require information satisfactory to PETRONAS in order to ascertain how users implement this requirement.

TABLE OF CONTENTS

1.	INTRODUCTION.....	1
1.1	OBJECTIVE	1
1.2	DISTRIBUTION, INTENDED USE AND REGULATORY CONSIDERATIONS.....	1
1.3	DEFINITIONS.....	1
1.4	INTRODUCTION TO STATISTICAL INFERENCE MODELLING	2
2.	STATISTICAL INFERENCE MODELLING WORKFLOW.....	3
3.	INFERENCE MODELLING	4
3.1	DATA PRE-PROCESSING	4
3.2	RUNNING INFERENCE MODELLING	5
3.3	GOODNESS OF FIT	6
3.4	VALIDATION	6
4.	CONFIGURATION FOR ON-LINE IMPLEMENTATION	7
4.1	NEED FOR UPDATING INFERENCE	7
4.2	CRITERIA FOR UPDATING INFERENCE	7
4.3	METHODS FOR UPDATING	7
4.4	OTHER GENERAL RULES.....	8
5.	INFERENCE MONITORING	9
5.1	ONCE-OFF VALIDATION OF EXISTING INFERENCE	9
5.2	CONTINUOUS MONITORING OF INFERENTIALS	10
6.	TOOLS AND CALCULATIONS.....	11
7.	REFERENCES	12

APPENDICES

APPENDIX A	SAMPLE OF CONTROL CHART	13
APPENDIX B	EXAMPLE OF INFERENCE VERIFICATION	14

1. INTRODUCTION

1.1 OBJECTIVE

The objective of this document is to provide a standard guideline for building and monitoring statistical inferential. It is intended for use by PETRONAS and Vendors and is applicable for building new inferential, updating existing inferential as well as monitoring new and existing inferential over time.

1.2 DISTRIBUTION, INTENDED USE AND REGULATORY CONSIDERATIONS

Unless otherwise authorised by PETRONAS, the distribution of this document is confined to companies forming part of or managed by PETRONAS, and to Contractors and Manufacturers/Suppliers nominated by them.

This PTS is intended for use in oil refineries, gas plants, chemical plants and, where applicable, in exploration and production facilities and supply/marketing installations. If national and/or local regulations exist in which some of the requirements may be more stringent than in this PTS, the Contractor shall determine by careful scrutiny which of the requirements are the more stringent and which combination of requirements will be acceptable as regards safety, environmental, economic and legal aspects. In all cases the Contractor shall inform the Principal of any deviation from the requirements of this PTS which is considered to be necessary in order to comply with national and/or local regulations. The Principal may then negotiate with the authorities concerned with the object of obtaining agreement to follow this PTS as closely as possible.

1.3 DEFINITIONS

1.3.1 General Definitions

The **Contractor** is the party which carries out all or part of the design, engineering, procurement, construction, commissioning or management of a project or operation of a facility. The Principal may undertake all or part of the duties of the Contractor.

The **Manufacturer/Supplier/Vendor** is the party which manufactures or supplies equipment and services to perform the duties specified by the Contractor.

The **Principal** is the party which initiates the project and ultimately pays for its design and construction. The Principal will generally specify the technical requirements. The Principal may also include an agent or consultant authorised to act for, and on behalf of, the Principal.

The word “**shall**” indicates a requirement.

The word “**should**” indicates a recommendation.

1.3.2 Specific Definitions

Inferential Property

Estimation of a property based on other measurements. Also commonly known as soft sensor, virtual analyser and quality estimator.

Statistical Inferential Property

Inferential property that is derived from using statistical methods (numerically derived from process measurements), such as multiple linear regression.

1.3.3 Abbreviations

APC	Advanced Process Control
ARC	Advanced Regulatory Control
LCL	Lower Control Limit
LIMS	Lab Information Management System
MLR	Multiple Linear Regression
PLS	Projection to Latent Structure, also known as Partial Least Squares.
R^2	Sum of Squared Residual (or SSR)
RMSE	Root Mean Squared of Error
Sigma	Standard Deviation of Error
UCL	Upper Control Limit
VIF	Variance Inflation Factor

1.4 INTRODUCTION TO STATISTICAL INFERENCE MODELLING

An inferential model, also known as a soft sensor, is an estimation of a property based on other measurements e.g. operating flow, temperature and pressure. In the oil and gas industry process control environment, it is a common component of Advanced Process Control (APC) and Advanced Regulatory Control (ARC) applications, but it can also stand on its own as a reference point for panel operator and engineers alike.

There are several reasons for building and using inferential, including:

- Validating instrumentation and providing early detection of drift or malfunction.
- Serving as a backup in the event of analyser problems
- Replacing physical sensors (transmitters) for process measurements.
- Improving control by providing feedback between lab samples or analyser results.
- Reducing lab costs.

