

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI[®]

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

NOTE TO USERS

This reproduction is the best copy available

UMI

**Knowledge-Assisted Stochastic Evaluation
of Sampling Errors in Mineral Processing Streams**

by

Chefi Ketata

A Thesis Submitted to the Faculty of Engineering
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Mining Engineering

APPROVED:

Dr. Maria C. Rockwell, Supervisor

Dr. Stephen D. Butt

Dr. Denis Riordan

Dr. Mary M. Poulton, University of Arizona, External Examiner

DALHOUSIE UNIVERSITY - DALTECH

Halifax, Nova Scotia, Canada

1998



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-39321-6

Canada


DALTECH LIBRARY

“AUTHORITY TO DISTRIBUTE MANUSCRIPT THESIS”

TITLE:

Knowledge-Assisted Stochastic Evaluation
of Sampling Errors in Mineral Processing Streams

The above library may make available or authorize another library to make available individual photo/microfilm copies of this thesis without restrictions.

Full Name of Author: Chefi Ketata
Signature of Author: 
Date: September/18/1998

DEDICATION

The Author would like to dedicate this thesis:

- To the memory of his grandparents, Ali and Aïcha Ketata, and his father, Abderrazzak Ketata.
- To his grandparents, Abdessalam Charfeddine and Jamila Ketata.
- To his other family members.
- To his instructors, teachers, and professors at home, and at the primary and secondary schools and the universities he attended.

TABLE OF CONTENTS

Dedication.....	iii
Table of Contents.....	iv
List of Tables.....	xi
List of Figures.....	xii
List of Abbreviations and Symbols.....	xvi
Acknowledgments.....	xxxi
Abstract.....	xxxii
1. Introduction.....	1
1.1. Background.....	1
1.1.1. Stream Sampling in Mineral Processing Plants.....	1
1.1.2. Modeling.....	4
1.2. Objective.....	8
1.3. Structure.....	9
2. Stream Sampling.....	14
2.1. Stream Sampling Procedure.....	14
2.2. Selection of the Stream Variables.....	16
2.3. Sampling Errors.....	19
2.4. Continuous Selection Error.....	22
2.4.1. Fundamental Error.....	26
2.4.1.1. Composition Factor.....	28
2.4.1.2. Liberation Factor.....	29
2.4.1.3. Particle Shape Factor.....	30
2.4.1.4. Granulometric Factor.....	30
2.4.1.5. Maximum Particle Size.....	31
2.4.2. Grouping and Segregation Error.....	31
2.4.2.1. Constitution Heterogeneity.....	32

2.4.2.2.	Grouping Factor.....	32
2.4.2.3.	Segregation Factor.....	33
2.4.3.	Long-Range Heterogeneity Fluctuation Error.....	33
2.4.4.	Periodic Heterogeneity Fluctuation Error.....	33
2.5.	Materialization Error.....	34
2.5.1.	Increment Delimitation Error.....	35
2.5.2.	Increment Extraction Error.....	36
2.5.3.	Sample Preparation Error.....	38
3.	Stochastic Modeling.....	40
3.1.	Normal Probability Distribution.....	40
3.2.	Central Limit Theorem.....	42
3.3.	Random and Systematic Errors.....	43
3.3.1.	Accuracy.....	44
3.3.2.	Precision.....	44
3.3.3.	Representativeness.....	45
3.4.	Probabilistic Selecting Process.....	45
3.5.	Stationary Time Series Models.....	46
3.5.1.	White Noise Processes.....	46
3.5.2.	Autoregressive Processes.....	47
3.5.3.	Moving Average Processes.....	49
3.5.4.	Dual Relationship between $AR(p)$ and $MA(q)$ Processes.....	53
3.5.5.	Autoregressive Moving Average Processes.....	54
3.6.	State Space Models.....	57
3.7.	State Representation of Input Variables.....	60
3.8.	Transfer from Component Flowrates to Compositions.....	62
3.9.	Variance Matrices.....	65
3.9.1.	Variance Matrix of Input Vector.....	65
3.9.2.	Variance Matrix of State Vector.....	66

3.9.3.	Variance Matrix of Output Vector.....	67
3.9.4.	Variance Matrix of Composition Vector.....	67
3.10.	Covariance Matrices.....	67
3.10.1.	Covariance Matrix for Input Vector.....	68
3.10.2.	Covariance Matrix for State Vector.....	69
3.10.3.	Covariance Matrix for Output Vector.....	71
3.10.4.	Covariance Matrix for Composition Vector.....	71
3.10.5.	Cross-Covariance Matrix between Input and State Vectors.....	72
3.10.6.	Cross-Covariance Matrix between State and Input vectors.....	73
4.	Stochastic Evaluation of Sampling Error.....	75
4.1.	Sampling Procedure.....	75
4.2.	Selection of Stream Variables.....	76
4.3.	Discretization of the Component Flowrate.....	78
4.4.	Discretization Error.....	80
4.5.	Sampling Error.....	83
4.6.	Sampling Error Variance.....	86
4.7.	Correlation between Sample Increments.....	89
4.8.	Comparison with Gy's Approach.....	92
4.9.	Numerical Examples.....	96
5.	Sampling Error Filtering by Material Balance.....	111
5.1.	Material Balance.....	111
5.2.	Flotation Process Unit Models.....	112
5.2.1.	Mixer Model.....	114
5.2.2.	Separator Model.....	115
5.3.	Mass Conservation Equations.....	116
5.4.	Data Redundancy.....	117
5.5.	Least-Squares Estimation.....	118
5.6.	Variance Analysis.....	121

5.7.	Stream Variables.....	122
5.8.	Discretization of the Component Flowrates.....	125
5.9.	Discretization Errors.....	126
5.10.	Sampling Errors.....	128
5.11.	Sampling Error Variance.....	131
5.12.	Correlation between Samples Increments.....	135
5.13.	Numerical Examples.....	138
6.	Fuzzy Evaluation of Sample Reliability.....	148
6.1.	Fuzzy Logic.....	148
6.2.	Sampling Correctness Conditions.....	148
6.2.1.	Cutter Geometry.....	150
6.2.2.	Cutter Speed.....	151
6.2.3.	Cutter Layout.....	153
6.2.4.	Cutter Path.....	156
6.3.	Fuzzy Sets.....	157
6.3.1.	Cutter Geometry.....	157
6.3.2.	Cutter Speed.....	158
6.3.3.	Cutter Layout.....	159
6.3.4.	Cutter Path.....	161
6.4.	Fuzzy Evaluation of Sample Reliability.....	162
6.5.	Fuzzy Correction of Erroneous Composition Value.....	164
6.6.	Numerical Example.....	166
7.	Expert Systems <i>Sampling Correctness Inspector</i> and <i>Sampling Error Evaluator</i>	168
7.1.	Expert Systems.....	168
7.2.	Expert System Evaluation.....	173
7.3.	Expert System Development.....	176
7.3.1.	Identification.....	176

