INTERNATIONAL STANDARD

ISO
18213-3

First edition
2009-03-01

Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy

Part 3:
Statistical methods

Technologie du combustible nucléaire — Étalonnage et détermination du volume de cuve pour la comptabilité des matières nucléaires

Partie 3: Méthodes statistiques

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 18213-3 was prepared by Technical Committee ISO/TC 85, *Nuclear energy*, Subcommittee SC 5, *Nuclear fuel technology*.

ISO 18213 consists of the following parts, under the general title *Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy*:

— *Part 1: Procedural overview*

— *Part 2: Data standardization for tank calibration*

— *Part 3: Statistical methods*

— *Part 4: Accurate determination of liquid height in accountancy tanks equipped with dip tubes, slow bubbling rate*

— *Part 5: Accurate determination of liquid height in accountancy tanks equipped with dip tubes, fast bubbling rate*

— *Part 6: Accurate in-tank determination of liquid density in accountancy tanks equipped with dip tubes*

Introduction

This part of ISO 18213 describes statistical procedures suitable for the treatment of tank calibration and volume measurement data for nuclear materials accountancy tanks. It is one part of a six-part International Standard that deals with the acquisition, analysis, standardization and use of calibration data to determine liquid volumes in process tanks for accountability purposes, and is intended for use in conjunction with other parts of ISO 18213. Other parts of ISO 18213 and their topics are ISO 18213-1 (procedural overview), ISO 18213-2 (data standardization), ISO 18213-4 (slow bubbling rate), ISO 18213-5 (fast bubbling rate), and ISO 18213-6 (in-tank determination of liquid density).

To someone without formal statistical training, the methods of ISO 18213-3 might appear to be unnecessarily complex. However, within the context of the data standardization model presented in other parts of ISO 18213, the statistical methods presented herein have been kept as simple as possible. Data collection, data standardization and statistical analysis go hand-in-hand. In order for one to meet the target uncertainty limits established for accountability purposes, it is necessary that the data standardization model be consistent with the measurement (instrument) capability and that the statistical error model likewise be compatible with the data standardization model. It makes no sense to use a highly refined data standardization model with crude measurement instruments. Conversely, the advantage of highly refined and precise measurement instruments is lost if a crude data standardization model is used in the subsequent analysis. Using a more refined measurement instrument, for example, does not improve results if the data standardization model fails, for example, to take proper account of the effects of temperature variation.

Similarly, it makes no sense to use a sophisticated statistical model with either crude measurements or a crude data standardization model. Conversely, an overly simple statistical model, or one that is inconsistent with the underlying data standardization model, yields poor results even when used with high-quality instrumentation and a refined data standardization model. Because of the important role volume determinations play in its overall accountability program, a facility typically devotes significant resources to instrumentation for tank calibration and volume determination. However, refined state-of-the-art measurement capability by itself is not sufficient to meet target uncertainty limits. Resources are also required to develop a data standardization model and statistical methods with quality comparable to that of the plant's measurement capability. The resources required for data analysis are typically much fewer than those allocated for instrumentation, but they are equally as important. In any event, adequate resources are required to engage someone with the necessary training to guide the development and application of computational and statistical methods that are comparable in sophistication to the measurements to which they are applied.

The statistical methods presented in this part of ISO 18213 are closely tied to the comprehensive state-of-the-art data standardization methodology presented in other parts of ISO 18213 and are therefore designed to be applicable over a wide range of measurement systems and operating conditions. As noted in the introduction to ISO 18213-1, it is not always necessary, or even possible, for the operator to develop the full model for all tanks in a given facility. Under these circumstances, the methods presented herein provide the framework for developing a "reduced" calibration model, including suitable estimates of uncertainty, that is consistent with the "reduced" standardization model developed for a particular tank.

Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy

Part 3:
Statistical methods

1 Scope

This part of ISO 18213 presents statistical procedures that can be applied to tank calibration and volume measurement data for nuclear materials accountancy tanks. In particular, this part of ISO 18213 presents

a) several diagnostic plots that can be used to evaluate and compare tank calibration data;

b) a procedure for estimating the uncertainties of tank calibration measurements (i.e., determinations of height and volume);

c) a model for estimating either a tank's calibration equation or its inverse (the measurement equation), together with related uncertainties, from a set of standardized tank calibration data (i.e., from a series of standardized height-volume determinations);

d) a method for computing uncertainty estimates for determinations of liquid volume.

It is intended that the methods in this part of ISO 18213 be used within the context of the other parts of ISO 18213. Specifically, the methods presented in this part of ISO 18213 are tailored to the general methodology described in ISO 18213-1 and to appropriate related algorithms in ISO 18213-2, ISO 18213-4, ISO 18213-5 or ISO 18213-6. Although the methodology in this part of ISO 18213 is intended for application specifically within the context of the other parts of ISO 18213, the methods are more widely applicable. In particular, the statistical model presented in Clause 6 for estimating the tank's measurement equation from a set of standardized calibration data can be applied, regardless of whether or not these data are acquired in accordance with the methods of ISO 18213. A similar statement holds for (propagation) methods of variance estimation: it is intended that the results in this part of ISO 18213 be applied to the specific models for which they were derived, but the methods themselves are more widely applicable.

This part of ISO 18213 provides a facility with the option to develop equivalent plant- or tank-specific methods of statistical analysis as an alternative. However, if a facility adopts ISO 18213 and chooses not to develop equivalent alternative methods of statistical analysis, it is necessary to use the methods of this part of ISO 18213.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 18213-1:2007, *Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy — Part 1: Procedural overview*

ISO 18213-4:2008, *Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy — Part 4: Accurate determination of liquid height in accountancy tanks equipped with dip tubes, slow bubbling rate*

ISO 18213-5:2008, *Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy — Part 5: Accurate determination of liquid height in accountancy tanks equipped with dip tubes, fast bubbling rate*

ISO 18213-6:2008, *Nuclear fuel technology — Tank calibration and volume determination for nuclear materials accountancy — Part 6: Accurate in-tank determination of liquid density in accountancy tanks equipped with dip tubes*

3 Symbols

The symbols used in this part of ISO 18213 are defined below. The symbols are listed in the first column of the table, approximately in order of appearance. Some symbols are introduced in groups, such as in connection with a particular equation. The ordering of symbols within such a group may differ from their appearance in the text if doing so makes the information easier to use. The location at which each symbol first appears is given in the corresponding row of the second column. The definition or usage of each symbol is presented in the third column.

Symbol	First reference	Definition/Usage
Y	5.2.1	response variable (either height or volume, height by convention)
X	5.2.1	control variable (either volume or height, volume by convention)
i	5.2.2	subscript that denotes either calibration increment number or observation number
Y_i	5.2.2	standardized elevation of a point in the tank above some pre-established reference point, typically associated with the standardized volume determined from the liquid added during the first i increments of a calibration run
X_i	5.2.2	standardized volume of the tank determined from the total volume of liquid added during the first i calibration increments, i.e., the standardized volume of the tank below Y_i
j	5.2.2	subscript
x_j	5.2.2	standardized volume of the jth increment of calibration liquid
(X_i, Y_i)	5.2.2	standardized volume-height data pair for the ith calibration increment
f or $f(...)$	5.2.2	generic function, the tank calibration equation, by convention
$\hat{Y} = \alpha + \beta X + \varepsilon$	5.2.3	equation that expresses height as a linear function of volume
α, β	5.2.3	equation parameters
ε	5.2.3	residual (height), the difference between the observed value of the response variable (Y) and the corresponding predicted value ($\alpha + \beta X$), $Y - \alpha - \beta X$
a, b	5.2.3	estimates of α, β
$\hat{Y}$	5.2.3	predicted response (height by convention) derived from some functional relationship between height and volume, $\hat{Y} = a + bX$
$Y_i - a - bX_i$	5.2.3	estimated residual, the estimated difference between observed and estimated values of the response variable for the ith calibration increment, $Y_i - \hat{Y}_i$

$Y = f(X)$	5.2.3	tank calibration equation
Δ	5.2.4	difference operator
ΔY	5.2.4	change (difference) in the response variable (height), typically between two calibration increments
ΔX	5.2.4	change (difference) in the control variable (volume), typically between two calibration increments
m_i	5.2.4	computed slope (change in height per unit change in volume) of calibration equation for the ith calibration increment, $\Delta Y_i / \Delta X_i$
f_1, f_2	5.2.5.1	generic functions, typically used to denote calibration equations or segments thereof
$\hat{f}$	5.2.5.2	estimate of the function f, the estimated calibration equation by convention
T_i	5.3.1	temperature, in either the tank or the prover, of the ith increment of calibration liquid
t_i	5.3.2	time associated with the ith calibration increment, e.g., time at start of increment
Δt_i	5.3.2	time required to complete the ith calibration increment, $t_i - t_{i-1}$
T_m	6.1, Eq. (6)	measured temperature of tank liquid
T_r	6.1, Eq. (6)	reference temperature established for calibration
H_M	6.1, Eq. (6)	height of a point in the tank at measured temperature T_m
H_r	6.1, Eq. (6)	height of a point in the tank at reference temperature T_r
ΔP	6.1, Eq. (6)	observed difference in pressure between the submerged bubbling probe and the reference probe
c_M	6.1, Eq. (6)	"corrections" that compensate for differences between the observed pressure at the manometer and the actual pressure at the tip of the submerged probe
ρ_M	6.1, Eq. (6)	average density of the liquid in the tank at the measured temperature T_m
$\rho_{\mathrm{a},s}$	6.1, Eq. (6)	average density of the air in the tank above the liquid surface at the prevailing pressure
g	6.1, Eq. (6)	local value of the acceleration due to gravity
α_ex	6.1, Eq. (6)	coefficient of linear thermal expansion for the dip tubes
ΔT_m	6.1, Eq. (6)	difference between the measured and reference temperatures, $T_\mathrm{m} - T_\mathrm{r}$
var(...)	6.1	variance operator, e.g., **var**(H_r) denotes the variance of H_r and **var**(ΔP) denotes the variance of ΔP, etc.
$\hat{f}^{-1}$	6.1	estimate of f^{-1}
f^{-1}	7.1	inverse of f, the measurement equation, by convention

$H = f(V)$	7.1	generic expression for the calibration equation
$V = f^{-1}(H)$	7.1	generic expression for the measurement equation
h or $h(...)$	7.2.1, Eq. (10)	generic function, f^{-1}, by convention
ε	7.2.1, Eq. (10)	residual, the difference between the observed value of the response variable (Y) and the corresponding predicted value $h(X)$, $Y - h(X)$
ε_i	7.2.1	residual difference between the observed value of the response variable (Y_i) and the corresponding predicted value $h(X_i)$ for the ith calibration increment, $Y_i - h(X_i)$
$\hat{h}$	7.2.1	estimate of h, typically $\hat{h} = \hat{f}^{-1}$
s	7.2.1, Eq. (11)	subscript
c_s	7.2.1, Eq. (11)	"cut point," point in the (height) range of the measurement equation
S	7.2.1, Eq. (11)	number of segments (intervals) into which the range of the measurement equation is partitioned by cut points
c_0	7.2.1, Eq. (11)	left-hand endpoint of the first segment, usually 0
c_S	7.2.1, Eq. (11)	the right-hand endpoint of the largest segment, usually the largest value of the control variable, i.e., $c_S = X_{max}$
h_s	7.2.1, Eq. (12)	function defined over the interval (c_{s-1}, c_s), i.e., function defined for values between c_{s-1} and c_s, where s ranges from 1 to S
β_i	7.2.1	model parameters (β_0 denotes the intercept)
n	7.2.1, Eq. (16)	total number of observations, i.e., total number of height-volume data pairs (X_i, Y_i)
$p + 1$	7.2.1, Eq. (16)	number of parameters in the specified model
Y	7.2.1, Eq. (16)	$n \times 1$ vector of (response variable) observations
H	7.2.1, Eq. (16)	$n \times (p + 1)$ design matrix
β	7.2.1, Eq. (16)	$(p + 1) \times 1$ vector of model parameters
ε	7.2.1, Eq. (16)	$n \times 1$ vector of residual differences, i.e., $n \times 1$ vector of fitting errors
$\sigma(\sigma^2)$	7.2.1	standard deviation (variance) of the components of **ε**
h_1, h_2, h_3	7.2.1	generic functions
θ	7.2.2	$(p + 1) \times 1$ vector of perturbations to the vector of model parameters, **β**
θ$_j$	7.2.2	$(p + 1) \times 1$ vector of perturbations to the vector of model parameters, **β**, attributable to the jth run
$\theta_{j,k}$	7.2.2	kth component of **θ**$_j$
β$_j$	7.2.2	$(p + 1) \times 1$ vector of model parameters for the jth run, $\boldsymbol{\beta}_j = \boldsymbol{\beta} + \boldsymbol{\theta}_j$

4

$\beta_{j,k}$	7.2.2	kth component of $\boldsymbol{\beta}_j$
$\mathbf{E}(\ldots)$	7.2.2	expectation operator
$\boldsymbol{\theta}'$	7.2.2, Eq. (18)	transpose of the vector $\boldsymbol{\theta}$
$\boldsymbol{\Phi}^2$	7.2.2, Eq. (18)	variance-covariance matrix of (the components of) $\boldsymbol{\theta}$
$(X_{j,i},\ Y_{j,i})$	7.2.2	ith standardized height-volume data pair from jth calibration run
n_j	7.2.2, Eq. (19)	total number of observations, i.e., total number of height-volume data pairs $(X_{j,i},\ Y_{j,i})$, from jth calibration run
$\mathbf{Y}_j$	7.2.2, Eq. (19)	$n_j \times 1$ vector of (response variable) observations from the jth calibration run
$\mathbf{H}_j$	7.2.2, Eq. (19)	$n_j \times (p + 1)$ design matrix for the jth calibration run
$\boldsymbol{\varepsilon}_j$	7.2.2, Eq. (19)	$n_j \times 1$ vector of residual differences (fitting errors) for the jth calibration run
r	7.2.2, Eq. (19)	number of calibration runs
$\mathbf{H}'_j$	7.2.2, Eq. (21)	transpose of the matrix $\mathbf{H}_j$
$\sigma_j\left(\sigma_j^2\right)$	7.2.2, Eq. (21)	standard deviation (variance) of the components of $\boldsymbol{\varepsilon}_j$
$\mathbf{I}$	7.2.2, Eq. (21)	$n_j \times n_j$ identity matrix
$Y_{j,i}$	7.2.2, Eq. (22)	ith component of the vector $\mathbf{Y}_j$
$\mathbf{h}'_{j,i}$	7.2.2, Eq. (22)	ith row of the design matrix $\mathbf{H}_j$
$\hat{\boldsymbol{\beta}}_j, \hat{\boldsymbol{\varepsilon}}_j^2, \hat{\sigma}_j^2,$ etc.	7.3.2	respective estimators of $\boldsymbol{\beta}_j, \boldsymbol{\varepsilon}_j^2, \sigma_j^2$, etc.
$\varepsilon_{j,i}$	7.3.2	ith component of $\boldsymbol{\varepsilon}_j$
$\hat{h}_j$	7.3.2, Eq. (26)	estimated measurement equation from data of the jth calibration run
$\hat{\boldsymbol{\beta}}$	7.3.3.1	estimator of $\boldsymbol{\beta}$
X_0	7.3.3.2	specified (unobserved) value of the control variable (volume by convention)
Y_0	7.3.3.2	value of the response variable (height) at X_0
$\mathbf{h}'_0$	7.3.3.2	row of the design matrix, $\mathbf{H}$, that corresponds to X_0
$\hat{\bar{Y}}_0$	7.3.3.2	predicted (mean) value of the response variable Y_0 at X_0, $\mathbf{h}'_0\hat{\boldsymbol{\beta}}$
$\hat{\boldsymbol{\Phi}}^2, \hat{\boldsymbol{\theta}}_j$	7.3.3.3	respective estimators of $\boldsymbol{\Phi}^2$ and $\boldsymbol{\theta}_j$
$\mathbf{v\hat{a}r}(\ldots)$	7.3.3.3	estimated variance, e.g., $\mathbf{v\hat{a}r}\left(\hat{\boldsymbol{\beta}}\right)$ denotes the estimated variance of $\hat{\boldsymbol{\beta}}$, etc.
$\hat{\sigma}^2$	7.3.3.4	estimator of σ^2
$\hat{\boldsymbol{\varepsilon}}_j$	7.3.3.4	estimator of $\boldsymbol{\varepsilon}_j$