A statistical inferential model is one that is built using statistical methods, ie numerically derived from process data. There is a broad suite of tools available for developing statistical inferential models. Many commercial tools offer the user a choice of similar and common modelling algorithms, such as multiple linear regression (MLR), and projection to latent structures (PLS). Less commonly applied methods include various forms of neural network and fuzzy logic. Much APC software incorporate inferential modelling tool as part of their suite of tools for APC development. In such cases, provision is usually readily available to facilitate inferential updating using both on-line analyzer and lab data.

This PTS focuses on statistical inferential modelling and is written to be non-specific to any commercial software. For completeness sake, it is hereby mentioned that a second category of inferential modelling exists and is called 1st principles. 1st principles models utilise fundamental chemical engineering relationships and equations in order to derive the calculated property. They are presently used in fewer numbers of sites across PETRONAS.

2. STATISTICAL INFERENCE MODELLING WORKFLOW

Workflow for development of statistical inferential modelling is graphically represented in the chart below. The workflow is applicable whether the inferential is used in an APC, ARC, or just as an indicator for panel operator or engineer's reference.

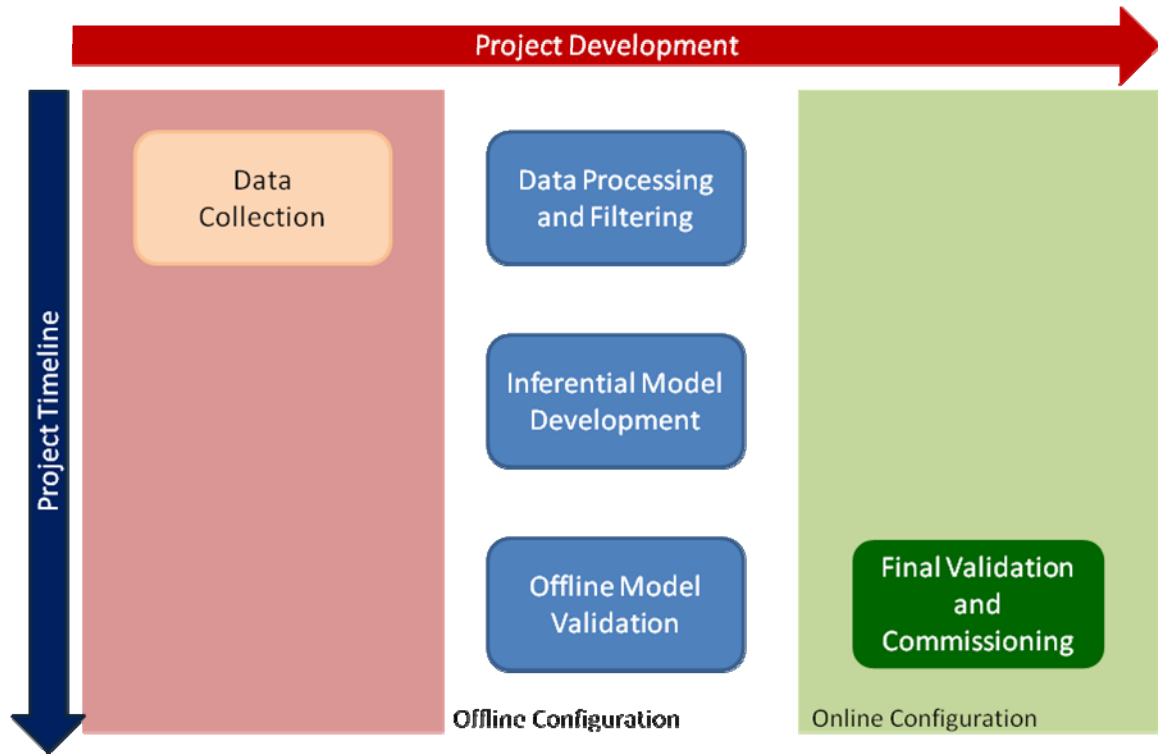


Figure 2.0 Workflow for building inferential.

3. INFERENCEAL MODELLING

This section focuses for data importing, data pre-treatment, data analysis, model development and validation