7.3.2.	Conceptualization.....	180
7.3.3.	Formalization.....	182
7.3.4.	Implementation.....	183
7.3.5.	Testing.....	184
7.4.	Expert System <i>Sampling Correctness Inspector</i>	186
7.4.1.	Solution Process.....	186
7.4.2.	Example Run of the <i>Sampling Correctness Inspector</i>	189
7.4.3.	Validation.....	194
7.5.	Expert System <i>Sampling Error Evaluator</i>	195
7.5.1.	Solution Process.....	195
7.5.2.	Example Run of the <i>Sampling Error Evaluator</i>	198
7.5.3.	Validation.....	201
8.	Conclusion.....	202
8.1.	General Conclusions.....	202
8.1.1.	Stream Sampling.....	202
8.1.2.	Stochastic Modeling.....	202
8.1.3.	Stochastic Evaluation of Sampling Error.....	203
8.1.4.	Sampling Error Filtering by Material Balance.....	204
8.1.5.	Fuzzy Evaluation of Sample Reliability.....	205
8.1.6.	Expert systems <i>Sampling Correctness Inspector</i> and <i>Sampling Error Evaluator</i>	206
8.1.7.	Comments.....	209
8.2.	Recommendations.....	210
8.2.1.	Uncertain Knowledge.....	210
8.2.2.	Cross-Stream Cutters.....	211
8.2.3.	Sampling Instrumentation.....	211
8.2.4.	Material Handling.....	211
	References.....	212

Appendices.....	221
Appendix A - APL Workspaces of the <i>Sampling Error Filter</i>	221
Appendix B - Influence of Composition Variations on Sampling Errors in Mineral Processing Streams.....	223
B1. Introduction.....	223
B2. Stream Sampling.....	224
B3. Sampling Error.....	225
B4. Sampling Error Variance.....	227
B5. Comparison with Gy's Method.....	230
B6. Conclusion.....	234
Appendix C - Influence of Composition and Mass Flowrate Variations on Sampling Errors in Mineral Processing Streams.....	236
C1. Introduction.....	236
C2. Stream Sampling.....	237
C3. Sampling Error.....	238
C4. Linearization of Average Stream Composition.....	239
C5. Linearization of Average Sample Composition.....	240
C6. Sampling Error Variance.....	242
C7. Comparison with Gy's Method.....	247
C8. Conclusion.....	250
Appendix D - Knowledge Base of the Expert System <i>Sampling Correctness Inspector</i>	252
Appendix E - Validation of the Expert System <i>Sampling Correctness Inspector</i>	274
E1. Case Study Number 1.....	274
E2. Case Study Number 2.....	278
E3. Case Study Number 3.....	283
E4. Case Study Number 4.....	288

E5.	Case Study Number 5.....	293
E6.	Case Study Number 6.....	298
E7.	Case Study Number 7.....	303
E8.	Case Study Number 8.....	307
E9.	Case Study Number 9.....	312
E10.	Case Study Number 10.....	317
E11.	Case Study Number 11.....	322
E12.	Case Study Number 12.....	327
E13.	Case Study Number 13.....	332
Appendix F - Knowledge Base of the Expert System <i>Sampling Error</i>		
	<i>Evaluator</i>	337
Appendix G - Validation of the Expert System <i>Sampling Error Evaluator</i>350		
G1.	Case Study Number 1.....	350
G2.	Case Study Number 2.....	353
G3.	Case Study Number 3.....	357
Appendix H - Computer Disk361		

LIST OF TABLES

1.1.	Strengths and weaknesses of some modeling techniques.....	6
2.1.	Nature of continuous selection errors.....	26
2.2.	Nature of materialization errors.....	35
4.1.	Comparison of the proposed method with Gy's procedure.....	96
7.1.	Twelve faulty sampling situations.....	194
A1.	Functions of the APL workspace CHAPTER4.AWS.....	221
A2.	Main function of the APL workspace CHAPTER5.AWS.....	222
B1.	Comparison of the proposed method with Gy's procedure.....	232
C1.	Comparison of the proposed method with Gy's procedure.....	250

LIST OF FIGURES

1.1.	Methods of combining paradigms for hybrid systems.....	7
1.2.	Block diagram of the <i>Intelligent Sampling Control System (ISCS)</i>	13
2.1.	Four principal steps of the incremental sampling process.....	15
2.2.	Sample increment extraction by a cross stream cutter.....	16
2.3.	Component flowrate per unit of width $x(y,t)$ at two different y positions.....	18
2.4.	Stream sampling.....	19
2.5.	Homogeneous constitution and distribution.....	23
2.6.	Heterogeneous constitution but homogeneous modular distribution.....	23
2.7.	Heterogeneous constitution and maximum distribution heterogeneity.....	24
2.8.	Heterogeneous constitution and natural distribution homogeneity.....	24
3.1.	Example of normal probability distribution.....	41
3.2.	Component flowrate as a simulated $AR(2)$ series, $(1 - 0.7B + 0.5B^2)x_t = \xi_t$, plus 25 t/h.....	49
3.3.	Autocorrelation function of the series $(1 - 0.7B + 0.5B^2)x_t = \xi_t$	50
3.4.	Component flowrate as a simulated $MA(2)$ series, $x_t = (1 - 0.7B + 0.5B^2)\xi_t$, plus 25 t/h.....	52
3.5.	Autocorrelation function of the series $x_t = (1 - 0.7B + 0.5B^2)\xi_t$	52
3.6.	Component flowrate as a simulated $ARMA(1,1)$ series, $(1 - 0.5B)x_t = (1 - 0.8B)\xi_t$, plus 25 t/h.....	56
3.7.	Autocorrelation function of the series $(1 - 0.5B)x_t = (1 - 0.8B)\xi_t$	56
4.1.	Discretization operator.....	79
4.2.	Decomposition of $x(t)$ into a sum of a moving average and a stationary noise.....	82
4.3.	Scheme of the process leading to the sampling error variance.....	84
4.4.	Autocorrelation functions of the AR series.....	97

4.5.	Autocorrelation functions of the four <i>MA</i> series.....	98
4.6.	Autocorrelation functions of the two <i>ARMA</i> series.....	99
4.7.	Relative sampling error variance of the <i>AR</i> series for a single-increment sample.....	100
4.8.	Relative sampling error variance of the <i>MA</i> series for a single-increment sample.....	100
4.9.	Relative sampling error variance of the <i>ARMA</i> series for a single-increment sample.....	101
4.10.	Relative sampling error variance of the <i>AR</i> series for a five-increment sample....	102
4.11.	Relative sampling error variance of the <i>MA</i> series for a five-increment sample...102	
4.12.	Relative sampling error variance of the <i>ARMA</i> series for a five-increment sample.....	103
4.13.	Relative sampling error variance of the <i>AR</i> series for a ten-increment sample....	103
4.14.	Relative sampling error variance of the <i>MA</i> series for a ten-increment sample....	104
4.15.	Relative sampling error variance of the <i>ARMA</i> series for a ten-increment sample.....	104
4.16.	Relative error for the <i>MA</i> (5) series, $x_i = (1 - 0.9B - \dots - 0.9B^5)\xi_i$	105
4.17.	Relative error for the <i>MA</i> (15) series, $x_i = (1 - 0.9B - \dots - 0.9B^{15})\xi_i$	106
4.18.	Relative error for the <i>MA</i> (30) series, $x_i = (1 - 0.9B - \dots - 0.9B^{30})\xi_i$	106
4.19.	Relative error for the <i>MA</i> (45) series, $x_i = (1 - 0.9B - \dots - 0.9B^{45})\xi_i$	107
4.20.	Ratio $\sigma_e^2(n, k)/\sigma_e^2(1, 48)$ where $nk = 48$ for the <i>AR</i> series.....	109
4.21.	Ratio $\sigma_e^2(n, k)/\sigma_e^2(1, 48)$ where $nk = 48$ for the <i>MA</i> series.....	109
4.22.	Ratio $\sigma_e^2(n, k)/\sigma_e^2(1, 48)$ where $nk = 48$ for the <i>MA</i> series.....	110
5.1.	Delay node.....	113
5.2.	Mixer node.....	113
5.3.	Separator node.....	113