$\hat{\varepsilon}_{j,i}$	7.3.3.4	ith component of $\hat{\boldsymbol{\varepsilon}}_j$, estimate of the ith component of $\boldsymbol{\varepsilon}_j$
n	7.3.3.4	total number of observations from all runs, $\sum_j n_j$
ε_0	7.4, Eq. (37)	prediction error for a new (future) value of Y_0, $\varepsilon_0 = Y_0 - h(X_0) = Y_0 - \mathbf{h}'_0(\boldsymbol{\beta} + \boldsymbol{\theta})$
$\hat{Y}_0$	7.4	estimated (predicted) value of Y_0
α	7.5.2.1, Eq. (40)	specified confidence level (typically 0,025 or 0,05)
$\hat{\sigma}_0\left(\hat{\sigma}_0^2\right)$	7.5.2.1, Eq. (40)	estimated standard deviation (variance) of $\hat{\bar{Y}}_0$, given by Equation (35)
ν	7.5.2.1, Eq. (40)	(approximate) degrees of freedom for the variance estimate $\hat{\sigma}_0^2$
$t_{\alpha/2}(\nu)$	7.5.2.1, Eq. (40)	$100(1 - \alpha/2)$ % point from the t-distribution with parameter (degrees of freedom) ν
S_1^2, S_2^2	7.5.2.1	quantities used to compute degrees of freedom, ν
ν_1, ν_2	7.5.2.1	component degrees of freedom
V, W	7.5.2.1	quantities used to compute degrees of freedom, ν
X	7.5.2.2	an arbitrary unspecified value of the control variable (volume)
$\hat{\bar{Y}}_X$	7.5.2.2	predicted (mean) value of the response variable (height) at X, $\mathbf{h}'_X\hat{\boldsymbol{\beta}}$
$\mathbf{h}'_X$	7.5.2.2	row of the design matrix $\mathbf{H}$ that corresponds to X
$\hat{\sigma}_X\left(\hat{\sigma}_X^2\right)$	7.5.2.2	estimated standard deviation (variance) of $\hat{\bar{Y}}_X$
ν	7.5.2.2	(approximate) degrees of freedom for the variance estimate, $\hat{\sigma}_X^2$
$F_\alpha(p + 1, \nu)$	7.5.2.2	$100(1 - \alpha)$ % point from the F-distribution with parameters $(p + 1)$ and ν
$\hat{\boldsymbol{\beta}}_{\text{new}}$	7.5.2.3	estimator of $\boldsymbol{\beta}$ computed from the data of a new calibration run
$\hat{\bar{Y}}_{X,\text{new}}$	7.5.2.3	predicted (mean) value of the response variable (height) at X obtained from the new calibration equation, $\mathbf{h}'_X\hat{\boldsymbol{\beta}}_{\text{new}}$
$\hat{\sigma}_{X,\text{new}}\left(\hat{\sigma}_{X,\text{new}}^2\right)$	7.5.2.3	estimated standard deviation (variance) of $\hat{\bar{Y}}_{X,\text{new}}$
ν_1, ν_2	7.5.2.3	(approximate) degrees of freedom for components of the variance estimate $\hat{\sigma}_X^2$
$\hat{\sigma}_X(\hat{\sigma}_X^2)$	7.5.3.2	estimated standard deviation (variance) of $\hat{Y}_X$
H_0	8.2.1	a standardized reference height (at reference temperature T_r), $H_0 = X_0$
V_0	8.2.1	standardized reference volume that corresponds to the height $H_0 = X_0$, $V_0 = Y_0$
$\mathbf{var}\left(V_{0,\text{pred}}\right)$	8.2.1, Eq. (55)	variance of the (mean) predicted volume obtained from the measurement equation at H_0
$\mathbf{var}\left(V_{0,\text{new}}\right)$	8.2.1, Eq. (55)	measurement component of variance of a (new) volume determination at H_0

$\mathbf{var}(V_{0,\text{trans}})$	8.2.1, Eq. (55)	component of variance of a (new) predicted volume resulting from the "transfer" of uncertainty in H_0 through the measurement equation
$\partial \hat{h}(H_0)/\partial(H_0)$	8.2.1, Eq. (57)	derivative of the estimated measurement equation $\hat{h}$, taken with respect to H and evaluated at $H = H_0$
V_M	8.2.1	the volume at temperature T_m of the standardized reference volume, V_0
T_3	8.2.2.2	specified temperature
V_3, ρ_3	8.2.2.2	respective density and volume, at temperature T_3, of a liquid that has volume V_M and density ρ_M at temperature T_m
V_1, V_2	8.3	specified standardized volumes
ΔV	8.3	difference between two specified volumes, $V_1 - V_2$
$\mathbf{h}'_1, \mathbf{h}'_2$	8.3	row vectors that correspond respectively to the standardized height H_1 and H_2

4 Data required

This part of ISO 18213 applies generally to data acquired during the process of data collection and analysis for tank calibration and volume determination as outlined in ISO 18213-1. Specific procedures apply either to particular subsets of these data or at various stages in the process pertaining to their acquisition, analysis, interpretation and use. The data to which a particular statistical procedure applies and the stage in the process at which the procedure should be used are identified in the subclause(s) where that procedure is discussed.

5 Diagnostic plots

5.1 Overview

Diagnostic plots are among the most powerful tools available for analyzing and verifying volume measurement data. Plots are particularly useful for identifying anomalous observations and measurements in a set of tank calibration data. They are also quite useful for comparing the (standardized) data from two calibration runs and for comparing two estimates of the tank's calibration equation. Plots that may be used to evaluate a set of calibration (height and volume) data are presented in 5.2. Plots of auxiliary data (time, temperature) are presented in 5.3. Examples of all plots are given in Annex A.

5.2 Calibration data

5.2.1 General comments

The plots discussed in 5.2 may be based on either the tank's calibration equation or its measurement equation. For plots based on the calibration equation, the response variable (denoted by Y) represents the height or elevation and the control variable (denoted by X) represents volume or increment number. For plots based on the measurement equation (the inverse of the calibration equation), the interpretation of Y and X is reversed: Y denotes volume and X denotes height. Both plot orientations can be useful in a particular application and both are illustrated in Annex A. For convenience, only the term "calibration equation" is used in 5.2, with the understanding that the discussion also applies to the measurement equation.

The plots of 5.2 can be constructed from data that are presented in various forms. A specific plot is typically constructed from the standardized data from a particular calibration run, but it can also be constructed from the corresponding raw data, or from "data points" obtained by evaluating the tank's estimated calibration (or measurement) equation at a number of selected points. Although it is possible to construct plots from raw calibration data, the analysis of raw data is generally not recommended because meaningful comparisons are difficult, particularly if measurement conditions vary significantly during the period(s) of data collection.

Finally, it is often useful for comparative purposes to superpose or overlay several plots on a set of common axes. A plot obtained by superposing profile variation plots of the calibration data from several individual runs, for example (see 5.2.3), is very useful for examining run-to-run variations. Likewise, overlaying a profile variation plot of the data from a new calibration run on that from a previous estimate of the calibration equation can be very helpful for verifying that the tank's calibration equation has not changed since the previous calibration.

5.2.2 Cumulative (Y vs. X) plots

A cumulative plot displays the relationship between the height or elevation of points in a tank above some pre-selected reference point, Y, and the corresponding volume of the tank, X, below these points. A cumulative plot shows the general features (shape) of the height-volume relationship for the tank.

As noted in 5.2.1, a cumulative plot can be constructed

a)   from the standardized data from a particular calibration run,

b)   from the corresponding raw data (generally not recommended), or

c)   from a previously defined calibration equation for the tank, expressed in functional or tabular form.

In the first two cases, the response of the tank's measurement system, e.g., "liquid height," is plotted for each increment of the calibration run against a measure of the total amount of calibration liquid, e.g., cumulative volume, required to reach this height. In the latter case, the plotted "data" are obtained by evaluating the tank's estimated calibration or measurement equation at a number of selected points.

In the notation of ISO 18213-2, Y_i denotes the (standardized) elevation, above some pre-established reference point, of a point in the tank determined by the liquid added during the first i increments of a calibration run. Similarly, X_i denotes the (standardized) total volume of the tank below that point, as determined from the volume of liquid added during the first i calibration increments, as given in Equation (1):

$$X_i = \sum_j x_j \tag{1}$$

where x_j denotes the standardized volume of the jth incremental addition of calibration liquid. A cumulative plot is obtained by plotting the standardized volume-height pairs (X_i, Y_i) derived from the raw data of a particular calibration run. Methods for computing the standardized values X_i and Y_i from a set of raw calibration data are described in ISO 18213-2.

It can be useful for comparative purposes to overlay plots of the standardized data from several calibration runs. It is also possible to include a cumulative plot derived from a previously defined calibration equation, f, in the overlay plot. This is done by plotting the points $[X_i, f(X_i)]$ obtained by evaluating the function at a suitable number of points.

One variation of the cumulative plot (and all other plots discussed in 5.2) is to plot the response variable against increment number, i, instead of against cumulative volume, X_i. However, when several plots are overlaid, a valid comparison is possible only when all data are plotted on a common scale.

Cumulative plots show the general features of the tank's profile (i.e., its height-volume relationship). They can also reveal gross differences in the data from several calibration runs, or among the data of several calibration runs and some pre-established calibration equation. However, the plotting range on the vertical scale is generally too large to provide adequate resolution for detecting

—   small differences in tank profile, or

—   outlying points in a set of calibration data.

Variations in tank profile and anomalous data points are more easily detected with the aid of the profile variation and incremental slope plots discussed in 5.2.3 and 5.2.4, respectively.

5.2.3   Profile variation [$(Y - a - bX)$ vs. X] plots

A profile variation plot shows the difference between the observed height, Y, and an estimate of height computed from an equation that expresses height as a linear function of volume, $\hat{Y} = a + bX$, versus the volume, X, of the tank below the corresponding height. In other words, the profile variation plot shows the variation in the free (unobstructed) cross-sectional area of the tank about its average free cross-sectional area. A profile variation plot provides greater resolution in the vertical ("height") scale than the cumulative plot, thereby revealing greater detail about the free cross-sectional area of the tank.

Like the cumulative plot, the profile variation plot can be constructed from

a)   the standardized data from a particular calibration run,

b)   the corresponding raw data (generally not recommended), or

c)   a previously defined calibration equation for the tank, expressed in a functional form.

In the first two cases, a profile variation plot is obtained by plotting, for each increment of a calibration run, the "residual height" against the corresponding cumulative volume. The residual heights are the differences between observed heights and corresponding estimates computed from a linear function chosen to describe or "fit" the relationship between height and cumulative volume in the selected data. In the later case, the plotted "data" are obtained by first evaluating the tank's estimated calibration or measurement equation at a number of selected points.

In the notation of 5.2.2, where Y denotes standardized (liquid) height and X denotes the corresponding standardized cumulative volume, the profile variation plot is obtained by plotting the following points:

$$(Y_i - \hat{Y}_i, X_i) = (Y_i - a - bX_i, X_i) \tag{2}$$

for all pairs of observations (X_i, Y_i) obtained during the calibration run. In Equation (2), a and b are estimates of the coefficients α and β in the linear relationship $\hat{Y} = \alpha + \beta X + \varepsilon$ employed to describe the data.

The main objective of the profile variation plot is to increase the resolution in the vertical (height) dimension, and the method used to estimate α and β is secondary to this objective. The coefficients a and b may be taken as least squares regression estimates of the intercept and slope for a straight-line fit to the calibration data (X_i, Y_i). Alternatively, a and b may be taken as the slope and intercept of a line that passes through some initial point, e.g., the second or third, and a terminal point, e.g., the next-to-last, of the run. It is generally advisable to avoid the first and last points because they tend to be more anomalous than other points in the run.

As with cumulative plots, it can be useful for comparative purposes to overlay profile variation plots of the standardized data from several calibration runs. It is possible to include a profile variation plot from a previously defined calibration equation, f, in the overlay plot by evaluating the function $Y = f(X)$ at suitable points, X_i, and computing $Y_i - a - bX_i$ for each. It is also possible to make profile plots from the raw data, but this is not recommended for the reasons cited in 5.2.1.

When data from several calibration runs or tank calibration equations are being compared, the linear coefficients a and b should be determined from the aggregated data from all runs or equations of interest to ensure that all data are plotted on a common scale.

5.2.4   Incremental slope ($\Delta Y / \Delta X$ vs. X) plots

The incremental slope plot displays the incremental changes in the slope of the calibration function, i.e., changes in height between successive calibration increments with respect to the corresponding incremental changes in volume, plotted relative to the volume of the tank below the associated height. In other words, an incremental slope plot displays the rate of change in liquid height in the tank per unit change in volume for each volume increment in a calibration run. Incremental slope plots reveal great detail, so they are very useful for detecting small changes in tank profile that would not be revealed by cumulative or profile variation plots.

Consequently, incremental slope plots are especially useful for detecting outliers and other small data anomalies.

As with previously discussed plots, the incremental slope plot can be constructed from

a) the standardized data from a particular calibration run,

b) the corresponding raw data (generally not recommended), or

c) a previously defined calibration equation for the tank, expressed in a functional form.

If Y denotes standardized height and X denotes standardized cumulative volume, the incremental slope, m_i, i.e. the change in height per unit change in volume, for the ith calibration increment is given in Equation (3):

$$
\begin{aligned}
m_i &= \Delta Y_i / \Delta X_i \\
&= \left(Y_i - Y_{i-1}\right) / \left(X_i - X_{i-1}\right) \\
&= \left(Y_i - Y_{i-1}\right) / x_i
\end{aligned}
\tag{3}
$$

where (X_{i-1}, Y_{i-1}) and (X_i, Y_i) are the cumulative volumes and heights, respectively, from the $(i-1)$th and ith calibration increments. The incremental slope plot is created by plotting m_i versus X_i for the selected calibration increments.

It can be useful for comparative purposes to overlay several incremental slope plots. Incremental slope plots can be overlaid to compare standardized data from the same dip tube acquired over several calibration runs. They can also be overlaid to compare data from several dip tubes acquired during a single run. Incremental slope plots of data from different runs can be especially helpful for detecting subtle changes in profile, whereas incremental slope plots of data from several probes collected during the same calibration run are helpful for detecting anomalous measurements. As with other plots, it is possible to include an incremental slope plot from a previously defined calibration equation in the overlay plot by evaluating the function $Y = f(X)$ at suitable points X_i.

5.2.5 Comparison and residual plots

5.2.5.1 Comparison $\{[Y - f(X)]$ vs. X or $[f_1(X) - f_2(X)]$ vs. $X\}$ plots

As its name implies, a comparison plot is a graphical comparison between two equations or two sets of data. Comparison plots are typically used to display differences between a set of calibration data and some reference function, such as the tank's previously estimated calibration equation. Comparison plots can also be used to display

a) differences between data from two calibration runs,

b) differences between two functions (e.g., f_1 and f_2), such as new and old estimates of the calibration equation, or

c) estimates of the calibration equation derived by the operator and the inspector.

It is often desirable to select one of the entities as a reference for the comparison.

As in 5.2.2, let Y denote the (standardized) liquid height and X denote the (standardized) cumulative volume. Then the plot for comparing a new set of calibration data with some function, f, of interest, such as a previously determined calibration equation, is obtained by plotting the following points:

$$
\left[Y_i - f\left(X_i\right), X_i\right]
\tag{4}
$$

for all pairs of observations (X_i, Y_i) obtained during the calibration run. It is clear from Equations (2) and (4) that the profile plot is a special case of the comparison plot in which the function, f, has the linear form $f(X) = \alpha + \beta X$. Moreover, the residual plot discussed in 5.2.5.2 is a special case of the comparison plot in which the reference function, f, is estimated from the data (X_i, Y_i) by statistical methods.