3.1 DATA PRE-PROCESSING

- a. The first step is data collection. Data on property to be inferred may come from laboratory or on-line analyser. Process data are also collected, with compression on the plant historian removed.
- b. Data should be at steady state and at normal operating condition when collected. A more superior inferential quality can be developed using data that is collected at steady state rather than from using data that is collected at fixed hours each day such as in routine sampling. This is because, while being easier for planning purposes, there is no guarantee that the plant is steady during those specific times of each day.
- c. There shall be a minimum of 60 good datapoints for inferential development - 30 of which are to be used for training and 30 for validation. This is the minimum number of samples to ensure statistical significance when building and validating the inferential.
- d. It is recommended that data be randomised in order to capture different operating modes before dividing them into training and validation datasets.
- e. Collecting more than 60 datapoints is recommended so that after data cleaning and removing invalid datapoint, there will be the minimum 60 datapoints left.
- f. For inferential modelling using lab sample, always use the timestamp that the sample was taken, not the time that the sample was tested in the lab.
- g. To account for inaccuracies in sample time recorded and to improve quality of inferential, the modeller should try taking averages of process data around the time of lab sample. The modeller may try the suggestions below, or other averages according to his own engineering judgement.
 - i. Taking half hour or one hour average, ending at the time the sample was taken
 - ii. Taking half hour or one hour average, mid-point at the time the sample was taken
- h. Inferential modelling against spot readings or 1-minute average data at the time the sample is taken shall be avoided.
- i. Data cleaning and screening is as important as model fitting. Some basic checks shall be applied to remove irrelevant datapoints. These may be situations of abnormal plant operations such as running at very low throughput, conducting a test run , major equipment failure and plant upsets.
- j. A statistically-derived inferential is typically only good within the operating envelope that it was developed in. Therefore developing an inferential using a wider operating envelope will make the inferential more robust.
- k. Any extrapolation of an inferential outside its normal envelope should be done with extreme caution and with sufficient validation effort.

3.2 RUNNING INFERENCE MODELLING

- a. A statistical inferential relates an inferred value (output variables, usually referred to as Y value) to values of other known process parameters (input variables, usually referred to as X values).
- b. Statistical inferential can be built using many readily available commercial tools from spreadsheets to more sophisticated modules usually available with APC software.
- c. Some software and algorithm require that inputs be uncorrelated, i.e. the inputs must all be independent of each other. Some software checks for input correlation and highlight to the user accordingly. Some algorithms (e.g. latent variables) are by design meant to handle input correlation gracefully. When using PLS, variables of high significance are automatically selected and input-to-input correlation detected and automatically managed. The user shall find out the features of the particular software and algorithm that he/she is using.
- d. For MLR, regression is done using an initial, larger set of inputs. Inputs of low significance are subsequently removed, one at a time, using Variance Inflationary Factor (VIF) and statistical P value as selection criteria.
 - i. Inputs with VIF of >5 indicates co linearity and should be removed.
 - ii. P value of >0.05 indicate low significance and are to be eliminated.
 - iii. The resulting Adjusted Sum of Squared Residual (Adjusted SSR or R^2) value should be >0.6 for the inferential to be fit for use in process control.

Depending on software used, the above parameters may all be automatically calculated, or the user may have to calculate its own values.

- e. In general, there are no hard and fast rules for minimum or maximum number of input variables. However, from a practical standpoint, maintaining 6 or less number of input variables is good as the more input variables used, the more likelihood that at some time in the future, one of the input variables will have a measurement problem and hence affect the inferred value.
- f. Depending on the application, linearizing some input variables using known key parameters (called secondary inputs) according to process knowledge and chemical engineering fundamentals may improve the inferential quality, for example
 - i. Using pressure-compensated temperature instead of temperature reading, natural logarithm of temperature or inverse of temperature
 - ii. Using ratio of steam to feed instead of just steam flow
 - iii. Using reflux ratio instead of reflux flow
- g. The selected inputs and their gain sign (+ve or -ve) must make sense from a process standpoint.

3.3 GOODNESS OF FIT

To indicate the goodness of fit between the inferential vs. lab or inferential vs. on-line analyser, the following shall be plotted:

1. Lab/analyser and inferential over time should be in sync, indicating a good fit.
2. Scatter plot (lab/analyser vs. inferential). A scatter close to 45 deg with minimal spread indicates a good fit.

Additionally, user shall focus on minimizing the difference between inferential and lab/on-line values. The following parameters are to be calculated.

- a. Root Mean Square of the prediction Error (RMSE).
- b. Standard Deviation of prediction Error (Sigma).
- c. Adjusted R^2 . A minimum value of 0.6 renders an inferential suitable for use in process control.

Some software has their own measures for goodness of fit. User may alternatively use these.

3.4 VALIDATION

The intent of validation is to check the goodness of fit of the inferential against a set of data that is different from the set of data used for model building. Some software allow the use to select data regions to be used for modelling and for validation. If user tools do not include this feature, the user shall do validation manually, using a different dataset.

- a. Minimum number of datapoints used for validation shall be 30
- b. The same set of model quality evaluation rules are applied to validation dataset.