5.4.	Junction node.....	113
5.5.	Flotation cell in flowsheet form.....	114
5.6.	Flotation cell in node form.....	114
5.7.	One product unit.....	116
5.8.	Process leading to the sampling error variance matrix.....	130
5.9.	Flotation circuit in flowsheet form.....	138
5.10.	Flotation circuit in node form.....	139
5.11.	Autocorrelation functions of the <i>AR</i> series.....	140
5.12.	Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ of the <i>AR</i> series for the first component of the stream 1 before material balancing.....	141
5.13.	Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ of the <i>AR</i> series for the first component of the stream 1 after material balancing without covariances.....	142
5.14.	Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ of the <i>AR</i> series for the first component of the stream 1 after material balancing with covariances.....	143
5.15.	Relative error variance before and after material balancing for the first component of the stream 1.....	144
5.16.	Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ before and after material balancing for the first component of the stream 1.....	145
5.17.	Relative error for the first component of the process streams: input, 1, and 4....	146
5.18.	Relative error for the first component of the process streams: 6, 8, and 10.....	146
6.1.	Example of crisp and fuzzy sets “reliable sample”.....	149
6.2.	Correct cutter geometry.....	151
6.3.	Correct cutter layout.....	154
6.4.	Incorrect cutter layout with stream thickness larger than cutter length.....	155
6.5.	Incorrect cutter layout with cutter idle position close to the stream.....	155
6.6.	Incorrect cutter layout with cutter reversing its motion soon.....	156
6.7.	Incorrect cutter path with taking part of the stream all of the time.....	156

6.8.	Definition of fuzzy set “very reliable cutter geometry”.....	158
6.9.	Cutter opening before (a) and after (b) obstruction by sticky materials.....	158
6.10.	Definition of fuzzy set “very reliable cutter speed”.....	159
6.11.	Definition of fuzzy set “very reliable cutter layout”.....	160
6.12.	Definition of fuzzy set “very reliable sampling direction”.....	161
6.13.	Factor importance ruler.....	163
6.14.	Definition of fuzzy set “very reliable sample”.....	165
6.15.	Definition of fuzzy set “very reliable sample”.....	167
7.1.	Components of an expert system.....	170
7.2.	Expert system development stages.....	177
7.3.	Network diagram of the <i>Sampling Correctness Inspector</i>	188
7.4.	Network diagram of the <i>Sampling Error Evaluator</i>	197
B1.	Stream sampling by a cross stream cutter.....	225
B2.	Autocorrelation functions of the three <i>MA</i> series.....	228
B3.	Relative sampling error variance σ_R^2 for a one-increment sample.....	229
B4.	Relative sampling error variance σ_R^2 for a five-increment sample.....	229
B5.	Relative sampling error variance σ_R^2 for a ten-increment sample.....	230
B6.	Relative error e_R for the <i>MA(5)</i> series.....	233
B7.	Relative error e_R for the <i>MA(15)</i> series.....	233
B8.	Relative error e_R for the <i>MA(45)</i> series.....	234
C1.	Relative sampling error variance σ_α^2 for a one-increment sample.....	245
C2.	Relative sampling error variance σ_1^2 for a one-increment sample.....	246
C3.	Relative sampling error variance σ_R^2 for a five-increment sample.....	246
C4.	Relative sampling error variance σ_R^2 for a ten-increment sample.....	247
C5.	Relative error e_R	249

LIST OF ABBREVIATIONS AND SYMBOLS

LATIN LETTERS

A :	1) sample section. 2) transition matrix of a linear time-invariant system X_t . 3) total cutter opening area.
a :	1) composition. 2) constant.
\tilde{A} :	1) fuzzy set “reliable sample”. 2) fuzzy set “very reliable cutter geometry”.
a^* :	overall mean value of composition a .
A_E :	effective cutter opening area.
A_i :	transition matrix of a linear time-invariant system $X_{i,t}$.
$a(i)$:	average composition of increment i .
a_j :	average composition of increment j .
a_j^c :	composition of a component c in a stream j .
a_j^{c*} :	true value of the composition of a component c in a stream j .
a_L :	mass fraction of the critical component in the lot.
APL:	A Programming Language.
AR:	APL function that computes the sampling errors of a simulated AR process as detailed in Chapter 4.
AR:	AutoRegressive.
AR48:	APL function that computes the sampling errors of a simulated AR process for a shift of eight hours as detailed in Chapter 4.
ARMA:	APL function that computes the sampling errors of a simulated $ARMA$ process as detailed in Chapter 4.

<i>ARMA</i> :	AutoRegressive Moving Average.
<i>ARMA48</i> :	APL function that computes the sampling errors of a simulated <i>ARMA</i> process for a shift of eight hours as detailed in Chapter 4.
<i>ARMA(p,q)</i> :	autoregressive moving average process of order (p,q) .
<i>AR(p)</i> :	autoregressive process or model of order p .
<i>ASCII</i> :	American Standard Code for Information Interchange.
<i>a_S</i> :	sample average composition.
<i>a_T</i> :	true average composition of the total amount of material considered in the time interval T .
<i>B</i> :	1) sample section. 2) time delay. 3) input matrix of a linear time-invariant system X_t .
<i>b</i> :	constant.
\tilde{B} :	fuzzy set “very reliable cutter speed”.
<i>B_i</i> :	input matrix of a linear time-invariant system $X_{i,t}$.
<i>C</i> :	name of a computer programming language.
<i>C</i> :	1) sample section. 2) output or observation matrix of a linear time-invariant system X_t . 3) set of the cutter sampling factors.
<i>c</i> :	1) composition factor. 2) stream component index.
\tilde{C} :	fuzzy set “very reliable cutter layout”.
<i>CE</i> :	continuous selection error.
<i>CE₁</i> :	short range heterogeneity fluctuation error.
<i>CE₂</i> :	long range heterogeneity fluctuation error.
<i>CE₃</i> :	periodic heterogeneity fluctuation error.
<i>C_i</i> :	output or observation matrix of a linear time-invariant system $X_{i,t}$.
<i>c_i</i> :	cutter sampling factor number i .