For examining the data from several calibration runs, it is particularly useful to overlay several plots in which the data from each run are compared with a common reference function, such as a previously determined estimate of the calibration equation. A plot of the differences between estimates of the tank's calibration equation from successive calibration exercises shows the extent to which the calibration has changed over time. Small differences indicate, for example, that the calibration equation has not changed significantly between calibrations. It is possible to construct various comparison plots that are quite useful for analysing tank calibration data. However, it is also easy to create plots that are confusing and difficult to interpret, so care is required to avoid useless or misleading plots.

5.2.5.2 Residual {[($Y - \hat{f}(X)$)] vs. X} plots

A residual plot is a special type of comparison plot that differs from a general comparison plot only in how the reference function is determined. Whereas any function may be used to make a comparison plot, the term residual plot is reserved for the case in which the reference function is obtained by "fitting" the data of interest.

The construction of a residual plot is identical to that of a comparison plot. If Y denotes standardized liquid height and X denotes standardized cumulative volume, then the residual plot for a set of calibration data is obtained by plotting the following points:

$$\left[Y_i - \hat{f}(X_i), X_i \right] \tag{5}$$

for all pairs of observations (X_i, Y_i) obtained during the calibration run. In this case, the function, $\hat{f}$, is estimated from the calibration data by means of some statistical fitting procedure such as, for example, least-squares regression.

Residual plots are used primarily for evaluating the "goodness" of various trial fits to a set of standardized calibration data in the search for a suitable estimate of the tank's calibration equation. Since residual plots show differences between the observed data points and the proposed estimate, they play a key role in the model fitting process of 7.3. When a model is fitted simultaneously to the data of several calibration runs, the corresponding residual plot also shows the extent to which run-to-run variation influences the resulting uncertainty estimates.

5.3 Auxiliary data

5.3.1 Temperature plots

A temperature plot displays the temperature associated with each calibration increment on the vertical scale versus increment number on the horizontal scale. The temperature of interest may be either that of the calibration liquid in the prover or the liquid in the tank. If T_i denotes the temperature of interest for the ith calibration increment, then the temperature plot is a plot of the values T_i versus i.

Depending upon which data are selected, a temperature plot provides an easy way to examine changes in the temperature of liquid in either the prover or the tank that occurred during a calibration run. Moreover, an overlay plot that shows, for each increment, the temperature of the liquid both in the prover and in the tank is convenient for determining the extent to which liquid temperature in the tank differed from that in the prover at any time during the run.

Temperature plots are very useful for examining the data collected during a single calibration run. Overlay plots of temperatures from several calibration runs are useful for evaluating the overall temperature variation during a calibration exercise.

5.3.2 Inter-increment time plots

The inter-increment time plot is a plot of the elapsed time between successive calibration increments versus increment number. If t_i denotes the time at which the ith calibration increment is completed, then the inter-increment time plot is simply a plot of the differences $\Delta t_i = t_i - t_{i-1}$ versus i. This plot provides an easy way to identify any excessive delays between successive increments of the calibration run. Only data from a single calibration run are used to construct an inter-increment time plot.

6 Uncertainty estimation for calibration data

6.1 Measurement system response (height)

Methods for estimating the uncertainty of measurements made with the tank's measurement system (manometer) are presented in this subclause. In particular, a formula is presented for estimating the uncertainty of height determinations derived from pressure measurements made with a tank's pneumatic system.

Expressions are given in ISO 18213-4:2008, Equations (7) to (9), or ISO 18213-5:2008, Equations (8) and (9), for determining H_r, the elevation of a point in the tank above a fixed reference point at some pre-determined reference temperature. A simplified form of that expression, in which H_r is given in terms of the pressure exerted by a column of liquid in the tank at the tip of a submerged dip tube at reference temperature, T_r, of a point in the tank whose height at the measurement temperature T_m was H_M, has the form given in Equation (6):

$$H_r = \left(\Delta P - c_M\right)\Big/\left[g\left(\rho_M - \rho_{a,s}\right)\left(1 + \alpha_{ex}\Delta T_m\right)\right] \tag{6}$$

where

ΔP　　　　　is the observed difference in pressure between the submerged bubbling probe and the reference probe that vents into the tank above the liquid surface;

c_M　　　　　denotes the "corrections" that compensate for differences between the observed pressure at the manometer and the actual pressure at the tip of the probe;[1]

ρ_M　　　　　is the average density of the liquid in the tank at the measured temperature T_m;

$\rho_{a,s}$　　　　is the average density of the air in the tank above the liquid surface at the prevailing pressure;

g　　　　　is the local value of the acceleration due to gravity;

α_{ex}　　　　is the coefficient of linear thermal expansion for the dip tube;

$\Delta T_m = T_m - T_r$　　is the difference between the measured and reference temperatures.

It is now possible to express the variance **var**(H_r) of the height determination, H_r, in terms of the quantities on the right-hand side of Equation (6). Each quantity is considered in turn. In practice, suitable estimates of these quantities are used to compute an estimate of **var**(H_r).

---

1)　It is not the purpose of this part of ISO 18213 to present a detailed discussion of the corrections denoted by c_M, except as they apply to variance estimation. A detailed development of Equation (6) is presented in ISO 18213-4:2008 for a slow bubbling rate and in ISO 18213-5:2008 for a fast bubbling rate.

At a given location, the acceleration due to gravity is constant. Consequently, any error in estimating this quantity appears as a bias in the determination of H_r, but does not otherwise contribute significantly to its variance. The factor, g, is therefore treated as a constant[2].

Similarly, the thermal correction factor, $(1 + \alpha_{ex}\Delta T_m)$, is retained as a bias correction, but is taken as a constant for variance estimation. Except in extreme cases, the contribution of the variability in this quantity to the total variability, **var**(H_r), is negligible because variability in $(1 + \alpha_{ex}\Delta T_m)$ due to temperature measurement variability is quite small. If it is further assumed that the quantities ΔP, c_M, ρ_M and $\rho_{a,s}$ are independent, then, to a first-order approximation, the variance of H_r can be expressed in terms of quantities on the right-hand side of Equation (6) as given in Equations (7) and (8):

$$\mathbf{var}\left(H_r\right) = \left[g\left(1+\alpha_{ex}\Delta T_m\right)\right]^{-2}\left[\left(\Delta P - c_M\right)\big/\left(\rho_M - \rho_{a,s}\right)\right]^2\left[\mathbf{var}\left(\Delta P - c_M\right)\big/\left(\Delta P - c_M\right)^2 + \mathbf{var}\left(\rho_M - \rho_{a,s}\right)\big/\left(\rho_M - \rho_{a,s}\right)^2\right] \tag{7}$$

or

$$\mathbf{var}\left(H_r\right) = \left(H_r\right)^2\left[\mathbf{var}\left(\Delta P - c_M\right)\big/\left(\Delta P - c_M\right)^2 + \mathbf{var}\left(\rho_M - \rho_{a,s}\right)\big/\left(\rho_M - \rho_{a,s}\right)^2\right] \tag{8}$$

In most cases, the contribution to **var**(H_r) of variability in the correction terms, c_M, is quite small, being on the order of 10^{-3} or less for each measurement. Similarly, the contribution of variability in $\rho_{a,s}$ to the total variability of H_r is typically quite small. If the quantities **var**(c_M) and **var**$(\rho_{a,s})$ are ignored [2], then Equation (8) takes on the form of Equation (9):

$$\mathbf{var}\left(H_r\right) = \left(H_r\right)^2\left[\mathbf{var}\left(\Delta P\right)\big/\left(\Delta P - c_M\right)^2 + \mathbf{var}\left(\rho_M\right)\big/\left(\rho_M - \rho_{a,s}\right)^2\right] \tag{9}$$

The following considerations apply to the estimation of the quantities in Equation (9).

— The standardized height, H_r, is computed from ΔP by means of either ISO 18213-4:2008, Equations (7) to (9) (for a slow bubbling rate), or ISO 18213-5:2008, Equations (8) and (9) (for a fast bubbling rate), as appropriate.

— ΔP is observed. Ideally **var**(ΔP) is estimated from replicated observations of ΔP made under typical measurement conditions. Alternatively, **var**(ΔP) is estimated from a statistical characterization of the properties of the instrument (manometer) used to measure ΔP (as supplied, for example, by the vendor and verified at the facility). The alternative is less desirable because an instrument is usually characterized under carefully controlled conditions that do not reflect actual measurement conditions in the facility where it will be used.

— The corrections denoted by c_M are computed by means of either ISO 18213-4:2008, Equations (7) to (9) (for a slow bubbling rate), or ISO 18213-5:2008, Equations (8) and (9) (for a fast bubbling rate), as appropriate.

— The determination of the density, ρ_M, depends on the liquid being measured. For calibration liquids that have been well characterized (e.g., demineralized water), the density may be determined either numerically or from pre-established tables (see ISO 18213-1:2007, 6.5.4). It is possible, for example, to obtain highly accurate values of the density of water from measurements of its temperature by means of the equation given in ISO 18213-6:2008, Clause 3.

NOTE This equation is also given in ISO 18213-2:2007, Annex A; ISO 18213-4:2008, Annex A; and ISO 18213-5:2008, Annex A.

---

2) The simplifying assumptions made here are generally acceptable under normal operating conditions. However, it is necessary to verify any assumptions for each application. If an assumption seems questionable or it cannot be verified in a particular situation, then it is necessary to retain the relevant terms in subsequent calculations.

For a process liquid that is not well characterized, its density may be determined either analytically (in the laboratory) or by means of in-tank measurements. If the density is reported by the laboratory at some standard reference temperature, then a suitable interpolation equation is required to determine the density of the liquid at its in-tank measurement temperatures. A method is given in ISO 18213-6 for computing accurate density estimates for process liquids at in-tank measurement temperatures. In the absence of a good working knowledge of how the density of the measured process liquid changes with temperature, it is possible that estimates derived from in-tank measurements are actually superior to those obtained from the laboratory. However, the accuracy of in-tank estimates depends critically on the quality of the probe-separation determination used in their computation. A procedure for making very precise determinations of probe separation is presented in ISO 18213-6:2008, 6.2.

— $\mathbf{var}(\rho_M)$ is obtained as appropriate, depending on how ρ_M is determined. If ρ_M is determined analytically, then the laboratory is expected to provide an estimate of the uncertainty for the density value it reports. For a process liquid, where the density is estimated from in-tank measurements, the variance of ρ_M can be estimated as given in ISO 18213-6. This estimate is further considered in 8.2.1, where an equation is given for estimating the variability of a volume determination in terms of the variability of the corresponding standardized height, H_r, and the estimated measurement equation, $V_r = \hat{f}^{-1}(H_r)$.

— The air density, $\rho_{a,s}$, is computed from measurements of prevailing atmospheric conditions as indicated in either ISO 18213-4:2008, Clause A.2, or ISO 18213-5:2008, Clause A.2. In the absence of direct measurements, $\rho_{a,s}$ may alternatively be estimated for "standard" or "typical" atmospheric conditions. The use of direct measurements is preferable, but differences between the two alternatives are typically quite small and the effect on Equation (9) is negligible (see either ISO 18213-4:2008, Clause A.2, or ISO 18213-5:2008, Clause A.2, for details). If an estimate of $\mathbf{var}(\rho_{a,s})$ is required, it may be derived in either case from Equation (A.3) in either ISO 18213-4:2008 or ISO 18213-5:2008 with the aid of propagation-of-variance methods.

6.2 Measurements of tank content (volume, mass)

As with manometer measurements of tank content made with the tank's pneumatic system, it is possible to estimate the uncertainties of independent measurements of tank content made with the prover. In practice, the uncertainties for determinations of tank content made with provers, either volumetric or gravimetric, are generally small enough to be ignored in subsequent uncertainty calculations. Because uncertainties are small and calculations are comparatively complex, equations for estimating the uncertainties associated with prover measurements are not presented in ISO 18213 (all parts).

7 Estimation of the measurement equation and associated uncertainties

7.1 Preliminaries

The height-volume relationship for a tank embodied in the calibration model of ISO 18213-1:2007, Clause 2, is too complex to compute from engineering drawings with sufficient accuracy for safeguards accountability purposes. Therefore, a calibration exercise is conducted to obtain data for "calibrating the tank", that is, for estimating either the tank's calibration equation or measurement equation. The twofold purpose of this clause is

a) to present a statistical model for estimating a tank's calibration or measurement equation from a set of (standardized) calibration data, and

b) to describe how this equation is subsequently used to make volume determinations.

Details concerning the construction of the model, its representation in matrix terms, procedures for estimating model parameters and the computation of associated uncertainties are presented. A method for comparing two or more estimates of the measurement equation is also presented.

The statistical measurement model presented in 7.2 may be used to estimate either the tank's calibration equation or its measurement equation. For estimating the calibration equation, the response variable, always denoted by Y, represents height or elevation and the control variable, denoted by X, represents volume. As with the plots in Clause 5, the interpretation of Y and X is reversed for estimating the measurement equation (the inverse of the calibration equation). It is theoretically correct to estimate the calibration equation, $H = f(V)$, and invert this estimate to make subsequent volume determinations. However, it is simpler to estimate the measurement equation, $V = f^{-1}(H)$, directly and the added complexity induced by the inversion step is seldom warranted in practice. For convenience, it is, therefore, assumed that the statistical model of 7.2 and 7.3 is being used to estimate the measurement equation, with the understanding that the discussion can also be applied to the calibration equation.

It is assumed throughout this clause that the calibration data have been suitably standardized to compensate for variations in measurement conditions during the calibration exercise (see ISO 18213-2 and either ISO 18213-4 or ISO 18213-5, as appropriate). Data that have not been standardized to a fixed set of reference conditions should not be used to estimate the measurement equation because it is possible to obtain an estimate that yields seriously biased estimates of volume and volume measurement uncertainty.

An expression is given in 7.4 [(Equation (39)] for the uncertainty of volume predictions made with the estimated calibration equation. In 8.2, this expression is combined with the uncertainty estimate for a new height determination as derived in 6.1 to obtain an expression for the total uncertainty of a volume determination made by evaluating the estimated measurement equation at the given height determination. Such determinations, which are the major product of the tank calibration and volume determination effort, are of fundamental importance in any accountability programme for bulk materials in a nuclear facility.

The modeling approach presented here represents an advance over conventional practice in that it enables the simultaneous estimation of the entire calibration equation [3]. This approach not only makes the model easier to use, but it also simplifies and improves the estimation of uncertainties for model parameters and volume determinations.

7.2   Measurement model

7.2.1   Model construction

A statistical model for estimating a tank's volume measurement equation is presented in this subclause and further developed in 7.2.2. The general form of the model is as given in Equation (10):

$$Y = h(X) + \varepsilon \tag{10}$$

where, by the convention of 7.1, the control variable X denotes standardized height, the response variable Y denotes standardized volume, and h denotes the measurement function f^{-1}. For each observation (X_i, Y_i), the variable $\varepsilon_i = Y_i - h(X_i)$ represents the difference between the observed value, Y_i, of the response variable and the corresponding predicted value, $h(X_i)$. The function, h, is described by various parameters that are "fitted" to, i.e., estimated from, a particular set of calibration data. To compute a suitable estimate, $\hat{h}$, of the measurement equation h, the range of the calibration is divided into several segments and a function, typically a low-order polynomial, is fitted to the data of each segment. The individual estimates for each segment are then combined to obtain an overall estimate of h.

---

3)   The conventional approach is to partition the tank profile into several segments, fit an equation to the data of each segment (independent of other segments), and combine the resulting segment-wise fits to obtain an overall fit.