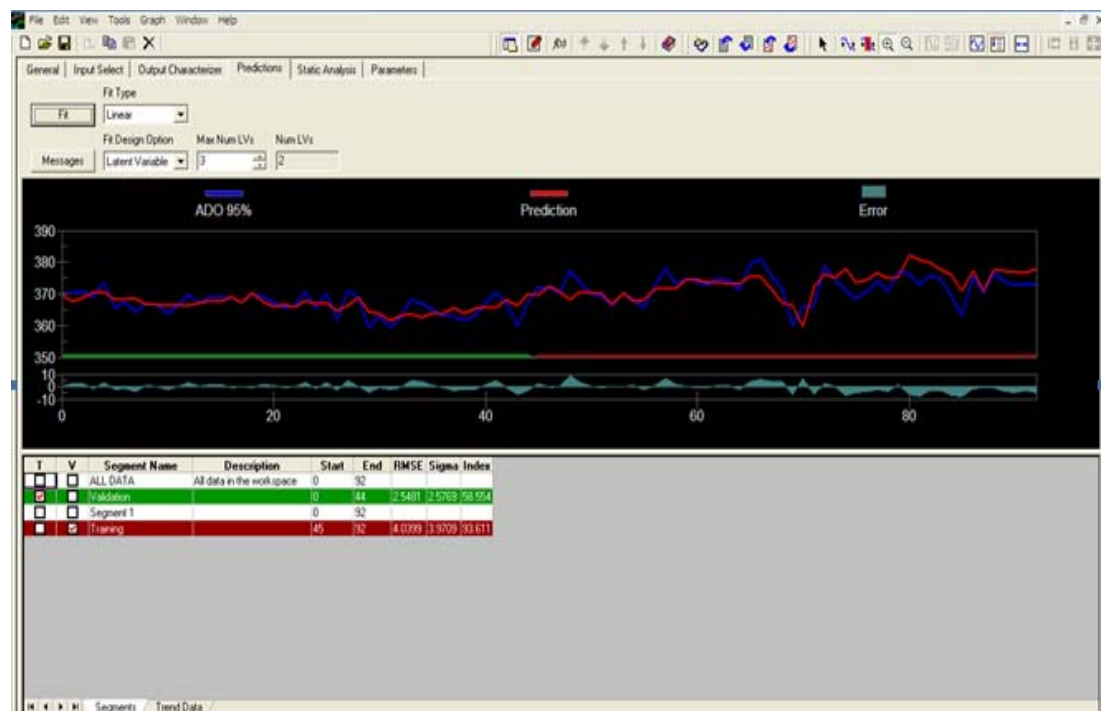


Figure 3.0 Example of dividing data into training and validation datasets.

4. CONFIGURATION FOR ON-LINE IMPLEMENTATION

For on-line implementation, provisions shall be made to keep inferential up-to-date and continuously matching lab or on-line analyser readings.

4.1 NEED FOR UPDATING INFERENTIAL

- a. If left untouched, inferential will eventually deviate from lab or on-line analyser reading that it is meant to represent.
- b. From time to time, it is necessary to update bias term of the inferential to keep it close to lab/on-line analyser.

4.2 CRITERIA FOR UPDATING INFERENTIAL

- a. It is NOT recommended to change the inferential bias term for every lab sample or on-line analyser reading to force a zero error, as this action would reduce the quality of the inferential model.
- b. The engineer should decide, based on his engineering judgement (and knowledge of his lab and measuring instruments' accuracy and repeatability), a tolerance within which there is no need to update bias at all.
- c. If error is larger than this tolerance, the inferential bias should be updated at a fraction of the error only. A fraction of 0.3 to 0.6 is recommended.

4.3 METHODS FOR UPDATING

- a. Automatic bias update :

If inferential is built on a commercial on-line platform which is linked to PI or LIMS, the inferential can be automatically biased upon data entry into PI or LIMS. These platforms typically have user-configurable validity checks to check the validity of lab data or on-line analyser reading. The engineer responsible for maintaining the inferential shall ensure he understands the update mechanism in the platform used.

- i. It is the responsibility of the engineer to configure the checks correctly e.g. to handle potential scenario of wrong lab data entry, timeout after no lab entry or bad value for on-line analyser reading.
- ii. For timeout due to no lab data entry, engineer shall estimate timeout based on frequency of routine samples, and decide on an allowable number missed sample before being considered as timeout. For example, if routine sample is taken once every 24 hours, and engineer intends to allow for 1 skipped sample, timeout shall be (48 hours + typical time lapse between sampling and lab data entry)
- iii. For timeout due to no on-line analyser reading, engineer shall decide on the number analyser cycle times he would allow to be missed before being considered as timeout. For example, if an on-line analyser takes half an hour to complete the test and engineer would allow for 5 cycles of no reading, timeout shall be 2.5 hours (i.e. 5X0.5 hours).

- iv. Engineer shall decide on next course of action after timeout. He may opt for the inferential to be dropped from the Controlled Variable (CV) list, or he may choose to configure an alarm to alert panel operator.

b. Manual check before bias update :

Most OPU's prefer that lab data is validated by process or APC personnel before bias is updated. Bias update is executed either

- i. when operations or APC personnel manually enter the lab value to enable the inferential to be biased (having validated the lab reading himself)
- ii. when operations or APC personnel manually accepts a value read from PIMS.