cm:	centimeter.
$\cos(z)$:	cosine of z .
$\text{cov}[x]$:	covariance of the variable x .
CV :	Pearson's coefficient of variation or relative standard deviation of a random variable.
d :	maximum particle size.
\tilde{D} :	fuzzy set "very reliable cutter direction".
DE :	delimitation error.
d_i :	liberation size.
e :	exponent.
e :	1) sampling error. 2) exponential function. 3) measurement error $(X_t - X_t^*)$.
$e_d(i)$:	discretization error of the component flowrate x .
EE :	extraction error.
$e_1(j)$:	sampling error associated with the j th increment (incremental error), that is, the sampling error when taking a single increment in a time interval containing k discrete values of the stream property.
e_R :	relative error.
et al.:	and others.
$E(x)$:	expected value of the random variable x .
F	transition matrix of a linear time-invariant system U_t .
f :	particle shape factor.
FE :	fundamental error.
F_i :	transition matrix of a linear time-invariant system $U_{i,t}$.
$f(x)$:	probability density of the random variable x .
G :	input matrix of a linear time-invariant system U_t .
g :	gram.

g :	granulometric factor.
g/cm^3 :	gram per cubic centimeter.
GE :	grouping and segregation error.
G_i	input matrix of a linear time-invariant system $U_{i,t}$.
H :	transfer matrix from component flowrates to compositions.
h :	hour.
$h(i)$:	heterogeneity contribution.
H_i^j :	diagonal matrix in the row number i and the column number j of matrix H .
i :	1) component index. 2) time index. 3) increment index. 4) index of a cutter sampling factor.
\tilde{I} :	fuzzy set representing the degrees of cutter factors importance.
i_i :	grade of membership of a factor c_i in \tilde{I} .
I_k :	$k \times k$ identity matrix.
$in.$:	inch.
$ISCS$:	Intelligent Sampling Control System.
J :	least-squares criterion.
j :	1) time index. 2) stream index. 3) index of a sample reliability index.
K :	constant of the mixer.
k :	1) time lag. 2) number of streams in a mineral processing system. 3) sampling period.
L :	1) width of mineral processing stream. 2) Lagrangian. 3) stream thickness.

l :	<ol style="list-style-type: none"> 1) liberation factor. 2) number of inputs to a mineral processing system. 3) stream component index. 4) time index of the first sample increment. 5) cutter length.
L_A :	actual cutter path length that is perpendicular to the stream direction and equal to the length of constant speed.
L_T :	total cutter path length that is perpendicular to the stream direction.
M :	<ol style="list-style-type: none"> 1) mass flowrate matrix. 2) common estimator of \bar{M} and M_T.
m :	meter.
m :	<ol style="list-style-type: none"> 1) average mass flowrate. 2) number of measured variables in a mineral processing system. 3) common estimator of \bar{m} and m_T.
\dot{M} :	overall mean value of the mass flowrate M .
\dot{m} :	overall mean value of the mass flowrate m .
\bar{M} :	average of the matrix M over the number of sample increments.
\bar{m} :	average mass flowrate of the slurry over the n sample increments.
M.4:	expert system shell.
MA:	APL function that computes the sampling errors of a simulated MA process as detailed in Chapter 4.
MA :	Moving Average.
MA48:	APL function that computes the sampling errors of a simulated MA process for a shift of eight hours as detailed in Chapter 4.
MAA:	APL function that computes the sampling errors of a simulated MA process as detailed in Appendix C.
$MA(q)$:	moving average process or model of order q .
max:	maximum.

MB :	APL function that computes the sampling errors of a simulated mineral process as detailed in Chapter 5.
ME :	materialization error.
M_i :	1) minimum increment mass. 2) incidence matrix of the mineral process network.
\bar{M}_i :	average flowrate of the reference phase over the n sample increments for the component i .
$M(i)$:	mass flowrate matrix at time index i .
$m(i)$:	mass flowrate of the slurry at time index i .
\min :	1) minimum. 2) minute.
m_j :	mass flowrate of the slurry at time index j .
$M_j(i)$:	$k \times k$ diagonal matrix of average mass flowrates of the reference phases at the time of the extraction of the i th samples increments for the component j .
$m_j(i)$:	average mass flowrates of the reference phase j at the time of the extraction of the i th samples increments.
M_L :	lot mass.
M_S :	sample mass.
m/s :	meter per second.
m_T :	average value of the slurry mass flowrate m in the time interval T .
N :	number of lot increments.
$N(0,1)$:	normal random variable with mean zero and variance one.
n :	1) number of sample increments. 2) number of independent random variables X_1, X_2, \dots , and X_n that have the same distribution function. 3) number of components in each mineral processing stream.

	4)	minimum number of sample increments to respect the requirement of sampling error variance ratio.
n_e :		number of mass conservation equations in a mineral process network.
n_i :	1)	number of independent variables in a mineral process network.
	2)	number of sample increments at time index i .
n_m :		number of measured variables in a mineral process network.
n_v :		number of variables in a mineral process network.
p :	1)	order of an autoregressive process or model.
	2)	time index.
PC:		personal computer.
PE:		preparation error.
$P\{E\}$:		probability of occurrence of the event E .
q :	1)	order of a moving average process or model.
	2)	time index.
R :	1)	recovery of the ore concentration.
	2)	set of sample reliability indices.
r :	1)	residual $(X_t - \hat{X}_t)$.
	2)	sampling error variance ratio.
\tilde{R} :		fuzzy relation.
r_i :		sample reliability index number i .
r_{ij} :		membership grade of the pair (c_i, r_j) .
$r^2(SE)$:		mean square of the total sampling error SE .
$r_0^2(SE)$:		standard of sampling representativeness.
S :		section depicted by the cutter movement.
s :		estimated standard deviation of a random variable.
s_0^2 :		standard of sampling precision.
\tilde{S} :		fuzzy set "very reliable sample".
SCI:		Sampling Correctness Inspector.

SE :	total sampling error.
SEE :	Sampling Error Evaluator.
SEF :	Sampling Error Filter.
SPI :	Sampling Performance Indexer.
$s^2(x)$:	estimated variance of the variable x .
T :	length of the period over which the stream is sampled.
t :	ton.
t :	time.
t_0 :	elapsed time since the beginning of the discretization interval when the cutter is started.
t_d :	discretization period of the component flowrate x .
t/h :	tons per hour.
$t(i)$:	sampling period of increment i .
T_p :	period of a periodic disturbance.
t_j :	sampling period of increment j .
t_p :	stream crossing time of the cutter.
t_s :	time interval between sample increments.
u :	vertical distance between the liberation point of the stream and the plane described by the cutter edges during sampling.
U_{it} :	system input vector number i at time t .
U_i :	system input vector at time t .
u_i :	system input variable.
v :	1) cutter speed. 2) area fraction of the cutter opening.
V° :	variance of $Y^{\circ}(i)$.
v_0 :	discontinuity component of the sampling error variance.
V' :	transpose of matrix V .
$\text{var}[x]$:	variance of the variable x .