Relative to conventional practice, in which the measurement equation is estimated in a piecewise fashion, the modelling procedure presented here yields a unified estimate of the entire measurement equation for all segments in a single computational step. The essence of the approach for simultaneously estimating all segments of the measurement equation is to view estimates for individual segments as functions that extend over the entire range of the calibration in a specified way, rather than as independent functions, each of which is defined for a single segment. It is now possible to express the measurement equation as the sum of these "extended" functions. Once this is accomplished, standard statistical procedures can be used to simultaneously estimate all parameters of the composite representation of h, together with the corresponding matrix of model parameter uncertainties, in a single mathematical calculation.

As under the conventional approach, the first step in the construction of a composite representation of h is to divide the range of the calibration into suitable segments. From an examination of the tank's profile, it is usually possible to identify several segments that correspond to distinct regions in the tank in which the cross-sectional area is either roughly constant or has an identifiable functional form.[4)] Let Equation (11)

$$c_0 \left(= 0\right), c_1, ..., c_{s-1}, c_s \left(= X_{max}\right) \tag{11}$$

denote the boundaries of segments 1, 2, .., S, respectively, and let $h_1, h_2, ..., h_S$ denote the functions selected to fit the data in these segments. Each h_s is usually selected to be a polynomial of low (first or second) degree, such as given in Equation (12) or Equation (13):

$$h_s \left(X\right) = \beta_1 \left(X - c_{s-1}\right) \tag{12}$$

$$h_s \left(X\right) = \beta_1 \left(X - c_{s-1}\right) + \beta_2 \left(X - c_{s-1}\right)^2 \tag{13}$$

The values, h_s, of the function are initially defined only for a particular segment. However, if the definition of h_s is extended so that $h_s(X) = 0$ for $X \leqslant c_{s-1}$ and $h_s(X) = h_s(c_s)$ for $X > c_s$, then it is possible to express the measurement equation, h, as given in Equation (14):

$$h\left(X\right) = \beta_0 + h_1\left(X\right) + h_2\left(X\right) + ... + h_S\left(X\right) \tag{14}$$

where X ranges from $c_0 = 0$ to $c_S = X_{max}$. The extensions of the segment-wise functions h_s (except the first) are constructed so that $h_s(c_{s-1}) = 0$, i.e., so that Equation (15) holds:

$$\beta_0 + h_1\left(c_{s-1}\right) + h_2\left(c_{s-1}\right) + ... + h_{s-1}\left(c_{s-1}\right) = \beta_0 + h_1\left(c_{s-1}\right) + h_2\left(c_{s-1}\right) + ... + h_{s-1}\left(c_{s-1}\right) + h_s\left(c_{s-1}\right) \tag{15}$$

This construction ensures that h is continuous at segment boundaries and makes it possible to write a design matrix that permits the simultaneous estimation of all "segment" functions, h_s.

It is convenient for computational purposes to write the model of Equations (10) and (14) in matrix form. For a given set of (standardized) calibration data pairs (X_i, Y_i), let

$$\mathbf{Y} = \mathbf{H\beta} + \mathbf{\epsilon} \tag{16}$$

where

Y is an $n \times 1$ vector of response variable observations Y_i, $i = 1, 2, ..., n$;

H is an $n \times (p + 1)$ design matrix whose individual rows are functions of control variable observations X_i[5)];

---

4) There is a certain amount of trial-and-error involved in the process of identifying suitable segments, and several iterations can be required during the fitting process to refine segment boundaries. Profile variation and incremental slope plots are extremely helpful for identifying suitable calibration segments (see 5.2.3 and 5.2.4).

5) The matrix form of the function h that relates the observed response Y_i to the corresponding observation X_i is $Y_i = h(X_i) = \mathbf{h'_i\beta}$, where $\mathbf{h'_i}$ is the row of **H** that corresponds to the ith observation (X_i, Y_i).

$\boldsymbol{\beta}$ is a $(p + 1) \times 1$ vector of model parameters [$(p + 1)$ is the number of parameters in the function h, including the intercept, β_0];

$\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of residual differences, i.e., fitting errors.

NOTE   The components of $\boldsymbol{\varepsilon}$ are assumed to be independent and identically distributed normal (Gaussian) random variables, each with expected value (mean) of zero and a variance, σ^2.

The design matrix $\mathbf{H}$ contains one row for each observation pair (X_i, Y_i) and one column for each parameter in the composite expression for h. The first column of $\mathbf{H}$ corresponds to the intercept term, β_0. Each additional column corresponds to a term in the extended polynomial expression selected to model the function h over a particular segment [see Equation (14)]. A linear function requires one column, a quadratic function requires two columns, etc.

Individual entries in the columns of $\mathbf{H}$ that correspond to linear functions are constructed as follows. If the segment has boundaries c_{s-1} and c_s [6]), then the corresponding column of $\mathbf{H}$ contains a zero for each observation X for which $X \leqslant c_{s-1}$. The column contains the value $(X - c_{s-1})$ for each observation in the given segment (those for which $c_{s-1} < X \leqslant c_s$). Finally, the column contains the value $(c_s - c_{s-1})$ for observations for which $X > c_s$. Entries in the column for a quadratic term are obtained by squaring the corresponding entries in the column for the associated linear term. Similarly, entries in the column for a third-degree term or higher are obtained by raising the corresponding entries in the column for the associated linear term to the appropriate power.

Thus, for example, the design matrix for a three-segment function, h, in which h_1 is quadratic, and h_2 and h_3 are linear, has five columns: one for the intercept, one each for the linear and quadratic terms in h_1, and one each for the linear terms in h_2 and h_3. The design matrix for this illustrative example has the following block structural form as given in Equation (17):

$$\mathbf{H} = \begin{bmatrix} 1 & X - c_0 & (X - c_0)^2 & 0 & 0 \\ 1 & c_1 - c_0 & (c_1 - c_0)^2 & X - c_1 & 0 \\ 1 & c_1 - c_0 & (c_1 - c_0)^2 & c_2 - c_1 & X - c_2 \end{bmatrix} \quad \begin{matrix} \text{Segment 1} & c_0 < X \leqslant c_1 \\ \text{Segment 2} & c_1 < X \leqslant c_2 \\ \text{Segment 3} & c_2 < X \leqslant c_3 \end{matrix} \qquad (17)$$

The basic idea of this example is readily extended to measurement equations with more than three segments and to polynomials of higher degree. However, it is rarely necessary in practice to use polynomials of the third degree or higher. Because statistical complexity increases with degree, the construction of higher degree polynomials is not recommended except in unusual situations [7]).

After a model has been specified for the measurement equation (i.e., the segment boundaries, c_s, are defined and the degree of the polynomial has been specified for each segment), the design matrix $\mathbf{H}$ for that model is easily generated by computer for a particular set of calibration data. Once the design matrix for a particular model has been generated, it is no longer necessary to keep track of which parameters are associated with a particular segment because this information is coded into the design matrix.

---

6)   The lower boundary of the first segment is denoted by $c_0 = 0$. The upper boundary of the last segment is denoted by $c_S = X_{\max}$.

7)   There is a trade-off between the number of segments and the maximum degree of the polynomial required to fit a given set of calibration data. The goal is to obtain a fit that yields a "straight-line" residual plot for each run with low degree polynomials. It is generally preferable to increase the number of segments rather than to fit polynomials of a degree greater than 2 or 3.

17

7.2.2 Run-to-run variation

In practice, data from several calibration runs are used to estimate a tank's measurement equation, and it is possible (likely) that these data exhibit significant variation from one run to another. To obtain realistic estimates of variability for predictions of the response variable Y, it is necessary that this run-to-run variation be taken into account in the statistical modelling process.

The effects of run-to-run variation are incorporated into the general model of Equation (16) by allowing the vector of model parameters, $\boldsymbol{\beta}$, to vary from one run to another. Specifically, it is assumed that parameter vector $\boldsymbol{\beta}$ is subjected to a small random perturbation, $\boldsymbol{\theta}$, for each run. Thus, a particular run, j, is actually governed by the parameter vector $\boldsymbol{\beta}_j = \boldsymbol{\beta} + \boldsymbol{\theta}_j$, where the components, $\theta_{j,k}$, of $\boldsymbol{\theta}_j$ represent perturbations in the corresponding components, $\beta_{j,k}$, of $\boldsymbol{\beta}$. It is further assumed that the perturbations, $\boldsymbol{\theta}$, are independent, normal (Gaussian), random vectors, each with expected value $\mathbf{E}(\boldsymbol{\theta}) = 0$ and a variance-covariance matrix as given in Equation (18):

$$\mathbf{var}(\boldsymbol{\theta}) = \mathbf{E}(\boldsymbol{\theta}\boldsymbol{\theta}') = \boldsymbol{\Phi}^2 \qquad (18)$$

The components of $\boldsymbol{\Phi}^2$ are the variances and covariances of the components of $\boldsymbol{\theta}$. Thus, run-to-run variability in the data is expressed in terms of the statistical properties of (the components of) the vector of perturbations, $\boldsymbol{\theta}$. It follows that $\boldsymbol{\beta}_j$ is a Gaussian random vector with expected value $\mathbf{E}(\boldsymbol{\beta}_j) = \boldsymbol{\beta}$ and variance-covariance matrix $\mathbf{var}(\boldsymbol{\beta}_j) = \mathbf{var}(\boldsymbol{\theta})$.

The model of Equation (16) is typically fit to the standardized calibration data $(X_{j,i},\ Y_{j,i})$ from several calibration runs. For data from the jth calibration run, this model has the matrix form as given in Equation (19):

$$\mathbf{Y}_j = \mathbf{H}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{H}_j(\boldsymbol{\beta} + \boldsymbol{\theta}_j) + \boldsymbol{\varepsilon}_j \qquad (19)$$

where

$\mathbf{Y}_j$	is an $n_j \times 1$ vector of response variable observations, $\mathbf{Y}_{j,i}$, from run j;
$\mathbf{H}_j$	is an $n_j \times (p+1)$ design matrix whose individual rows are functions of observations of the control variable, $X_{j,i}$,[8];
$\boldsymbol{\beta}_j = (\boldsymbol{\beta} + \boldsymbol{\theta}_j)$	is a $(p+1) \times 1$ vector of model parameters for the jth run;
$\boldsymbol{\varepsilon}_j$	is an $n_j \times 1$ vector of residual differences (fitting errors).
j	is the run number ($j = 1, 2, ..., r$).

Thus, the quantities $\mathbf{Y}_j$, $\mathbf{H}_j$, and $\boldsymbol{\varepsilon}_j$ are defined exactly as for Equation (16), except that they now pertain only to the calibration data $(X_{j,i},\ Y_{j,i})$ from the jth calibration run.

Here, the vector of model parameters $\boldsymbol{\beta}_j$ is interpreted as the realization of $\boldsymbol{\beta}$ that governs the data of the jth calibration run. The $\boldsymbol{\beta}_j$s can now be estimated using conventional statistical methods. As noted above, the statistical properties of the $\boldsymbol{\theta}_j$s are used to estimate the run-to-run component of variability in the calibration data.

---

8) The matrix form of the function $\mathbf{h}_{j,i}$ that relates the observed response $Y_{j,i}$ to the corresponding observation $X_{j,i}$ is written as $Y_{j,i} = h_{j,i}(X_{j,i}) = \mathbf{h}'_{j,i}\boldsymbol{\beta}_j = \mathbf{h}'_{j,i}(\boldsymbol{\beta} + \boldsymbol{\theta}_j)$, where $\mathbf{h}'_{j,i}$ is the row of $\mathbf{H}_j$ that corresponds to the ith observation from the jth calibration run.

In accordance with 7.2.1, the components of $\boldsymbol{\varepsilon}_j$ are assumed to be independent and identically distributed, normal (Gaussian), random variables, each with an expected value (mean) of zero and a variance, σ_j^2. It is further assumed that the distribution of measurement errors is the same for all runs, so that σ_j^2 can be interpreted as the realization of σ^2 [see Equation (16)] obtained from the data of the jth calibration run. Finally, it is assumed that errors from different runs are independent.

Under the model of Equation (19), the vector of responses $\mathbf{Y}_j$ has the expected value as given in Equation (20):

$$\mathbf{E}\left(\mathbf{Y}_j\right) = \mathbf{H}_j\boldsymbol{\beta} \tag{20}$$

and the variance as given in Equation (21):

$$\mathbf{var}\left(\mathbf{Y}_j\right) = \mathbf{H}_j\boldsymbol{\Phi}^2\mathbf{H}_j' + \sigma_j^2\mathbf{I} \tag{21}$$

where

$\boldsymbol{\Phi}^2 = \mathbf{E}(\boldsymbol{\theta}\boldsymbol{\theta}') = \mathbf{var}(\boldsymbol{\theta})$ is the variance-covariance matrix of the vector of perturbations, $\boldsymbol{\theta}$, [see Equation (18)];

$\mathbf{I}$ is the $n_j \times n_j$ identity matrix;

σ_j^2 is the variance of the components of the error vector, $\boldsymbol{\varepsilon}_j$.

The statistical properties of individual components of the vector, $\mathbf{Y}_j$, are derived by means of Equations (20) and (21). In particular, the ith component, $Y_{j,i}$, of $\mathbf{Y}_j$ has the expected value as given in Equation (22):

$$\mathbf{E}\left(Y_{j,i}\right) = \boldsymbol{\beta}\mathbf{h}'_{j,i} \tag{22}$$

and the variance as given in Equation (23):

$$\mathbf{var}\left(Y_{j,i}\right) = \mathbf{h}'_{j,i}\boldsymbol{\Phi}^2\mathbf{h}_{j,i} + \sigma_j^2 \tag{23}$$

where $\mathbf{h}'_{j,i}$ is the ith row of $\mathbf{H}_j$ (the row that corresponds to $X_{j,i}$). All other quantities in Equations (22) and (23) are defined as in Equations (20) and (21).

It is important to note that when the data from several calibration runs are modeled, the same model is fit to the data of each calibration run. This means, in particular, that the same segments (segment boundaries) are specified for each run and that the polynomials specified for the corresponding segments of all runs have the same degree.

7.3 Estimation of model parameters

7.3.1 Preliminaries

Parameters in the model for a particular run are estimated by "fitting" the model [Equation (19)] to the standardized calibration data from that run. Fitting is the process of determining those parameter values (called "least-squares" estimates) that minimize the sum of squared differences between the observed and predicted responses, i.e., that minimize the residual sum of squares, where the summation extends over all observed values of the control variable. Standard statistical methods are employed to compute the least-squares estimates of model parameters.

Parameters in the general model [Equation (16)] are estimated from the parameters fitted to the data of the individual runs. In particular, the parameter vector β is estimated by averaging the estimates of the parameter vectors, β_j, for the individual runs.

7.3.2 Individual runs

Let $(X_{j,i}, Y_{j,i})$, $i = 1,2, \ldots, n_j$, denote the standardized data from the jth calibration run. In matrix form, the least-squares estimate of parameter vector $\beta_j = (\beta + \theta_j)$ that fits the measurement model of Equation (19) to these data is given in Equation (24):

$$\hat{\beta}_j = \left(H'_j H_j\right)^{-1} H'_j Y_j \tag{24}$$

It follows from the assumed model that $\hat{\beta}_j$ has the expected value $E\left(\hat{\beta}_j\right) = \beta$ and variance-covariance matrix as given in Equation (25):

$$var\left(\hat{\beta}_j\right) = \sigma_j^2 \left(H'_j H_j\right)^{-1} + \Phi^2 \tag{25}$$

where

H_j is defined as for Equation (19);

Φ^2 is defined by Equation (18);

NOTE   Estimation of Φ^2 is deferred to 7.3.3 because it is not possible to estimate this quantity from the data of a single run.

σ_j^2 is the variance of the residual differences (errors) $\varepsilon_{j,i}$.