Similar to automatic bias updating, engineer shall establish his work process and actions for timeout.

c. Manual bias update :

In some applications, the bias term is updated by the authorised person directly, upon his own monitoring eg in cases where the inferential is built as a simple DCS calculation tag.

4.4 OTHER GENERAL RULES

In any method of inferential updating, lab samples taken when the plant is in upset condition shall not be used for comparing and updating.

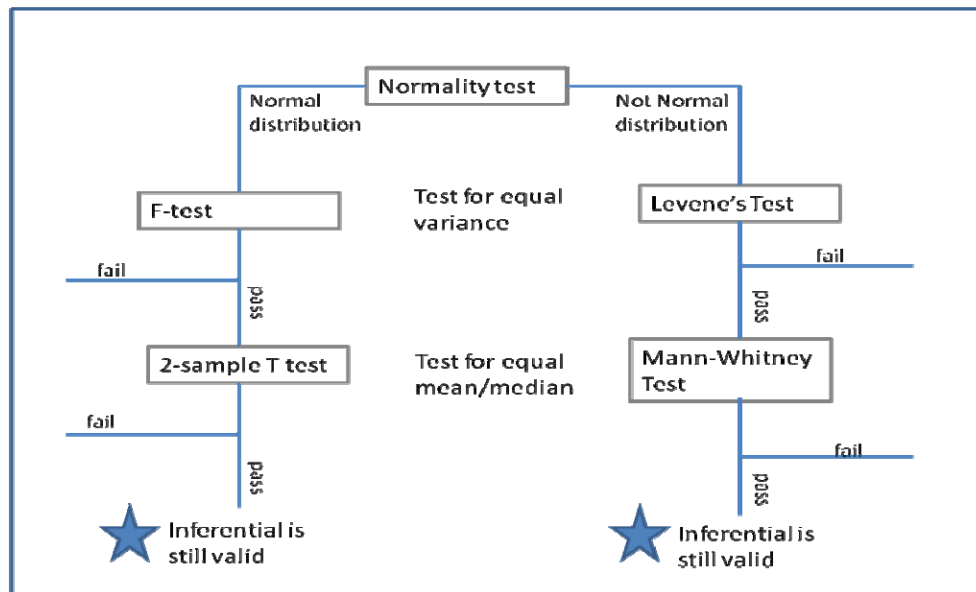
Inferentials shall also not be updated using reading from on-line analyser which itself is undergoing calibration or validation.

5. INFERENCEAL MONITORING

Subsequent to building a satisfactory inferential, an inferential shall be monitored to ensure that it is still well correlated against lab or on-line analyser. Timely action shall be taken in doing the needful to ensure continuously valid inferential.

5.1 ONCE-OFF VALIDATION OF EXISTING INFERENCEAL

If an existing inferential has not been validated for some time, a once-off validation shall be carried out to test if the inferential is still valid. Inferential shall be validated against lab or analyser; testing both populations for equality of mean/median as well as equality of median. Selection of the correct set of tests depends on whether the sample population has normal distribution and is represented in the diagram below.



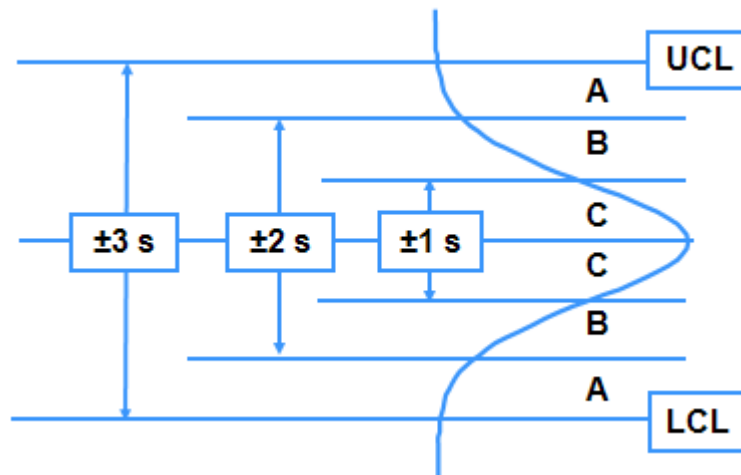
- i. Normality test.
The intent of this test is to determine whether the sample distribution follows the normal distribution (Gaussian distribution, graphically also known as the bell curve). Several tests can be done to establish normality, from an informal test of comparing histograms of residuals to a normal probability curve, to more rigorous test such as Anderson–Darling and Shapiro–Wilk test.
- ii. F-test
This test analyses the variance between two samples, for normally distributed samples. P-value > 0.05 indicates that both samples are assumed equal and representative of each other.
- iii. 2-sample T-test
This test compares means for two samples, for normally distributed samples. The two populations are considered the same when the test passes at P-value>0.05.
- iv. Mann-Whitney Test
This test for equal medians, for samples not following a normal distribution, should yield P-value>0.05 for the first hurdle that the inferential is still valid vs. lab reading.
- v. Levene's Test
This test for equal variance should also yield P-value>0.05.