V_d :	variance matrix of Y_d .
V_e :	variance of sampling error Z_e .
$V_e(n,k)$:	variance of sampling error Z_e as a function of n and k .
$V_1(i)$	covariance between the Y incremental errors of increments separated by ik time intervals t_d .
$v(k)$:	variogram at time lag k .
V_r :	variance matrix of r .
V_{re} :	cross-covariance matrix between r and e .
V_U :	variance matrix of vector U_t .
$V_{UU}(k)$:	covariance matrix of vector U_t at time lag k .
V_{UX} :	cross-covariance matrix between vectors U and X at time lag zero.
$V_{UX}(k)$:	cross-covariance matrix between vectors U and X at time lag k .
V_X :	variance matrix of X_t^m .
V_Y :	variance of $Y(i)$.
V_{YY}^* :	autocovariance of $Y^*(i)$.
$V_{YY}(k)$:	autocovariance of $Y(i)$ at time lag k .
V_ε :	variance matrix of ε .
w :	1) width of cutter opening. 2) simple integral of a variogram.
\tilde{W} :	fuzzy composition.
w' :	double integral of a variogram.
w_i :	grade of membership of a sample reliability index r_i in \tilde{W} .
$W(k)$:	sampling error generator as a function of k .
$w'(k)$:	double integral of a variogram as a function of k .
$w(l)$:	simple integral of a variogram as a function of l .
$W(n,k)$:	sampling error generator as a function of n and k .
X :	sum of independent random variables X_1, X_2, \dots , and X_n that have the same distribution function.

x :	1) composition. 2) mass flowrate of the selected component. 3) constant-speed length fraction of the cutter.
\bar{x} :	overall mean value of the component flowrate x .
\bar{x} :	estimated average of the random variable x .
x_C :	component flowrate in a concentrate.
x_F :	component flowrate in a feed.
$\{X_i\}$:	set of independent random variables X_1, X_2, \dots , and X_n that have the same distribution function.
x_i :	1) component flowrate at time index i . 2) component flowrate of stream i .
x_i^* :	true value of the component flowrate of stream i .
\hat{x}_i :	maximum likelihood estimate of the component flowrate of stream i .
$x(i)$:	discretized value of the component flowrate x at time index i .
$x^*(i)$:	true value of x in the time interval $[(i-1)t_d, it_d]$.
$X_{i,t}$:	system output number i at time t .
$x_{i,t}$:	input component flowrate number i at time t .
x_j^i :	mass flowrate of a component i in a stream j .
x_j^{i*} :	true value of the mass flowrate of a component i in a stream j .
X-ray:	electromagnetic radiation of short wavelengths, able to pass through opaque bodies.
X_t :	1) system output at time t .
X_t :	2) state vector at time t .
x_T :	tailings component flowrate.
x_t :	1) component flowrate at time t . 2) stationary process.
\hat{X}_t :	estimate of X_t .

X_t^m :	system output vector at time t .
\hat{X}_t^m :	estimate of X_t^m .
$x(y,t)$:	mass flowrate of the selected component per unit of stream width at position y and time t .
Y :	1) normal random variable with mean zero and variance one. 2) component flowrate vector.
y :	1) axial position of an increment in the stream width. 2) component flowrate. 3) cutter length fraction.
Y^* :	overall mean value of the component flowrate Y .
Y_d :	discretization error of the component flowrate Y .
$Y_d(i)$:	discretization error of the component flowrate Y at time index i .
$Y_{e1}(j)$:	Y sampling error when taking a single increment in a time interval containing k discrete values of the streams property for the j th increments.
$Y(i)$:	component flowrate vector at time index i .
$Y^*(i)$:	true value of component flowrate vector at time index i .
Y_{it} :	system output vector number i at time t .
Y_t :	system output vector at time t .
y_t :	component flowrate at time t .
$Y(y,t)$:	component flowrate vector per unit of stream width at position y and time t .
$Y_e(t)$:	disturbance around the moving average of the component flowrate Y at time t .
$Y_\mu(t)$:	moving average of the component flowrate Y at time t .
Z :	composition vector.
z :	1) coordinate axis. 2) cutter deviation angle, in degrees, from the correct cutter path that is perpendicular to the stream.

Z^* :	overall mean value of composition Z .
Z_e :	sampling error vector.
$Z(i)$:	composition vector at time index i .
$Z_{i,s}$:	sample average composition vector of component i .
$Z_{i,t}$:	composition state vector number i at time t .
$Z_j(i)$:	k vector of average mass fractions of the component j at the time of the extraction of the i th samples increments.
Z_S :	state vector of average sample compositions for a mineral process network.
Z_T :	vector of true average compositions of the total amount of material considered in the time interval T for a mineral process network.
Z_t :	composition state vector at time t .

GREEK LETTERS

α :	1)	value of $\left(M_j V_x M_j'\right)^{-1}$.
	2)	composition value such as $\mu_{\bar{x}}(\alpha) = 0.5$.
	3)	composition value such as $\mu_{\bar{y}}(\alpha) = 0$.
	4)	constant given by Equation C20.
β :	1)	composition value such as $\mu_{\bar{y}}(\beta) = 1$.
	2)	correlation term neglected by Gy.
Δt :		time interval between two consecutive time indices.
ε :		estimation error ($e-r$).
$\varepsilon(t)$:		disturbance around the moving average of the component flowrate.
ϕ :		autoregressive polynomial coefficient for an $AR(1)$.
ϕ_i :		autoregressive polynomial coefficient number i .
$\Phi_p(B)$:		autoregressive polynomial of order p in B .

γ :	1) grouping factor. 2) composition value such as $\mu_{\tilde{S}}(\gamma) = 1$ and $\gamma = \beta$.
γ_k :	autocovariance function of a random variable at time lag k .
λ :	Lagrange multiplier.
μ :	true unknown mean of a random variable.
μ_0 :	standard of sampling accuracy.
$\mu_{\tilde{A}}(x)$:	membership function characterizing the fuzzy set \tilde{A} .
$\mu_{\tilde{B}}(x)$:	membership function corresponding to the fuzzy set \tilde{B} .
$\mu_{\tilde{C}}(y)$:	membership function of the fuzzy set \tilde{C} .
$\mu_{\tilde{D}}(z)$:	membership function of the fuzzy set \tilde{D} .
$\mu_{\tilde{R}}(c_i, r_j)$:	membership function of the fuzzy relation \tilde{R} .
μ_r :	degree of data redundancy.
$\mu_{\tilde{S}}(a)$:	membership function of the fuzzy set \tilde{S} .
$\mu(t)$:	moving average of the component flowrate $x(t)$.
μ_{ξ} :	mean of the white noise process ξ_t .
π :	3.1459.
$\Pi(B)$:	infinite order AR polynomial of order ∞ in B .
π_i :	inifinite order AR polynomial coefficient number i .
θ_i :	moving average polynomial coefficient number i .
$\Theta_q(B)$:	moving average polynomial of order q in B .
ρ :	average density of the ore.
$\rho_a(k)$:	autocorrelation function of the random variable $a(i)$ at time lag k .
ρ_c :	density of the critical component in the lot.
ρ_g :	mean density of the noncritical components.
ρ_k :	autocorrelation function of a random variable at time lag k .