This residual variance, σ_j^2, is estimated as given in Equation (26):

$$\begin{aligned} \hat{\sigma}_j^2 &= \sum_i \left[Y_{j,i} - \hat{h}_j\left(X_{j,i}\right) \right]^2 \Big/ \left[n_j - (p+1) \right] \\ &= \sum_i \hat{\varepsilon}_{j,i}^2 \Big/ \left[n_j - (p+1) \right] \end{aligned} \tag{26}$$

In matrix form, this estimator, $\hat{\sigma}_j^2$, is written as given in Equation (27):

$$\begin{aligned} \hat{\sigma}_j^2 &= \left(Y_j - H_j \hat{\beta}_j \right)' \left(Y_j - H_j \hat{\beta}_j \right) \Big/ \left[n_j - (p+1) \right] \\ &= \hat{\varepsilon}'_j \hat{\varepsilon}_j \Big/ \left[n_j - (p+1) \right] \end{aligned} \tag{27}$$

7.3.3 Several runs

7.3.3.1 Vector β

The parameter vector, β, in the general model of 7.2.1 [see Equation (16)] is estimated by averaging estimates of the parameter vectors $\beta_j = (\beta + \theta_j)$ obtained by individually fitting the data of several calibration runs as given in Equation (28):

$$\hat{\beta} = r^{-1} \sum_{j=1}^{r} \hat{\beta}_j \tag{28}$$

It follows from 7.3.2 and the assumed independence of the $\hat{\boldsymbol{\beta}}_j$s that the estimator, $\hat{\boldsymbol{\beta}}$, is a random vector with an expected value $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ and a variance [see Equation (25)] as given in Equation (29):

$$
\begin{aligned}
\mathbf{var}(\hat{\boldsymbol{\beta}}) &= r^{-2}\sum\nolimits_{j=1}^{r}\mathbf{var}(\hat{\boldsymbol{\beta}}_j) \\
&= r^{-2}\sum\nolimits_{j=1}^{r}\left[\sigma_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + \boldsymbol{\Phi}^2\right] \\
&= r^{-2}\sum\nolimits_{j=1}^{r}\sigma_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + r^{-1}\boldsymbol{\Phi}^2
\end{aligned}
\tag{29}
$$

7.3.3.2   Predicted mean response, $\overline{Y}_0$

The predicted (mean) value of the response variable, Y_0, at a specified new (unobserved) value of the control variable X_0 is estimated as given in Equation (30):

$$
\hat{\overline{Y}}_0 = \mathbf{h}'_0\hat{\boldsymbol{\beta}}
\tag{30}
$$

where $\mathbf{h}'_0$ denotes the row vector of $\mathbf{H}$ that corresponds to X_0.

The quantity $\hat{\overline{Y}}_0$ has the expected value as given in Equation (31):

$$
\mathbf{E}\left(\hat{\overline{Y}}_0\right) = \mathbf{h}'_0 E(\hat{\boldsymbol{\beta}}) = \mathbf{h}'_0\boldsymbol{\beta}
\tag{31}
$$

and a variance [see Equation (29)] as given in Equation (32):

$$
\begin{aligned}
\mathbf{var}\left(\hat{\overline{Y}}_0\right) &= \mathbf{var}\left(\mathbf{h}'_0\hat{\boldsymbol{\beta}}\right) \\
&= r^{-2}\mathbf{h}'_0\sum\nolimits_{j=1}^{r}\sigma_j^2\left[\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + r\boldsymbol{\Phi}^2\right]\mathbf{h}_0
\end{aligned}
\tag{32}
$$

7.3.3.3   Variances of $\hat{\boldsymbol{\beta}}$ and $\hat{\overline{Y}}_0$

An estimate of the covariance matrix, $\boldsymbol{\Phi}^2$, is required to compute the variances of $\hat{\boldsymbol{\beta}}$ and $\hat{\overline{Y}}_0$. This matrix is estimated as given in Equation (33):

$$
\begin{aligned}
\hat{\boldsymbol{\Phi}}^2 &= r^{-1}\sum\nolimits_{j=1}^{r}\left(\hat{\boldsymbol{\beta}}_j - \hat{\boldsymbol{\beta}}\right)\left(\hat{\boldsymbol{\beta}}_j - \hat{\boldsymbol{\beta}}\right)' \\
&= r^{-1}\sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j
\end{aligned}
\tag{33}
$$

The elements of the $(p+1)\times(p+1)$ matrix $\hat{\boldsymbol{\Phi}}^2 = r^{-1}\Sigma_j\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j$ are the estimated variances and covariances of the run-to-run variance parameters, $\theta_{j,k}$, in the model of 7.2.2.

Estimates of $\mathbf{var}(\hat{\boldsymbol{\beta}})$ and $\mathbf{var}\left(\hat{\overline{Y}}_0\right)$ are now obtained by substituting the estimators for $\boldsymbol{\Phi}^2$ and σ_j^2 from Equations (33) and (27) into Equations (29) and (32) as given in Equations (34) and (35), respectively:

$$
\begin{aligned}
\hat{\mathbf{var}}(\hat{\boldsymbol{\beta}}) &= r^{-2}\left[\sum\nolimits_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}_j\mathbf{H}'_j\right)^{-1} + r\hat{\boldsymbol{\Phi}}^2\right] \\
&= r^{-2}\left[\sum\nolimits_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}_j\mathbf{H}'_j\right)^{-1} + \sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right]
\end{aligned}
\tag{34}
$$

$$\hat{\text{var}}\left(\hat{\bar{Y}}_0\right) = \hat{\text{var}}\left(\mathbf{h}'_0\hat{\boldsymbol{\beta}}\right)$$

$$= r^{-2}\mathbf{h}'_0\left[\sum_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}_j\mathbf{H}'_j\right)^{-1} + \sum_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right]\mathbf{h}_0 \tag{35}$$

7.3.3.4   Variance, σ^2

The error variance estimates, σ_j^2, given by Equation (27) have the same statistical properties for all runs, so they are pooled to obtain the estimate for σ^2 as given in Equation (36):

$$\hat{\sigma}^2 = \sum_{j=1}^{r}\left[n_j - (p+1)\right]\hat{\sigma}_j^2 \Big/ \sum_{j=1}^{r}\left[n_j - (p+1)\right]$$

$$= \sum_{j=1}^{r}\sum_{i=1}^{n_j}\hat{\varepsilon}_{j,i}^2 \Big/ \left[n - r(p+1)\right] \tag{36}$$

$$= \sum_{j=1}^{r}\hat{\boldsymbol{\varepsilon}}'_j\hat{\boldsymbol{\varepsilon}}_j \Big/ \left[n - r(p+1)\right]$$

where $n = \sum_{j}n_j$ is the total number of observations from all runs [9].

7.4   Volume determinations and variance estimates

In the model of 7.2.2, the response Y_0 for a single new (future) value of the control variable X_0 is expressed as given in Equation (37):

$$Y_0 = h(X_0) + \varepsilon_0$$

$$= \mathbf{h}'_0(\boldsymbol{\beta} + \boldsymbol{\theta}) + \varepsilon_0 \tag{37}$$

where

$\mathbf{h}'_0$ is the row vector of the design matrix $\mathbf{H}$ that corresponds to X_0;

NOTE   The specific form of $\mathbf{h}'_0$ depends on the segment of the measurement equation into which X_0 falls.

ε_0 is the residual prediction error.

Let $\hat{Y}_0$ denote the estimated value of the new response, Y_0. Then, the estimated (mean) value of $\hat{Y}_0$ is given by Equation (30). The variance of $\hat{Y}_0$ has two components:

a)   variance of its estimated mean;

b)   variance of a (new) individual observation.

---

9)   The variance estimators, $\hat{\sigma}_j^2$, for the individual runs all have the same statistical properties, so it is theoretically correct to replace each $\hat{\sigma}_j^2$ in Equation (35) by $\hat{\sigma}^2$. In practice, differences between the two computational alternatives are small.

It follows from Equations (32) and (23) that the theoretical variance of the new predicted response at X_0 is as given in Equation (38):

$$\mathbf{var}\left(\hat{Y}_0\right) = \mathbf{var}\left(\hat{\bar{Y}}_0\right) + \mathbf{var}\left(Y_0\right)$$
$$= \mathbf{var}\left(\mathbf{h}'_0\hat{\boldsymbol{\beta}}\right) + \mathbf{var}\left(Y_0\right) \tag{38}$$
$$= r^{-2}\mathbf{h}'_0\left[\sum_{j=1}^{r}\sigma_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + r\boldsymbol{\Phi}^2\right]\mathbf{h}_0 + \mathbf{h}'_0\boldsymbol{\Phi}^2\mathbf{h}_0 + \sigma^2$$

An estimator of $\mathbf{var}\left(\hat{Y}_0\right)$ is obtained by substituting the estimators of Equations (27), (36) and (33) for the corresponding theoretical quantities into Equation (38) as given in Equation (39):

$$\mathbf{var}\left(\hat{Y}_0\right) = r^{-2}\mathbf{h}'_0\left[\sum_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + \sum_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right]\mathbf{h}_0 + r^{-1}\mathbf{h}'_0\left[\sum_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right]\mathbf{h}_0 + \hat{\sigma}^2$$
$$= r^{-2}\mathbf{h}'_0\left[\sum_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + (r+1)\left(\sum_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\right]\mathbf{h}_0 + \hat{\sigma}^2 \tag{39}$$

The decomposition in Equation (39) shows explicitly how run-to-run variability contributes to the variance of a new, observed value of the response variable. This decomposition is relevant not only because run-to-run variability in the calibration data contributes to the uncertainty of the estimated measurement equation, but also because it contributes to the uncertainty of any future volume determinations. Equations (35) and (39) are the basic expressions from which all subsequent statistical results in this part of ISO 18213 are derived. In particular, these expressions are used to obtain both the confidence regions and the prediction intervals presented in 7.5, as well as the uncertainty estimates for volume determinations given in 8.2.

At a given value X_0 of the control variable, the quantity $\mathbf{h}'_0\hat{\boldsymbol{\theta}}_j = \mathbf{h}'_0\left(\hat{\boldsymbol{\beta}}_j - \hat{\boldsymbol{\beta}}\right)$ that appears implicitly in the second term in brackets in the first line of Equation (39) is the difference between the predicted value of the response variable at X_0 for the jth run and that of the overall (average) fit. Thus, the estimate of the run-to-run component of variability in Y_0 at a given value of X_0 is simply the estimated variance of the predicted values $\hat{\bar{Y}}_0$ obtained from the fits of the calibration model to the data of the individual calibration runs.

If the statistical model of 7.2 is used estimate to the measurement equation (where Y = volume, X = height), then Equation (30) yields the volume estimate for a particular (standardized) determination of height, and Equation (39) yields the estimated variance of that volume estimate. Conversely, if the statistical model is used to estimate the calibration equation (where Y = volume, X = height), then Equation (30) yields the height that corresponds to a particular volume, and Equation (39) yields the estimated variance of that height determination. Since the ultimate goal is to determine the volume associated with a particular height determination, it is necessary to "invert" the results in the latter case. The latter approach is theoretically correct, but the additional complexity introduced by choosing this alternative is seldom warranted in practice.

7.5 Confidence regions and prediction intervals

7.5.1 General

It is possible to construct confidence regions for estimates of the measurement equation and prediction intervals for values of the response variable. These confidence regions and prediction intervals are useful for determining the significance of differences between

a) individual observations,

b) observations and an estimate of the measurement equation,

c) two estimates of the measurement equation.

They are especially helpful when plotted in connection with the diagnostic plots presented in Clause 5. All of the prediction intervals and confidence regions presented here are based on Equation (30) and Equation (39), and are derived by the methods in Reference [4]. In particular, the variance of the estimated (mean) response $\hat{\bar{Y}}_0 = \mathbf{h}'_0\hat{\boldsymbol{\beta}}$ is given by Equation (35) for any value of the control variable, X_0.

To ensure that the computed confidence regions and prediction intervals yield valid comparisons, it is assumed throughout that all data have been standardized to the same set of reference conditions.

7.5.2   Confidence regions for the measurement equation

7.5.2.1   Confidence interval for the measurement equation at a single value of the control variable

An approximate $100(1 - \alpha)$ % confidence interval for the predicted (mean) value of the measurement equation at a single specified value $X = X_0$ of the control variable, X, has the form as given in Equation (40):

$$\mathbf{h}'_0\hat{\boldsymbol{\beta}} \pm \hat{\sigma}_0 t_{\alpha/2}(\nu) \tag{40}$$

where

$\mathbf{h}'_0$ is the row vector of the design matrix $\mathbf{H}$ that corresponds to X_0;

$\hat{\boldsymbol{\beta}}$ is given by Equation (28);

$\hat{\sigma}_0^2$ is the estimated variance of the prediction $\hat{\bar{Y}}_0 = \mathbf{h}'_0\hat{\boldsymbol{\beta}}$ given by Equation (35);

$t_{\alpha/2}(\nu)$ is the $100(1 - \alpha/2)$ % point from the t-distribution with parameter (degrees of freedom) ν;

ν is the (approximate) degrees of freedom for the variance estimate $\hat{\sigma}_0^2$.

It is necessary to compute the degrees of freedom, ν, before the confidence interval of Equation (40) can be computed. The number of degrees of freedom depends upon

a)   the relative magnitudes of the components of $\hat{\sigma}_0^2$,

b)   the number of parameters in the model $(p + 1)$,

c)   the number runs, r,

d)   the number of observations in each run that are used to estimate the model parameters, n_j.

The computation is typically accomplished by means of the Welch-Satterthwaite equation (see Annex B). In this case, the approximate degrees of freedom for (the distribution of) $\hat{\sigma}_0^2$ is computed from the two terms on the right-hand side of Equation (35) as given in Equations (41) and (42):

$$S_1^2 = \mathbf{h}'_0\left[\sum_{j=1}^{r} \hat{\sigma}_j^2 \left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} \right]\mathbf{h}_0 \tag{41}$$

$$S_2^2 = \mathbf{h}'_0\left[\sum_{j=1}^{r} \left(\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)^{-1} \right]\mathbf{h}_0 \tag{42}$$

The approximate degrees of freedom for S_1^2 and S_2^2 are given by Equations (43) and (44) [10], respectively:

$$\nu_1 = n - r(p + 2) \tag{43}$$

$$\nu_2 = r \tag{44}$$

The value of ν is now computed by means of Equation (B.2), with $V = S_1^2/\left[n - r\left(p + 2\right)\right]$ and $W = S_2^2/r$. Typical values of α are 0,025 and 0,05.

7.5.2.2 Confidence intervals for the measurement equation that hold simultaneously for an indeterminate number of values of the control variable

Approximate $100(1 - \alpha)$ % point confidence intervals for the predicted (mean) value of the measurement equation that hold simultaneously for an indeterminate number of values of the control variable, X, have the form as given in Equation (45):

$$\mathbf{h}'_X \hat{\boldsymbol{\beta}} \pm \hat{\sigma}_X \left[\left(p + 1 \right) F_\alpha \left(p + 1, \nu \right) \right]^{0,5} \tag{45}$$

where

$\mathbf{h}'_X$ is the row vector of the design matrix $\mathbf{H}$ that corresponds to X;

$\hat{\boldsymbol{\beta}}$ is given by Equation (28);

$\hat{\sigma}_X^2$ is the estimated variance of the prediction $\hat{\bar{Y}}_X = \mathbf{h}'_X \hat{\boldsymbol{\beta}}$ given by Equation (35);

$F_\alpha \left(p + 1, \nu \right)$ is the $100(1 - \alpha)$ % point from the F-distribution with parameters $(p + 1)$ and ν;

ν is the (approximate) degrees of freedom for the variance estimate $\hat{\sigma}_X^2$.

The quantity $(p + 1)$ is the number of parameters in the underlying general measurement model (see 7.2). For a particular value of X, the parameter ν is computed by means of the Welch-Satterthwaite formula (see Annex B) exactly as for Equation (40). The values of $\hat{\sigma}_X^2$ and the corresponding values of ν obtained by this computation vary from one value of X to another. Instead of computing these quantities for each X, the smallest value of $\hat{\sigma}_X^2$ (produced by the smallest value of S_1^2) and the corresponding value of ν can be used for all values of X. This alternative approximation produces confidence intervals that are somewhat narrower than those obtained when $\hat{\sigma}_X^2$ and ν are determined for each X.