When conditions are met for F-test AND 2-sided T-test, OR Mann-Whitney AND Levene's Test (depending on whether the sample is normally distributed), the inferential can be considered still valid against lab/on-line analyser.

5.2 CONTINUOUS MONITORING OF INFERENTIALS

Inferential shall be subjected to continuous monitoring to ensure they are continuously valid and fit for use.

- Monitoring shall be by means of control charting, i.e. plotting inferential-lab (inferential minus lab reading) over time.
- Standard control chart run rules (e.g. Nelson's test, Western Electric test) apply in detecting abnormal conditions whereupon the engineer is expected to investigate the cause of the abnormality and take corrective action if necessary.
- An explanation of the Nelson test for control charting is listed below. Other tests offer slightly different variations of the rules.



where

“s” represents standard deviation of the error at a time the inferential was built.
UCL is the upper control limit which is equivalent to +3 standard deviations.
LCL is the lower control limit which is equivalent to -3 standard deviations.

According to Nelson's test, abnormalities are detected following the rules below :

- Any one point outside control limits
- 9 consecutive points on same side of centerline
- 6 consecutive points increasing or decreasing
- 2 of 3 points in same Zone A or beyond
- 4 of 5 points in same Zone B or beyond
- 14 consecutive points alternating up and down
- 15 consecutive points in either Zone C
- 8 points in a row outside Zone C, same side of centerline

6. TOOLS AND CALCULATIONS

Calculations for Anderson-Darling Test, Shapiro-Wilk Test, Mann-Whitney Test, Levene's Test and control charting can be readily sourced through many publications, both hardcopies and on-line, and may be calculated by hand.

For rapid and cost effective deployment across PETRONAS, an in-house spreadsheet-based tool is being developed by Group Technology Solutions. This programme can be installed on desktops or laptops can be run on an ad-hoc basis, extracting information from Plant Information Management System (PIMS) or Lab Information Management System (LIMS), or be connected live to PIMS/LIMS for continuous monitoring.

Commercial statistical tools are also available. These offer a wide variety of statistical tests including the above tests. They may be used if already available and user is familiar with and has a preference for them.

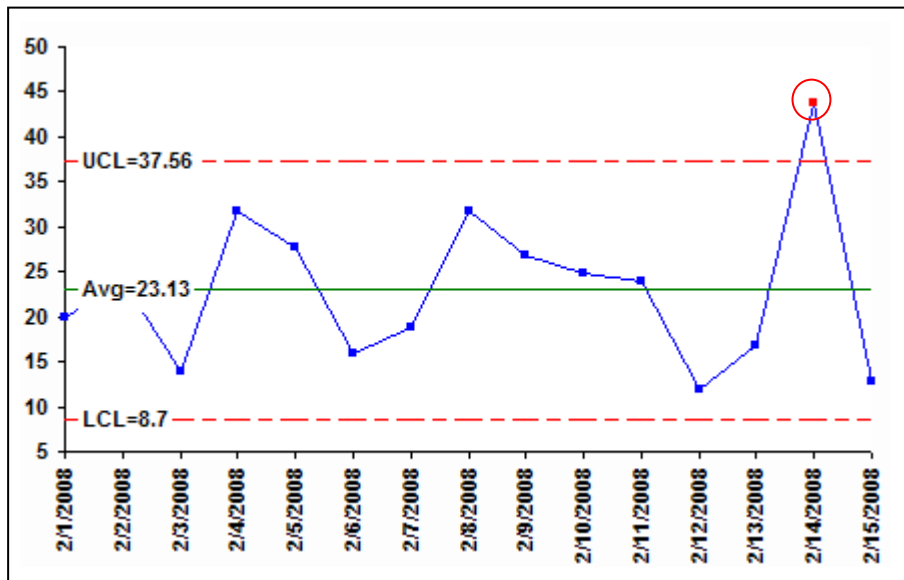
Note :

For information about tool or its deployment, please contact directly
APC Group, Process Automation and Optimisation, Group Technology Solutions.
or through
APC Community of Practices

7. REFERENCES

1. Application of Statistical Quality Control for Correlation of Analytical Results between Online Analytical Instrumentation and Laboratory Analysis, PETRONAS.
2. Process Control Training, Section 15, Whitehouse Consulting
3. Robust Quality Estimator manual
4. Introduction to SPC, INSTEP
5. Six Sigma and Minitab, A Complete Toolbox Guide for all Six Sigma Practitioners