$\rho(k)$:	1) autocorrelation function of the component flowrate $x(i)$ at time lag k .
	2) autocorrelation function of the composition $a(i)$ at time lag k .
$\rho^*(k)$:	autocorrelation function of $x^*(i)$ at time lag k .
$\rho_m(k)$:	autocorrelation function of the random variable $m(i)$ at time lag k .
σ :	true unknown standard deviation of a random variable.
σ^2 :	1) variance of the component flowrate $x(i)$.
	2) variance of the composition $a(i)$.
$\sigma_{x^*}^2$:	variance of $x^*(i)$.
σ_1^2 :	relative sampling error variance.
σ_a^2 :	variance of the random variable $a(i)$.
σ_d^2 :	variance of the discretization error e_d .
σ_e^2 :	variance of the sampling error e .
$\sigma_e^2(n,k)$:	variance of the sampling error e as a function of n and k .
$\sigma_1(i)$:	covariance between the sampling errors of increments separated by ik time intervals t_d .
σ_m^2 :	variance of the random variable $m(i)$.
σ_R^2 :	relative sampling error variance.
$\sigma^2(x)$:	variance of the variable x .
σ_x^2 :	variance of the variable x .
σ_a^2 :	relative sampling error variance.
σ_ξ^2 :	variance of the white noise process ξ .
τ :	1) starting time of extraction at position y .
	2) time constant of the mixer.
τ' :	ending time of extraction at position y .
ξ :	segregation factor.
ξ_i :	white noise process at time index i .

$\xi_{i,t+1}$ white noise input vector of a linear time-invariant system $U_{i,t}$.
 ξ_t : white noise process at time t .
 $\{\xi_t\}$: white noise process.
 $\Psi(B)$: infinite order MA polynomial of order ∞ in B .
 ψ_i : infinite order MA polynomial coefficient or weight number i .

ACKNOWLEDGMENTS

The Author would firstly like to express his most sincere and grateful thanks to his supervisor, Dr. Maria C. Rockwell, and his guiding committee members, Dr. Stephen D. Butt and Dr. Denis Riordan, for their continued support and guidance throughout the period of this research work. Their enthusiasm and assistance have been much appreciated.

He would like to thank his supervisor, Dr. Maria C. Rockwell, for her financial assistance throughout the duration of this research work.

He is also grateful to the DalTech Office of Graduate Studies and Research for its financial assistance for publishing the results of this thesis.

Finally, he would like to thank his stepfather, Habib Ketata, his mother, Hamida Charfeddine, his father-in-law, Ridha Kallel, his mother-in-law, Nedra Moalla, his wife, Raoudha Kallel, and his other family members for their support and encouragement throughout the duration of this research work.

ABSTRACT

The analysis of the data obtained from stream sampling is a crucial step in understanding the performance of a mineral processing plant. To control the sampling process efficiently, it is very important to minimize sampling errors and estimate them. The factors influencing these errors are divided into two categories: material properties and cutter features. This thesis includes four innovations dealing with knowledge-assisted stochastic evaluation of sampling errors in mineral processing streams. The sampling error is the difference between the sample composition and the composition of the total material that flows during the sampling period.

First, the influence of the data variation on sampling errors is studied. The sampling process is assumed to be correct. Then, the sampling errors are random. Thus, they are described by their variances. A new expression for sampling error variances is developed. It is essential for the application of data reconciliation techniques and for the estimation of process performance indicator reliability. In addition, the proposed evaluation method is compared with that of Gy. For some low sampling periods compared to the correlation length of the component flowrate signal, the approximation by Gy's expression is inaccurate. The difference between the formula developed in this thesis and Gy's expression increases with the number of sample increments.

Secondly, the influence of the data variation on sampling errors throughout a two-stage flotation circuit is analyzed. The material balance technique is used to upgrade the measured variables. The weighted least-squares method is applied to minimize the estimation errors. The sampling errors are evaluated before and after material balancing for comparison. In one case, the covariance terms between the different components in a stream and between the different streams in the flotation circuit are included. In the other case, these terms are not included. This proves that, by including the covariance terms in

the calculation, the variances of the sampling errors are reduced and therefore the reliability of the material balancing is improved. Consequently, a *Sampling Error Filter* is constructed.

Thirdly, fuzzy logic can be employed to assess the sampling performance index that is the sample reliability. The latter is influenced by the sampler features describing the sampling conditions. Fuzzy logic with its intuitive nature and closeness to the natural languages offer significant advantages over traditional approaches in the appraisal of sampling conditions. Fuzzy logic allows sampling situations to be described and processed in linguistic terms such as *very reliable*, *reliable*, *adequate*, *doubtful*, and *very doubtful*. These fuzzy sets lead to the sampling performance index value for each sample increment. It can be used as a weight in further calculations of average stream compositions or correct composition values. Hence, a *Sampling Performance Indexer* is proposed.

Finally, two expert systems, *Sampling Correctness Inspector* and *Sampling Error Evaluator*, are developed. The first one is intended to inspect the correctness of sampling operations in a mineral processing plant. The second one is destined for the evaluation of sampling errors in a mineral processing plant. The knowledge is collected from experts publications in sampling of mineral processing streams in addition to the author's expertise in the considered domain. These expert systems take into account the stream properties, the cutter features, and the sampling manner.

The resulting hybrid system is an intelligent sampling controller. It assists the mineral processing plant operators in their decision-making process to correct sampling conditions and minimize sampling errors.

CHAPTER 1

INTRODUCTION

1.1. BACKGROUND

1.1.1. Stream sampling in mineral processing plants

The analysis of mineral processing plant performances requires the estimation of average values of process variables for given periods of time. The key variables are generally the stream flowrates and compositions. There are three different ways for estimating process variable means:

- 1) Averaging of sensor records.
- 2) Stream sampling and sample analysis.
- 3) Inference from values of other process variables using a model such as mass conservation equations.

Usually the three approaches are simultaneously used since only a few streams are instrumented or accessible for sampling.

Therefore the data reconciliation or filtering techniques are useful. These techniques need a prior knowledge of process behavior and measurement uncertainty. This last aspect of data processing is very important since it is an essential part of the following:

- Data reconciliation techniques (Hodouin and Everell, 1980; Hodouin et al., 1989, 1991).
- Gross error detection (Crowe, 1992).

- Reliability evaluation of the process performance indices that are calculated from the data (Hodouin and Flament, 1991).

The analysis of the data obtained from stream sampling is a crucial step in understanding the performance of a mineral processing plant. To control the sampling process efficiently, it is very important to minimize sampling errors and estimate them. The factors influencing these errors are divided into two categories:

- 1) Material properties.
- 2) Cutter features.

The material properties can be particle composition, mineral liberation, particle size distribution, and particle shape. These factors cannot be controlled due to material heterogeneity. However, the cutter features, such as cutter geometry, speed, layout, and path, are controllable. They can be responsible for gross errors if the cutter is not designed and used according to sampling correctness requirements.