Typical values of α are 0,025 and 0,05.

An approximate confidence region for arbitrarily many values of the control variable is obtained by first evaluating Equation (45) for a series of values of X selected to cover some range of interest, and then interpolating between adjacent points. The resulting confidence region is especially useful when used in connection with the plots presented in 5.2. The plot is most easily interpreted when $\hat{\bar{Y}}_X = \mathbf{h}'_X \hat{\boldsymbol{\beta}}$ is used as a reference, and upper and lower confidence limits are plotted as differences relative to this equation (see Annex A for examples).

---

10) The total number of observations from all runs is $n = \sum_j n_j$.

7.5.2.3 Confidence bands for the difference between two estimates of the measurement equation

Suitable confidence bands are particularly useful for comparing two estimates of the measurement equation. The two estimates may, for example, be those from two calibration exercises conducted at different times. In this case, one estimate is typically from a new calibration undertaken to verify an existing (old, reference) measurement equation obtained from a previous calibration exercise. The two estimates may also be those from two analyses of a single set of calibration data, such as those produced by the operator and the inspector.

An approximate $100(1-\alpha)$ % confidence interval for the difference between two independent estimates of the measurement equation at some arbitrary value of the control variable, X, is given by Equation (46):

$$\mathbf{h}'_X\left(\hat{\boldsymbol{\beta}}_{\text{new}}-\hat{\boldsymbol{\beta}}\right)\pm\left[\left(\hat{\sigma}^2_{X,\text{new}}+\hat{\sigma}^2_X\right)(p+1)F_\alpha\left(p+1,\nu\right)\right]^{0,5} \tag{46}$$

where

$\mathbf{h}'_X$ is the row vector of the design matrix $\mathbf{H}$ that corresponds to X;

$\hat{\bar{Y}}_{X,\text{new}}=\mathbf{h}'_X\hat{\boldsymbol{\beta}}_{\text{new}}$ is the predicted value of the response variable at X obtained from the new measurement equation;

$\hat{\bar{Y}}_X=\mathbf{h}'_X\hat{\boldsymbol{\beta}}$ is the predicted value of the response variable at X obtained from the reference (old) measurement equation.

The estimated variances of $\mathbf{h}'_X\hat{\boldsymbol{\beta}}_{\text{new}}$ and $\mathbf{h}'_X\hat{\boldsymbol{\beta}}$, denoted respectively by $\hat{\sigma}^2_{X,\text{new}}$ and $\hat{\sigma}^2_X$, are computed from Equation (35), each with values of r, $\hat{\sigma}^2$, $\mathbf{H}_j$, and $\hat{\boldsymbol{\theta}}_j$ appropriate to its respective calibration run (new or old).

The quantity $F_\alpha\left(p+1,\nu\right)$ is the $100(1-\alpha)$ % point from the F-distribution with parameters (degrees of freedom) $(p+1)$ and ν. The degrees of freedom parameter, ν, is computed from the individual degrees of freedom for $\hat{\sigma}^2_{X,\text{new}}$ and $\hat{\sigma}^2_X$ by means of the Welch-Satterthwaite formula. If ν_1 and ν_2 denote the degrees of freedom for the quantities $\hat{\sigma}^2_{X,\text{new}}$ and $\hat{\sigma}^2_X$, respectively, then ν_1 and ν_2 are computed for their respective calibrations from Equation (35) by means of the Welch-Satterthwaite equation exactly as for Equation (40). The parameter ν is now computed by a second application of Equation (B.2), with $V=\hat{\sigma}^2_{X,\text{new}}\big/\nu_1$ and $W=\hat{\sigma}^2_X\big/\nu_2$.

In the case where one estimate of the measurement equation (e.g., the new one) is derived from the data of a single calibration run, it is not possible to obtain a direct estimate of the run-to-run component of variability for $\hat{\sigma}^2_{X,\text{new}}$. If the same procedure is used to derive both estimates of the measurement equation, then it may be possible to estimate the run-to-run variability for the new equation from the estimate computed for the reference equation.

Two estimates of the measurement equation, e.g., new and reference, can be compared visually by plotting the confidence bounds of Equation (46) for a series of values of X and interpolating between adjacent points. If at any point the confidence band for the difference between the new and reference equations does not contain zero, then the two equations are significantly different at the specified significance level.

7.5.3 Prediction intervals for future observations

7.5.3.1 Prediction interval for the response variable at a single future observation of the control variable

Under the model of 7.2.2, the response variable, Y_0, at a single new (future) observation of the control variable, X_0, is given by Equation (37). The predicted (mean) value of Y_0, as given by Equation (31), is $\hat{\bar{Y}}_0=\mathbf{h}'_0\hat{\boldsymbol{\beta}}$ and, from Equation (39), the estimated variance of $\hat{\bar{Y}}_0$ is given in Equation (47):

$$\hat{\sigma}_0^2 = \hat{\sigma}^2 + r^{-2}\mathbf{h}'_0\left[\sum\nolimits_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + (r+1)\left(\sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\right]\mathbf{h}_0 \tag{47}$$

Thus, an approximate $100(1-\alpha)$ % prediction interval for the response variable at a single future observation X_0 is as given in Equation (48):

$$\mathbf{h}'_0\hat{\boldsymbol{\beta}} \pm \hat{\sigma}_0 t_{\alpha/2}(\nu_0) \tag{48}$$

where $\hat{\sigma}_0^2$ is given by Equation (47). Other quantities are defined as in 7.5.2.1. For computing the degrees of freedom ν_0, it is convenient to write Equation (47) as given in Equation (49):

$$\hat{\sigma}_0^2 = r^{-2}\mathbf{h}'_0\left[\sum\nolimits_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + \left(\sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\right]\mathbf{h}_0 + \hat{\sigma}^2 + r^{-1}\mathbf{h}'_0\left(\sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\mathbf{h}_0 \tag{49}$$

The first term on the right-hand side of Equation (49) is given by Equation (35). Thus, the degrees of freedom associated with this term, say ν_1, is computed by means of the Welch-Satterthwaite formula exactly as for Equation (40). The degrees of freedom for the last two terms, say ν_2, is computed by means of the Welch-Satterthwaite formula with $V = \hat{\sigma}^2\big/\left[n - r(p+2)\right]$ and $W = r^{-2}\mathbf{h}'_0\left(\sum_j\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\mathbf{h}_0$. Finally, ν_0 is computed by means of a third application of the Welch-Satterthwaite formula, this time with V and W as given in Equations (50) and (51), respectively:

$$V = r^{-2}\mathbf{h}'_0\left[\sum\nolimits_{j=1}^{r}\hat{\sigma}_j^2\left(\mathbf{H}'_j\mathbf{H}_j\right)^{-1} + \sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right]\mathbf{h}_0\Big/\nu_1 \tag{50}$$

$$W = \left[\hat{\sigma}^2 + r^{-1}\mathbf{h}'_0\left(\sum\nolimits_{j=1}^{r}\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\mathbf{h}_0\right]\Big/\nu_2 \tag{51}$$

7.5.3.2 Prediction intervals that hold simultaneously for an indeterminate number of future observations of the control variable

It is possible to construct prediction intervals that hold simultaneously for a specified number of future observations of the control variable. These intervals are not especially useful in practice because the specific number of new predictions is rarely known in advance. Moreover, the confidence intervals become quite wide when the number of predictions is large. Nevertheless, it is useful to have a means of comparing data from a new calibration to an existing measurement equation. This is done with the aid of a family of tolerance intervals that hold simultaneously at all values of the control variable. Details of construction are given in Reference [4]. A somewhat simpler approximation that works well in practice is given here.

The variance of a single new prediction of volume, $\hat{Y}_0$, obtained from the estimated measurement equation at X_0 is given by Equation (47). Thus, approximate $100(1-\alpha)$ % confidence intervals for the predicted responses that hold jointly (simultaneously) for an indeterminate number of new values of the control variable X have the form as given in Equation (52):

$$\mathbf{h}'_X\hat{\boldsymbol{\beta}} \pm \hat{\sigma}_X\left[(p+1)F_\alpha(p+1,\nu)\right]^{0,5} \tag{52}$$

For a particular value of X, $\hat{\sigma}_X^2$ and its associated degrees of freedom, ν, are computed exactly as for Equation (48). The quantities $\hat{\sigma}_X^2$ and ν vary from one value of X to another. An alternative to computing $\hat{\sigma}_X^2$ and ν for each X is to use the smallest value of $\hat{\sigma}_X^2$ and the corresponding value of ν for all values of X. This alternative approximation produces confidence intervals that are somewhat narrower than those obtained when ν is determined for each X.

Confidence bands for an arbitrary number of new predictions are obtained by plotting the confidence bounds of Equation (52) for a series of selected values of X and interpolating between adjacent points. These confidence bands can be used to compare data from a new calibration run with an existing calibration. If any observation from the new run falls outside the confidence region for the calibration equation, the new data are inconsistent with the old calibration at the specified probability level. The plot is most easily interpreted when $\hat{\bar{Y}}_X = \mathbf{h}'_X \hat{\boldsymbol{\beta}}$ is used as a reference and upper and lower confidence limits are plotted as differences relative to this equation (see Annex A for examples).

8 Uncertainty estimates for volume determinations

8.1 Overview

The purpose of a tank calibration exercise is to develop an estimate of the tank's measurement equation that can subsequently be used to determine the volume of process liquid in the tank associated with a given measure of pressure. The quantity of interest may be either the volume of liquid contained in the tank (a contained volume) or the volume of liquid involved in a transfer operation (a transfer volume). Equations are presented in 8.2 and 8.3, respectively, for estimating the uncertainty of both contained and transfer volume determinations.

The procedure for estimating the volume of process liquid contained in a tank is broken into two major steps. First, the response that the liquid produces in the tank's measurement (manometer) system is observed and converted to a measure of height at standard reference conditions (as outlined in 6.1; see ISO 18213-4 or ISO 18213-5 for additional details). Second, this measure of height is then substituted into the tank's estimated measurement equation to obtain the corresponding determination of volume (at reference conditions). It is possible to express the total variability for the resulting volume determination as the sum of three components:

a) variance of the predicted (mean) volume attributable to the estimated measurement equation as a result of the calibration process (i.e., variability in the measurement equation due to the calibration process);

b) variance of a new volume determination not associated with the calibration process;

c) variability in the predicted volume that is propagated or transferred through the measurement equation as a result of measurement uncertainty in the underlying height determination.

8.2 Contained volumes

8.2.1 Estimation of volume variance

Let $X_0 = H_0$ denote a standardized reference height for which the associated volume is desired, let $Y_0 = V_0$ denote the corresponding standardized volume (at reference temperature T_r), and let $\hat{h} = \hat{f}^{-1}$ denote the estimated measurement equation. The reference height [see Equation (6)] can be written as given in Equation (53):

$$X_0 = H_0 = \left(\Delta P - c_{\mathsf{M}}\right)\Big/\left[g\left(\rho_{\mathsf{M}} - \rho_{\mathsf{a},s}\right)\left(1 + \alpha_{\mathsf{ex}}\Delta T_{\mathsf{m}}\right)\right] \tag{53}$$

The predicted (mean) volume obtained from the measurement equation [see Equation (30)] is as given in Equation (54):

$$V_0 = \hat{\bar{Y}}_0 = \hat{h}\left(H_0\right) = \mathbf{h}'_0\hat{\boldsymbol{\beta}} \tag{54}$$

In accordance with 8.1, the total variability of V_0 can be written as given in Equation (55):

$$\textbf{var}\left(V_0\right) = \textbf{var}\left(V_{0,\text{pred}}\right) + \textbf{var}\left(V_{0,\text{new}}\right) + \textbf{var}\left(V_{0,\text{trans}}\right)$$
$$= \textbf{var}\left(\hat{Y}_0\right) + \textbf{var}\left(V_{0,\text{trans}}\right)$$

(55)

$\textbf{var}\left(\hat{Y}_0\right)$ represents the variance of a single predicted new response obtained from the measurement equation; an estimate of this quantity, based on Equation (39), is given in Equation (56):

$$\hat{\textbf{var}}\left(\hat{Y}_0\right) = r^{-2}\textbf{h}'_0\left[\hat{\sigma}^2\sum\nolimits_{j=1}^r\left(\textbf{H}'_j\textbf{H}_j\right)^{-1} + (r+1)\left(\sum\nolimits_{j=1}^r\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\right]\textbf{h}_0 + \hat{\sigma}^2$$

(56)

The last term on the right-hand side of Equation (55) has the form as given in Equation (57):

$$\hat{\textbf{var}}\left(V_{0,\text{trans}}\right) = \left[\partial\hat{h}\left(H_0\right)\big/\partial\left(H_0\right)\right]^2\textbf{var}\left(H_0\right)$$

(57)

and is estimated, based on Equation (9), as given in Equation (58):

$$\hat{\textbf{var}}\left(V_{0,\text{trans}}\right) = \left[\partial\hat{h}\left(H_0\right)\big/\partial\left(H_0\right)\right]^2\left(H_0\right)^2\left[\textbf{var}\left(\Delta P\right)\big/\left(\Delta P - c_\text{M}\right)^2 + \textbf{var}\left(\rho_\text{M}\right)\big/\left(\rho_\text{M} - \rho_{\text{a},s}\right)^2\right]$$

(58)

The following expression for the estimated total variability of the new volume determination V_0 associated with the standardized height determination H_0 is obtained by substituting Equations (56) and (58) into Equation (55), as given in Equation (59):

$$\hat{\textbf{var}}\left(V_0\right) = r^{-2}\textbf{h}'_0\left[\hat{\sigma}^2\sum\nolimits_{j=1}^r\left(\textbf{H}'_j\textbf{H}_j\right)^{-1} + (r+1)\left(\sum\nolimits_{j=1}^r\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}'_j\right)\right]\textbf{h}_0 + \hat{\sigma}^2 + ...$$
$$... + \left[\partial\hat{h}\left(H_0\right)\big/\partial\left(H_0\right)\right]^2\left(H_0\right)^2\left[\textbf{var}\left(\Delta P\right)\big/\left(\Delta P - c_\text{M}\right)^2 + \textbf{var}\left(\rho_\text{M}\right)\big/\left(\rho_\text{M} - \rho_{\text{a},s}\right)^2\right]$$

(59)

For Equation (59), the following hold.

— H_0 is the reference (standardized) height for the liquid of interest. This quantity is determined from measurements of the pressure ΔP as given by Equation (53) (see 6.1 and corresponding discussions in ISO 18213-4 or ISO 18213-5, respectively, for details).

— $\hat{h}$ denotes the estimated measurement equation, $\hat{f}^{-1}$.

— $V_0 = \hat{h}\left(H_0\right)$ is the standardized volume that corresponds to H_0. Equivalently, $V_0 = \textbf{h}'_0\hat{\boldsymbol{\beta}}$, where $\textbf{h}'_0$ is the vector of the design matrix that corresponds to the standardized height, H_0.

— $\hat{\boldsymbol{\beta}}$ is the vector of parameters that characterize the estimated measurement equation (see 7.3).

The quantity $\partial\hat{h}\left(H_0\right)\big/\partial\left(H_0\right)$ denotes the derivative of the estimated measurement equation, $\hat{h}$, taken with respect to H and evaluated at $H = H_0$. The square of this quantity propagates or "transfers" uncertainty (variance) in H_0 through the measurement equation to the volume estimate, V_0. The value of $\partial\hat{h}\left(H_0\right)\big/\partial\left(H_0\right)$ depends on the coefficients of the measurement equation, h, for the segment in which H_0 falls. Thus, for a particular value of H_0, this quantity is computed by taking the derivative with respect to the appropriate coefficients in the vector of parameter estimates, $\hat{\boldsymbol{\beta}}$.

The remaining quantities in Equation (59) are defined in 7.2 or 7.3, and all are computed at the model fitting stage.