APPENDIX A SAMPLE OF CONTROL CHART



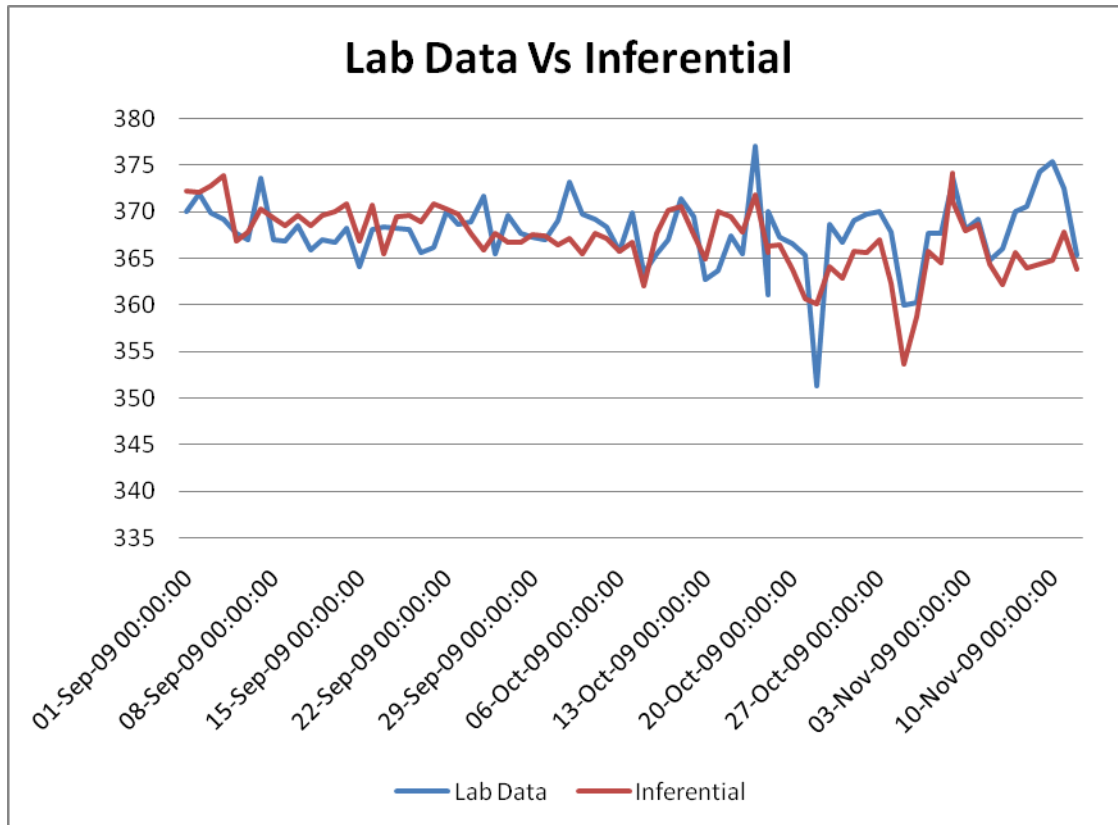
Above is a sample of a control chart in which rule number (i) is violated and the point is circled in red.

APPENDIX B EXAMPLE OF INFERENTIAL VERIFICATION

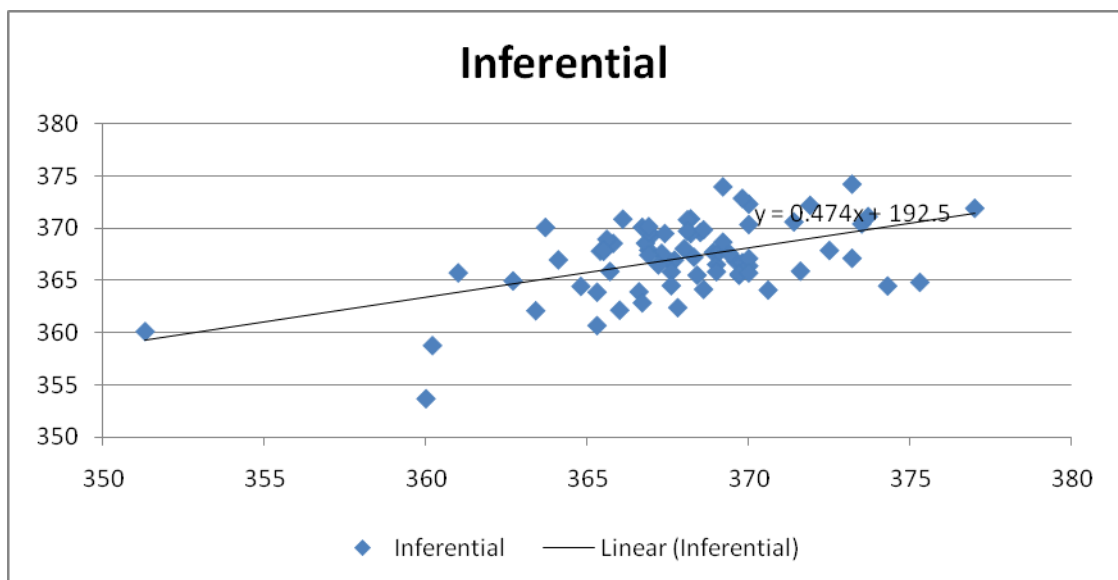
Lab Data Versus Inferential:

Sample Time	Lab Data	Inferential
01-Sep-09 06:00:00	370.00	372.28
02-Sep-09 06:00:00	371.90	372.17
03-Sep-09 06:00:00	369.80	372.83
04-Sep-09 06:00:00	369.20	373.93
05-Sep-09 06:00:00	367.70	366.94
06-Sep-09 06:00:00	366.90	367.87
07-Sep-09 06:00:00	373.50	370.39
08-Sep-09 06:00:00	367.00	369.31
09-Sep-09 06:00:00	366.80	368.54
10-Sep-09 06:00:00	368.50	369.57
11-Sep-09 06:00:00	365.80	368.53
12-Sep-09 06:00:00	366.90	369.62
13-Sep-09 06:00:00	366.70	370.06
14-Sep-09 06:00:00	368.20	370.87
15-Sep-09 06:00:00	364.10	366.95
16-Sep-09 06:00:00	368.10	370.78
17-Sep-09 06:00:00	368.40	365.49
18-Sep-09 06:00:00	368.20	369.45
19-Sep-09 06:00:00	368.10	369.70
20-Sep-09 06:00:00	365.60	368.92
21-Sep-09 06:00:00	366.10	370.84
22-Sep-09 06:00:00	370.00	370.33
23-Sep-09 06:00:00	368.60	369.82
24-Sep-09 06:00:00	368.90	367.66
25-Sep-09 06:00:00	371.60	365.87
26-Sep-09 06:00:00	365.40	367.78
27-Sep-09 06:00:00	369.60	366.82
28-Sep-09 06:00:00	367.60	366.69
29-Sep-09 06:00:00	367.30	367.58
30-Sep-09 06:00:00	366.90	367.42
01-Oct-09 06:00:00	369.00	366.51
02-Oct-09 06:00:00	373.20	367.10
03-Oct-09 06:00:00	369.70	365.49
04-Oct-09 06:00:00	369.10	367.77
05-Oct-09 06:00:00	368.30	367.21
06-Oct-09 06:00:00	365.70	365.84
07-Oct-09 06:00:00	369.80	366.69
08-Oct-09 06:00:00	363.40	362.09

Sample Time	Lab Data	Inferential
09-Oct-09 06:00:00	365.50	367.77
10-Oct-09 06:00:00	366.90	370.13
11-Oct-09 06:00:00	371.40	370.58
12-Oct-09 06:00:00	369.40	367.54
13-Oct-09 06:00:00	362.70	364.94
14-Oct-09 06:00:00	363.70	370.05
15-Oct-09 06:00:00	367.40	369.47
16-Oct-09 06:00:00	365.50	367.88
17-Oct-09 06:00:00	377.00	371.88
18-Oct-09 02:31:00	361.00	365.71
18-Oct-09 06:00:00	370.00	366.37
19-Oct-09 06:00:00	367.20	366.47
20-Oct-09 06:00:00	366.60	363.90
21-Oct-09 06:00:00	365.30	360.67
22-Oct-09 06:00:00	351.30	360.11
23-Oct-09 06:00:00	368.60	364.14
24-Oct-09 06:00:00	366.70	362.84
25-Oct-09 06:00:00	369.00	365.82
26-Oct-09 06:00:00	369.70	365.60
27-Oct-09 06:00:00	370.00	367.08
28-Oct-09 06:00:00	367.80	362.38
29-Oct-09 06:00:00	360.00	353.69
30-Oct-09 06:00:00	360.20	358.76
31-Oct-09 06:00:00	367.60	365.80
01-Nov-09 06:00:00	367.60	364.49
02-Nov-09 06:00:00	373.20	374.18
02-Nov-09 13:58:00	373.70	371.09
03-Nov-09 06:00:00	368.00	368.02
04-Nov-09 06:00:00	369.20	368.64
05-Nov-09 06:00:00	364.80	364.42
06-Nov-09 06:00:00	366.00	362.15
07-Nov-09 06:00:00	370.00	365.71
08-Nov-09 06:00:00	370.60	364.07
09-Nov-09 06:00:00	374.30	364.46
10-Nov-09 06:00:00	375.30	364.81
11-Nov-09 06:00:00	372.50	367.85
12-Nov-09 06:00:00	365.30	363.86



Inferential Scatter Plot:



Statistical Test

1. Test for Normality

Result: Lab data and Inferential is not normally distributed.

2. Test for equal median using Mann Whitney Test

Result:

P value = 0.2163

Conclusion:

P value > 0.05, therefore both lab data and inferential have equal median.

3. Test for equal variance using Levene Test

Result:

P value = 0.9960

Conclusion:

P value > 0.05, therefore both lab data and inferential have equal variance.

Conclusion :

Inferential is still valid.