In mineral processing, sampling errors have been extensively studied by Gy (1979, 1988, 1992). Gy's theory of particulate materials sampling contains a detailed analysis of the sources of sampling errors. These errors are related to the material constitution (fundamental error) and distribution (grouping and segregation error) heterogeneities as well as to the sample materialization and preparation processes. His analysis of the errors, related to the heterogeneity of a flowing particulate material, is essentially derived from geostatistical concepts (Matheron, 1965; David, 1977; Journel and Huijbregts, 1978).

The sampling theory of Gy is a reference in mineral processing (Smith, 1985; Lyman, 1986; Bilonick, 1986; Saunders et al., 1989; Pitard, 1992, 1993). Other contributions to

the analysis of sampling errors for particulate materials can be found in Visman (1972), Rose (1983), Merks (1985), and Holmes (1991).

The objective of stream sampling is to collect a sample that represents the average properties of the stream during a given time interval T , for example 1 hour, 1 shift of 8 hours, or 1 day. This sample is obtained by gathering and blending n sub-samples collected during the time interval T . The sample to be analyzed for its composition is called “composite sample” and the sub-samples “sample increments” or simply “increments”.

Gy defines three main strategies for incremental sampling:

- Random sampling: the n increments are extracted at n uniformly distributed random times over the interval T .
- Systematic sampling: the increments are extracted from the stream every $t_s = T/n$ time units.
- Stratified random sampling: the increments are randomly extracted within the n time intervals t_s .

Only the systematic sampling strategy, which is the usual one, is studied here.

A sample should be correct, that is, all the particles should have the same probability of falling into the cutter. Assuming sampling correctness is equivalent to neglecting the errors of increment delimitation, extraction, and reunion as defined by Gy.

A stream is sampled for the determination of its average composition. Usually, the composition is defined, for a set of stream components, as the mass fractions of these components with respect to a reference phase. The latter could be either the overall stream material or a part of it. For example, the following composition definitions are used:

- Mass fraction of solids in the slurry.
- Mass fraction of one metal in the solid phase.
- Mass fraction of one mineral in a given particle size interval.

1.1.2. Modeling

Many areas of research, both pure and applied, require the creation of models to explain or predict phenomena of interest. Such models can be utilized for areas including controlling sampling processes and assessing sampling quality.

Some common elements to these particular modeling applications, and most others, are:

- 1) Acceptable magnitude of errors.
- 2) Amount of time available for model development.
- 3) Amount of data and knowledge available.
- 4) Possibility or necessity of extracting knowledge from the experts or models.
- 5) Need for the models to adapt as circumstances change.

For any given application a model needs to be developed, and implemented, meeting these requirements.

A single modeling technique may not always be able to meet all requirements and for these cases a hybrid approach may be more suitable. The focus of this thesis is on adaptive hybrid model development where all sources of information, including both data and knowledge, are employed to supplement the model development process. All aspects of the model development process are examined, ranging from problem formulation through to implementation, which includes adaptation of some type.

From the model development and implementation perspective many techniques are available. Those of interest here are grouped into two broad categories.

- **Statistical techniques.** These include algorithms for deriving probability, applied statistics, signal processing, and *ARMA* time series modeling. Some techniques provide automatic optimization of the model according to some criteria, such as least squares when the format of the model is already specified.
- **Artificial intelligence techniques.** Examples of these are fuzzy logic and expert systems.

Each of these modeling techniques has associated with it a set of advantages and disadvantages. This leads to its degree of suitability for each niche of problems as shown in Table 1.1 (Gray and Kasabov, 1997).

Most models are developed using only one of the above mentioned techniques. As the problem evolves the chosen technique may become unsuitable due to a shift of focus away from the technique's strengths towards its weaknesses. For example, initially data may be paramount for a model, making an *ARMA* time series modeling ideal, and then later knowledge extraction and conceptual verification may become more important, suiting a fuzzy expert system model.

All the techniques have associated weaknesses and strengths. For some problems, such as sampling, one technique is not capable of delivering all that is required. Therefore, some combination of two or more techniques is more suitable than one used alone.

Table 1.1. Strengths and weaknesses of some modeling techniques.

	Strengths	Weaknesses
Statistical techniques	<ol style="list-style-type: none"> 1) Can offer exact solutions in some cases (for example, least-squares method guarantees minimization of the error criteria). 2) Can use expert knowledge as well as data. 3) Provide degree of confidence in estimates in many cases. 4) Well-developed and understood theory behind methods. 	<ol style="list-style-type: none"> 1) Usually require the formulation of the structure of the model in advance. 2) Many opportunities for misuse (such as overfitting models to inflate correlations). 3) Limited interpretability.
Artificial intelligence techniques	<ol style="list-style-type: none"> 1) Claimed to follow the human reasoning process more closely. 2) Robust within certain ranges. 3) Allow use of both data and expert knowledge. 4) Provide an easily interpreted and verified solution to the problem. 5) Very suitable for representing uncertain knowledge. 	<ol style="list-style-type: none"> 1) Difficult to derive the best membership functions, rules, and the weights attached to the rules. 2) Exact mappings of complex functions require extreme numbers of rules and membership functions, which defeats the issue of interpretability. 3) Do not always behave intuitively.

Three general approaches to such hybridization are possible. Firstly, two or more techniques may be used in a single structure, such as a fuzzy-expert system. The second possibility is to use the techniques as interacting and cooperating modules. For example an expert system may examine and validate mineral processing plant data. Finally, the modules can be used in a chain where one module's outputs become another's inputs. An example of this would be an expert system that processes data to perform some feature extraction and then passes these features on to a statistical system. All of these possibilities are shown in Figure 1.1.

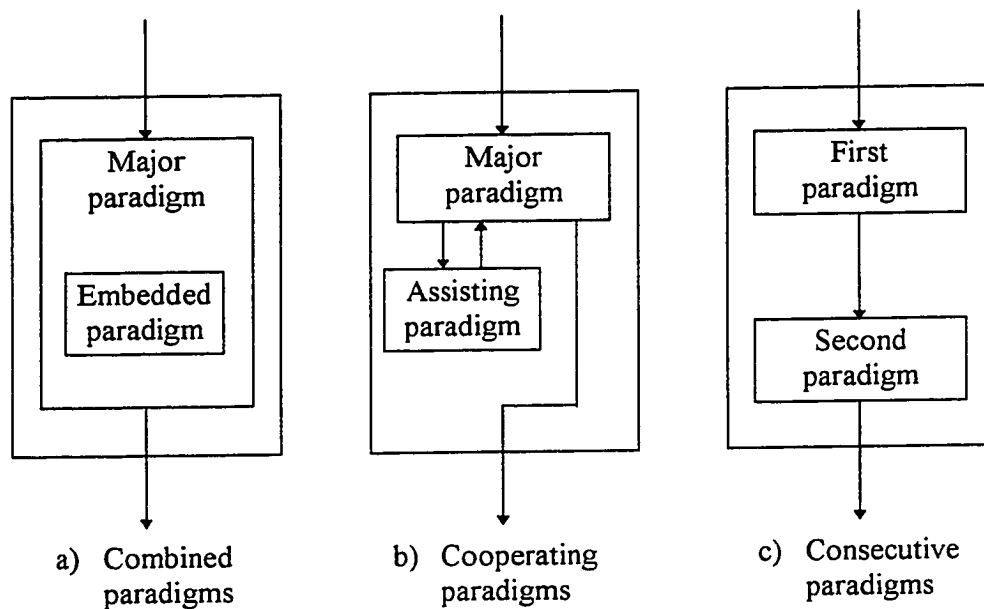


Figure 1.1. Methods of combining paradigms for hybrid systems.