Beginning with a measurement of pressure and an estimate of the tank's measurement equation, the corresponding standardized volume is determined by means of the steps described in ISO 18213-1:2007, Clause 6 (see also 6.1 in this part of ISO 18213). These steps, which apply to process liquids as well as calibration liquids, are summarized as follows.

a) Observe the differential pressure (ΔP) produced by the liquid of interest and record the additional information on ambient conditions, etc., that is required to perform the necessary standardization calculations.

b) Compute the height of liquid, at measurement temperature, that corresponds to the observed pressure measurement as given in Equation (60):

$$H_M = (\Delta P - c_M)/\big[g(\rho_M - \rho_{a,s}) \big]$$ (60)

c) Compute the corresponding (standardized) height at the reference temperature T_r: $H_0 = H_M/(1 + \alpha_{ex}\Delta T_m)$

d) Use the tank's estimated measurement equation to determine the corresponding standardized volume at the calibration reference temperature: $V_0 = \hat{h}(H_0)$ or $V_0 = \mathbf{h}_0'\hat{\boldsymbol{\beta}}$.

e) Compute the estimated uncertainty for the resulting volume determination, V_0, as given by Equation (59).

The estimation of each quantity in Equation (59) is discussed in turn. The quantity $\partial\hat{h}(H_0)\big/\partial(H_0)$ has already been discussed.

It is possible to estimate the variance of H_0 by means of the steps outlined in 6.1 [see Equation (9)] as given in Equation (61):

$$\mathbf{var}(H_0) = (H_0)^2 \Big[\mathbf{var}(\Delta P)\big/(\Delta P - c_M)^2 + \mathbf{var}(\rho_M)\big/(\rho_M - \rho_{a,s})^2 \Big]$$ (61)

The major components of uncertainty in this equation are those for ΔP and ρ_M. Uncertainty in the measurement ΔP depends upon both the capabilities of the tank's measurement system (manometer) and the number of readings, and is determined at the time of measurement. If the same equipment and procedures are used for measurements of process liquid as were used for calibration, then the estimate of $\mathbf{var}(\Delta P)$ obtained at the time of calibration may also be suitable for process measurements (see 6.1).

Unlike the situation with a calibration liquid, the density, ρ_M, of a process liquid is not known *a priori* and it is necessary, therefore, to estimate it at the time of measurement. This can be done either analytically in the laboratory or by means of in-tank measurements. An appropriate estimate of $\mathbf{var}(\rho_M)$ depends upon the method used to determine ρ_M. When the procedure of ISO 18213-6 is used to determine liquid density (at its tank temperature) from in-tank measurements of pressure, the corresponding variance estimate $\mathbf{var}(\rho_M)$ shall be computed as indicated in ISO 18213-6:2008, Clause 7. If the density is determined by chemical analysis, then the laboratory is expected to provide a suitable estimate of the variance for the density value that it reports. If the density is reported by the laboratory at some standard reference temperature, a suitable interpolation equation is required to determine the density of the liquid at its in-tank measurement temperature.

The remaining quantities required to compute $\mathbf{v\hat{a}r}(V_0)$ are $\hat{\sigma}^2$, $\sum_j\big(\mathbf{H}_j'\mathbf{H}_j\big)^{-1}$ and $\sum_j\hat{\boldsymbol{\theta}}_j\hat{\boldsymbol{\theta}}_j'$. These quantities are best computed at the time the model is fit to the calibration data as indicated in 7.3.

Once the reference volume, V_0, and its estimated variance have been computed, it is sometimes necessary to determine the corresponding quantities at measurement temperature T_m. At its measurement temperature, the estimated reference volume V_0 becomes as given in Equation (62):

$$V_M = (1 + 3\alpha_{ex}\Delta T_m)\hat{h}(H_0)$$
$$= (1 + 3\alpha_{ex}\Delta T_m)V_0 \tag{62}$$

where $\Delta T_m = T_m - T_r$, and the estimated variance of V_M is as given in Equation (63):

$$\mathbf{v\hat{a}r}(V_M) = (1 + 3\alpha\Delta T_m)^2 \mathbf{v\hat{a}r}(V_0) \tag{63}$$

where $\mathbf{v\hat{a}r}(V_0)$ is given by Equation (59).

8.2.2 General procedural considerations

8.2.2.1   Beginning with an observation of pressure, it is possible to combine all the indicated calculations for determining V_M into a single equation of the form as given in Equation (64):

$$V_M = (1 + 3\alpha_{ex}\Delta T_m)\hat{h}(H_0)$$
$$= (1 + 3\alpha_{ex}\Delta T_m)\hat{h}\left\{(\Delta P - c_M)\Big/\left[G(\rho_M - \rho_{a,s})(1 + \alpha_{ex}\Delta T_m)\right]\right\} \tag{64}$$

Thus, it is possible to obtain the volume, V_M, of process liquid at its measured temperature directly from the measurement equation at reference temperature by means of Equation (64).

8.2.2.2   The step described in 8.2.2.1 yields the volume, V_M, of the process liquid at its measurement temperature, T_m. It is sometimes necessary to relate the volume of this liquid to its volume at some other temperature (e.g., the ambient temperature in the laboratory, say T_3). This calculation requires the density of the liquid at both temperatures T_m and T_3, and is accomplished by means of the usual equation for conservation of mass as given in Equation (65):

$$V_3 = V_M\rho_M/\rho_3 \tag{65}$$

The uncertainty of V_3 is determined from the uncertainties of the quantities on the right-hand side of Equation (65) by means of standard propagation-of-variance methods.

8.2.2.3   Except for measurement variability in ΔP, the equations presented in 8.2.1 do not take account of variability in the prover measurements. It is possible for prover measurements to be a significant source of variability in certain high-precision measurement and calibration systems. In case it is necessary to consider the prover measurement variability for a particular application, this can be done by means of appropriate propagation-of-variance calculations. Since prover measurement variability can safely be ignored in nearly all applications, and calculations are complex, the estimation of prover measurement uncertainty is not pursued in this part of ISO 18213.

8.2.3 Heel volume

The reference volume, V_0, obtained from the measurement equation is a measure of only the calibrated volume of the tank below the point H_0. In particular, the volume of any liquid in the tank at the start of a calibration (the heel) is not included in the estimate V_0 unless this volume has been independently measured and incorporated into the cumulative volumes obtained during the calibration process. Thus, the reference volume, V_0, is an accurate measure of actual or total (i.e., contained) volume of the tank below the point H_0 (at reference temperature) only if the heel volume of the tank has been incorporated into the calibration process or the tank was completely empty at the start of the calibration exercise.

Although determinations of transfer volumes are independent of any liquid in the tank at the start of the calibration exercise, determinations of contained volume are typically required for inventory purposes. It is, therefore, important to incorporate accurate determinations of heel volume into estimates of the calibration and measurement equations if these equations are used to make volume determinations for inventory purposes.

8.3   Transfer volumes

A transfer volume is simply the difference between two measurements of contained volume. Let V_1 denote the before-transfer (initial) volume and let V_2 denote the after-transfer (final) volume. It is assumed for convenience that the transfer is out of the tank, in which case V_1 is the larger of the two volumes and the transfer volume is the difference as given in Equation (66):

$$\Delta V = V_1 - V_2 \tag{66}$$

In terms of the measurement equation, the transfer volume is expressed as given in Equation (67):

$$\begin{aligned} \Delta V &= \hat{h}(H_1) - \hat{h}(H_2) \\ &= \mathbf{h}'_1 \hat{\boldsymbol{\beta}} - \mathbf{h}'_2 \hat{\boldsymbol{\beta}} \\ &= (\mathbf{h}_1 - \mathbf{h}_2)' \hat{\boldsymbol{\beta}} \end{aligned} \tag{67}$$

where $\mathbf{h}'_i$ for $i = 1$, 2, is the (row) vector that corresponds to the standardized height, H_i, from which V_i is determined.

If the measurement errors, ε_i, in the two observations of height are assumed to be independent, the estimated variance of ΔV is given by Equation (68), which is derived with the aid of Equation (59):

$$\begin{aligned} \hat{\mathbf{var}}(\Delta V) = 2\hat{\sigma}^2 + r^{-2}(\mathbf{h}_1 - \mathbf{h}_2)' \left[\hat{\sigma}^2 \sum_{j=1}^{r} (\mathbf{H}'_j \mathbf{H}_j)^{-1} + (r+1)\left(\sum_{j=1}^{r} \hat{\boldsymbol{\theta}}_j \hat{\boldsymbol{\theta}}'_j \right) \right] (\mathbf{h}_1 - \mathbf{h}_2) + ... \\ ... + \left[\partial \hat{h}_1(H_1) / \partial(H_1) \right]^2 (H_1)^2 \left[\mathbf{var}(\Delta P_1) / (\Delta P_1 - c_{M,1})^2 + \mathbf{var}(\rho_{M,1}) / (\rho_{M,1} - \rho_{a,s})^2 \right] + ... \\ ... + \left[\partial \hat{h}_2(H_2) / \partial(H_2) \right]^2 (H_2)^2 \left[\mathbf{var}(\Delta P_2) / (\Delta P_2 - c_{M,2})^2 + \mathbf{var}(\rho_{M,2}) / (\rho_{M,2} - \rho_{a,s})^2 \right] \end{aligned} \tag{68}$$

The first two terms on the right-hand side of Equation (68) represent the combined uncertainty in the predicted volume difference ΔV that is attributable to

a)   the prediction uncertainty due to the measurement equation,

b)   the variability in the difference between the two new observations.

The last two terms on the right-hand side of Equation (68), represent the respective contributions to the total variance of the two height determinations, H_1 and H_2, as propagated through the estimated measurement equation. All terms in Equation (68) are computed exactly as outlined in 8.2 for contained volumes.

Except for segments between those in which H_1 and H_2 fall, inclusive, the components of the vector $(\mathbf{h}_1 - \mathbf{h}_2)'$ are zero. In particular, the components are zero for any initial segments common to the two measurements. Since the heel volume does not enter into the computations of Equations (67) and (68), it is clear that determinations of transfer volume are independent of the tank's heel volume.

Annex A
(informative)

Examples of diagnostic plots

Figure A.1 shows a cumulative plot in which height is plotted against volume. Data from seven calibration runs are plotted. Prior to plotting, all data were standardized to a predefined set of reference conditions and aligned to compensate for differences in heel volume between successive runs. The plot shows that initially height increases rapidly with volume. The rate of increase gradually decreases until approximately 1 500 l. Thereafter, the increase in height with volume appears to be approximately linear. Figure A.2 shows the same data as Figure A.1, except that volume has been plotted against height. The cumulative plots show the gross departure from linearity in the height-volume relationship at the bottom of the tank, but finer details are obscured by the wide plotting range on the vertical axis.

Key
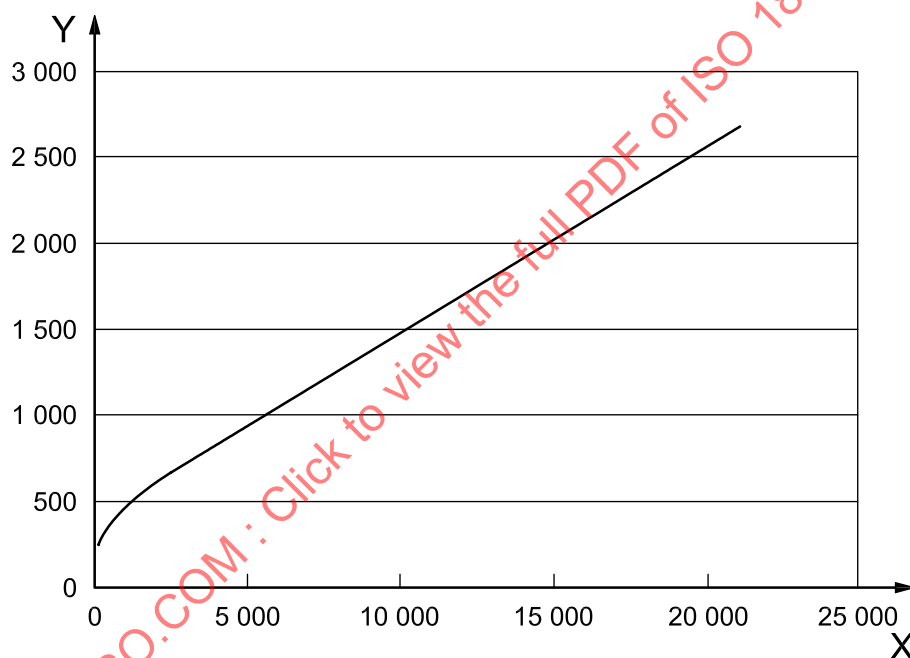
X   volume, expressed in litres

Y   height, expressed in millimetres

NOTE   The curve includes the data of seven runs.

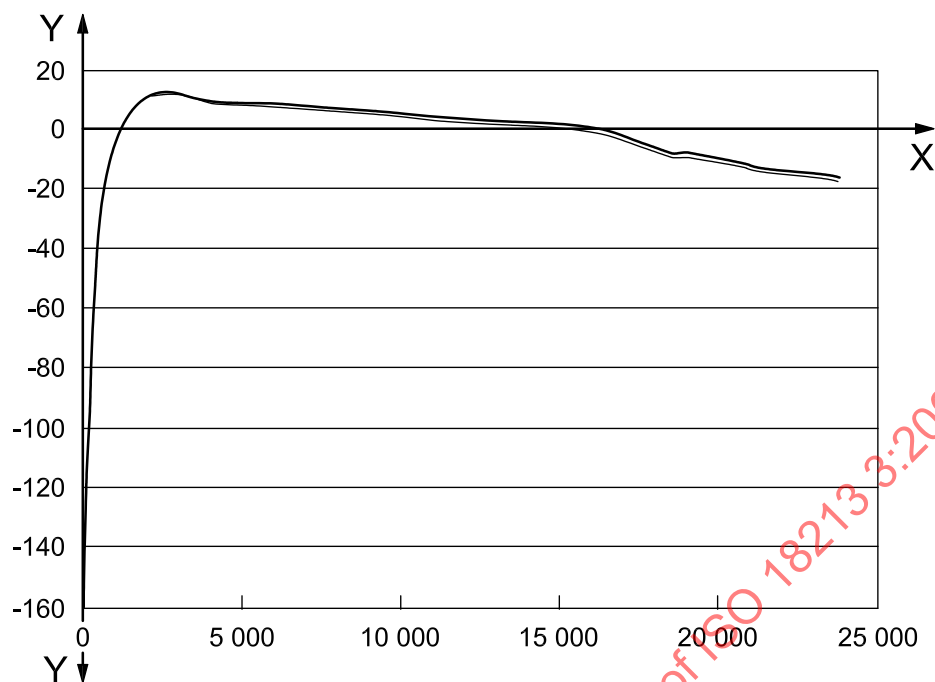**Figure A.1 — Cumulative plot of height vs. volume**

Key
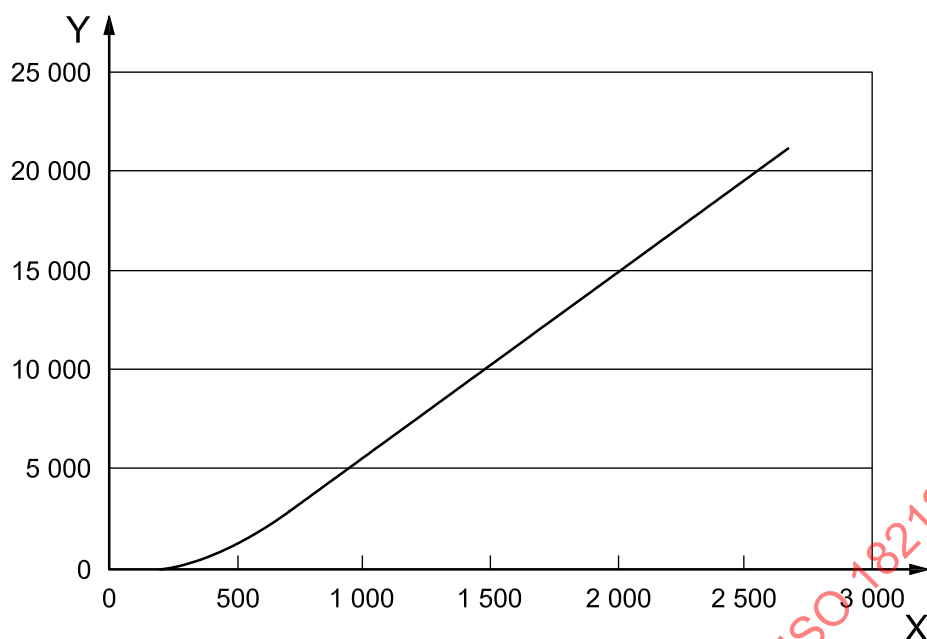
X   height, expressed in millimetres

Y   volume, expressed in litres

NOTE   The curve includes the data of seven runs.

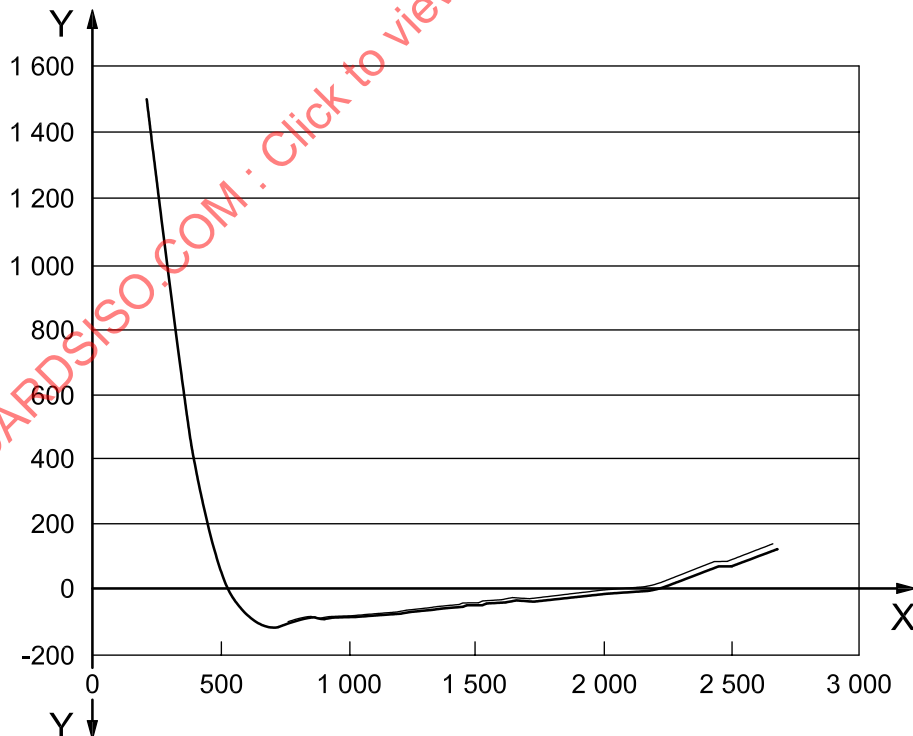**Figure A.2 — Cumulative plot of volume vs. height**

Figure A.3 shows a profile variation plot created from the data of Figure A.1. This plot shows the residuals obtained by fitting a linear regression equation to the combined data (height vs. volume) of all seven runs. As a result, the vertical range has been decreased from approximately 2 700 mm in Figure A.1 to 160 mm in Figure A.3. The reduced range on the vertical axis makes it possible to see detail in the profile that is not evident in the cumulative plot. The tank's profile is not nearly as linear as it appears in Figure A.1, and several distinct regions are evident in the portion of the tank that appears to be linear in the cumulative plot. For this reason, the profile variation plot is quite helpful for identifying segments when one is fitting a calibration or measurement equation to the data. Figure A.4 is similar to Figure A.3, except that the plot shows the residuals obtained by regressing volume on height. The two plots reveal the same information and the choice of which to use is largely a matter of personal preference.

Key

X   volume, expressed in litres

Y   residual height, expressed in millimetres

NOTE   The curve includes the data of seven runs.

Figure A.3 — Profile variation plot of height vs. volume
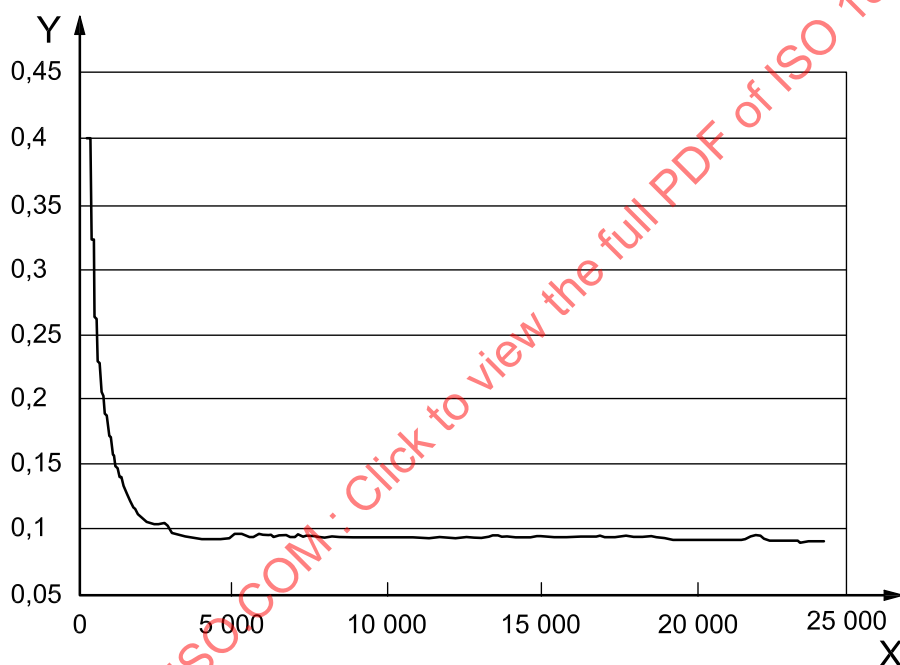

Key

X   height, expressed in millimetres

Y   residual volume, expressed in litres

NOTE   The curve includes the data of seven runs.

Figure A.4 — Profile variation plot of volume vs. height

35

Figures A.5 and A.6 show incremental slope plots of the same data shown in Figures A.1 to A.4. The first expresses incremental changes in height (the change in height per unit change in volume) for each volume increment added to the tank during a calibration run, while the second shows incremental changes in volume for each observed change in height. Figures A.5 and A.6 show the slopes of the calibration and measurement equations, respectively, at each calibration increment. Both plots provide the same information, but the measurement scales are typically such that the second is generally easier to read.

Incremental slope plots reveal the fine detail in the tank's profile and are, therefore, very helpful for identifying or confirming the locations of pipes, agitators and other internals that have a local effect on the free cross-sectional area of the tank. Both plots reveal an abrupt change in profile at a height of approximately 700 mm or a volume of approximately 2 000 l. (Note that this feature is more evident in Figure A.6 than in Figure A.5.) The engineering drawings for the tank confirm that an agitator exists at this height in the tank. Another major structural feature is evident at a height of approximately 2 500 mm. The jagged appearance in the profile between approximately 850 mm and 2 200 mm is caused by internal heating and cooling coils. The fact that all seven runs show these features confirms that the variation is due to structural features in the tank and not to measurement variation. The features are so repeatable across runs that it would be possible (but tedious and perhaps unnecessary) to model individual coils when fitting a measurement equation to these data.
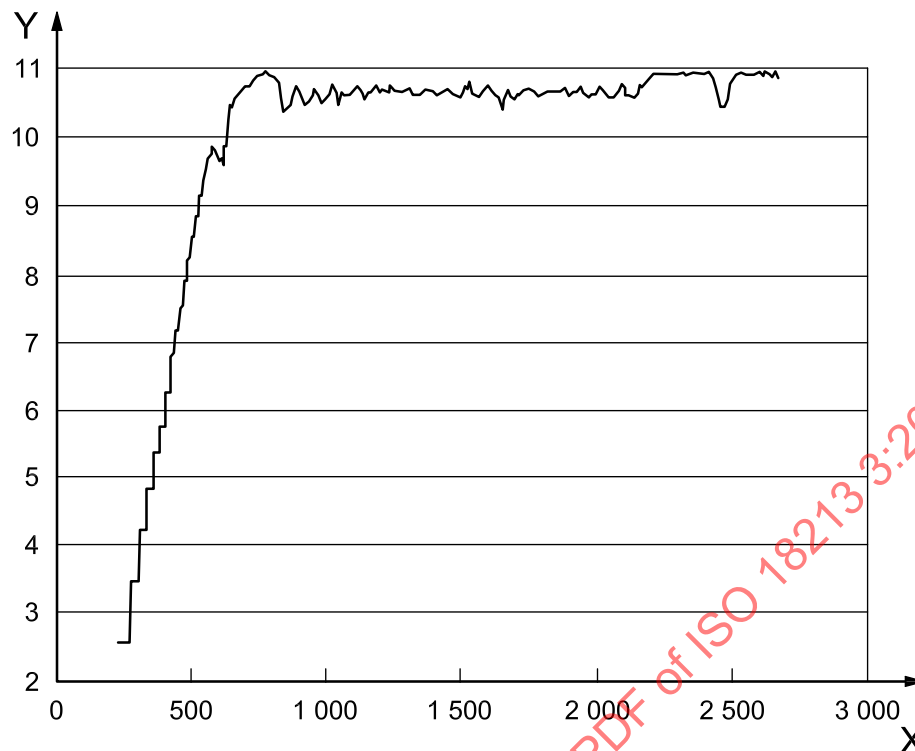

Key

X   volume, expressed in litres

Y   slope

NOTE   The curve includes the data of seven runs.

Figure A.5 — Incremental slope plot of height vs. volume

Key

X   height, expressed in millimetres

Y   slope

NOTE   The curve includes the data of seven runs.

Figure A.6 — Incremental slope plot of volume vs. height

In this example, the incremental slope plots for all runs are nearly identical, indicating that the (standardized) data are in very good agreement. Run-to-run discrepancies revealed by an incremental slope plot are an indication that one or more of the runs are anomalous. In particular, if the slopes for one run differ from those of other runs at an isolated point or two, the plot provides a strong indication that one or more measurements are erroneous in the anomalous run.

The incremental slope plot, particularly that of the measurement equation (Figure A.6), is perhaps the most useful of all plots for "fine-tuning" the segments identified with the aid of a profile variation plot. The previously mentioned features at approximately 700 mm and 2 200 mm require special attention, as does the region at the top of the heel between 700 mm and 900 mm, approximately. Moreover, the region at the top of the coils between approximately 2 150 mm and 2 400 mm is distinct from the coil region. By examining the calibration data with the aid of suitable incremental slope plots, it is possible to make quite precise determinations of the segment boundaries (cut points) used for model fitting.

Since incremental slope plots show derivative (slope) information, they are also very useful for identifying the degree of the polynomial that fits the data of a given segment of the calibration or measurement equation. Consider Figure A.6, which shows incremental changes in volume for observed changes in height (the measurement form). Segments in which the slopes are constant can be fit with a linear equation. Segments in which the slopes are linear can be fit, at least initially, by a second-degree polynomial, and similarly for higher degrees. The slopes in Figure A.6 are quite linear in the initial segment ranging from 250 mm to 700 mm, approximately. This suggests that a second-degree polynomial fits the measurement equation in this region of the tank. Similarly, a linear equation should fit the measurement equation in the region at the top of the coils between 2 150 mm and 2 400 mm, approximately. It should be clear, especially from Figure A.6, that incremental slope plots are quite useful, not only for identifying segment boundaries, but also for determining the degree of the polynomial that fits the data in a particular segment. Clearly, the amount of trial and error required to achieve a good fit to a set of calibration data can be greatly reduced with the aid of incremental slope plots.

As the name implies, comparison plots are used primarily to compare two equations. Figure A.7 shows a plot of the data from two runs (run 6 and run 7) in which a fit to one (run 7) was selected as the reference function for plotting purposes. These runs are part of the same series of runs shown in Figures A.1 to A.6. It is clear from Figure A.7 that these runs differ from each other by an amount that is significantly greater than differences among runs shown in previous figures. Further investigation reveals that both run 6 and run 7 are, in fact, anomalous. In a complete calibration exercise, the reasons for these anomalies should be investigated to determine whether or not data from these runs are suitable for inclusion in future analyses. Provided that they are large enough, the differences shown in the comparison plot are also revealed by a profile variation plot. However, it is unlikely that differences such as those shown in Figure A.7 will be revealed by an incremental slope plot because the two profiles are so nearly parallel.

Key
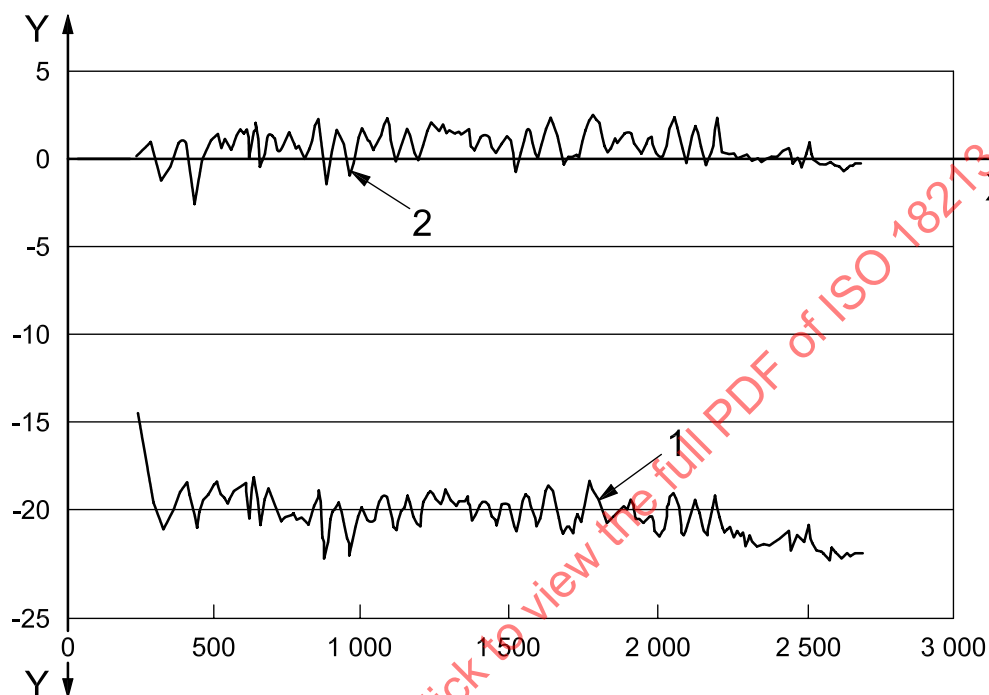
X   height, expressed in millimetres

Y   residual volume, expressed in litres

1   run 6

2   run 7

Figure A.7 — Comparison plot

After segment boundaries were identified and the degree of the polynomial for fitting each segment was established, the measurement equation (height vs. volume) was fitted to the data of plots in Figures A.1 to A.6. The residuals from this fit are shown in Figure A.8, together with the segment boundaries (vertical lines) used to define the fit. The residual traces for each run are approximately linear, indicating that little improvement is to be expected by defining additional segments. At capacity (approximately 22 000 l), the residuals exhibit a spread of approximately 8 l. For this fit, confidence limits for the predicted volume at capacity are approximately ± 5 l. The corresponding prediction error is in the order of 0,02 %, a result that is quite acceptable in any safeguards programme.

Figure A.9 shows the same information as Figure A.8, but results are presented in the form of a comparison plot. To create this plot, the fitted measurement equation was selected as a reference function and the observed (standardized) calibration data were plotted relative to this function. The purpose of this plot is to demonstrate that the residual plot is simply a particular type of comparison plot.