The thesis focuses on the hybrid that involves a combination of statistical (probability, applied statistics, and *ARMA* time series modeling) and artificial intelligence (fuzzy logic and expert systems) techniques. This is seen as a fruitful area of research, largely untapped

due to the lack of researchers conversant with both statistical and artificial intelligence techniques.

1.2. OBJECTIVE

The objective of this thesis is to evaluate sampling errors in mineral processing streams and develop an intelligent sampling controller. Rather than using basic geostatistical concepts, signal processing, fuzzy logic, and expert system techniques are integrated to appraise sampling errors. This approach helps in the understanding of sampling theory and illustrates complementary aspects of the problem. Consequently, the results are explained based on data and experts knowledge in sampling of mineral processing streams.

Mean values of stream compositions in mineral processing plants are frequently estimated by the analysis of a composite sample. The sampling error is the difference between the sample composition and the true composition of the total material that flows during the sampling period.

Besides, the nature of the uncertainties, attached to the procedure of process variable estimation by stream sampling, is analyzed. In this situation, the properties of the flowing material dynamically vary. The knowledge of the structure of this estimation error is fundamental for determining the stream sampling strategy and, as previously said, to reconcile data and calculate performance indices.

If the sampling process is correct, then the sampling errors are assumed to be random. This means that these errors are characterized by given means, that are nil, and given variances, that are never nil. The sampling error variance is determined assuming that the component flowrate random variations are described by their variances and autocorrelation functions. The estimation of the sampling error variance is essential for the

application of data reconciliation or filtering techniques, for gross error detection, and for the estimation of process performance indicator reliability.

As a result, the variances of the sampling errors are computed employing both conventional and knowledge-based techniques. Also a parameterization of the autocorrelation functions based on *ARMA* models is used to generate examples, the results of which are compared to those of Gy's approach. Furthermore, two expert systems, *Sampling Correctness Inspector* and *Sampling Error Evaluator*, are constructed to enhance the decision-making process when practicing sampling. The first one is intended to inspect the correctness of sampling operations in a mineral processing plant. The second one is destined for the evaluation of sampling errors in a mineral processing plant.

1.3. STRUCTURE

This thesis includes six main chapters:

- Stream sampling.
- Stochastic modeling.
- Stochastic evaluation of sampling error.
- Sampling error filtering by material balance.
- Fuzzy evaluation of sample reliability.
- Expert systems *Sampling Correctness Inspector* and *Sampling Error Evaluator*.

First, stream sampling is introduced. The objective of stream sampling is to collect a sample that represents the average properties of the stream during a given time interval. This sample is obtained by gathering and blending the extracted increments following the systematic sampling strategy. In addition, the sampling errors due to material heterogeneity and cutter design are presented.

Secondly, stochastic modeling is investigated. It deals with stochastic processes. A stochastic process is a family of time random variables. The theory of probability and statistics plays such an important role in sampling of mineral processing streams that some knowledge and understanding of the most important and useful concepts are required. Indeed, for instance, normal probability distribution is defined. It is the most common model in use in conventional statistics.

In addition, random and systematic errors are described. Further, probabilistic selecting process is explained. Moreover, stationary time series and state space models and their properties are specified. A time series is an ordered sequence of observations through time. The state space representation of a system is a fundamental concept in modern control theory (Wei, 1990). The state of a mineral process is defined to be a minimum set of information from the present and past. Thus, the future behavior of the system can be completely described by the knowledge of the present state and the present input.

Thirdly, the influence of the data variation on sampling errors is studied for a stream component. Besides, the proposed evaluation method is compared with that of Gy since these two procedures are different. A stream is sampled for the determination of its average composition. The latter is obtained by an averaging process of component mass flowrates. If the stream mass flowrate is constant through time, then the average stream composition is acquired by an averaging process of incremental compositions. The generated algorithms are developed using the programming language APL of Manugistics, Inc. (1992).

Fourthly, the influence of the data variation on sampling errors throughout a two-stage flotation circuit is analyzed. The material balance technique is used to upgrade the measured variables. The weighted least-squares method is applied to minimize the estimation errors. The sampling errors are evaluated before and after material balancing

for comparison. In one case, the covariance terms between the different components in a stream and between the different streams in the flotation circuit are included. In the other case, these terms are not included. The produced algorithms are elaborated using the programming language APL of Manugistics, Inc. (1992).

Fifthly, fuzzy logic can be used to assess the sampling performance index that is the sample reliability. The latter is influenced by the sampler features describing the sampling conditions. Fuzzy logic with its intuitive nature and closeness to the natural languages offer significant advantages over traditional approaches in the appraisal of sampling conditions. Fuzzy logic allows sampling situations to be described and processed in linguistic terms such as *very reliable*, *reliable*, *adequate*, *doubtful*, and *very doubtful*. These fuzzy sets lead to the value of the sampling performance index.

As a result, a value of sampling performance index is attached to each sample increment. It can be utilized as a weight in further calculations of average stream compositions or correct composition values for example. In this way, the degree of sample reliability is estimated by applying fuzzy logic.

Finally, two expert systems for sampling correctness inspection, called *Sampling Correctness Inspector (SCI)*, and sampling error evaluation, named *Sampling Error Evaluator (SEE)*, are illustrated. They are developed through the utilization of an expert system shell M.4 of Teknowledge Corporation (1993, 1995, 1996). The knowledge is collected from experts publications in sampling of mineral processing streams (Gy, 1965, 1979, 1988, 1992; Merks, 1985; Holmes, 1991; Pitard, 1992, 1993) in addition to the author's expertise in the considered domain (Ketata, 1991; Hodouin and Ketata, 1994; Ketata and Rockwell, 1998). These expert systems take into account the stream properties, the sampler features, and the sampling manner.

The expert systems, also known as knowledge-based systems, emulate part of the human reasoning capabilities based on the knowledge of an expert or a specialist to solve problems. The task of developing expert systems is not easy but there are large benefits to be realized if the appropriate approach is taken. As research into the development of these systems continues, the practical benefits are being understood and used in industry.

Given the previous description of the thesis structure, an intelligent sampling controller (see Figure 1.2) is developed. Consequently, the actions taken by the operator to correct sampling conditions are assisted by the intelligent sampling hybrid system composed of:

- 1) *Sampling Error Filter (SEF)*. This is the product of Chapters 4 and 5.
- 2) *Sampling Performance Indexer (SPI)*. This is the product of Chapter 6.
- 3) *Sampling Correctness Inspector (SCI)*. This the product of a part of Chapter 7
- 4) *Sampling Error Evaluator (SEE)*. This is the product of a part of Chapter 7.

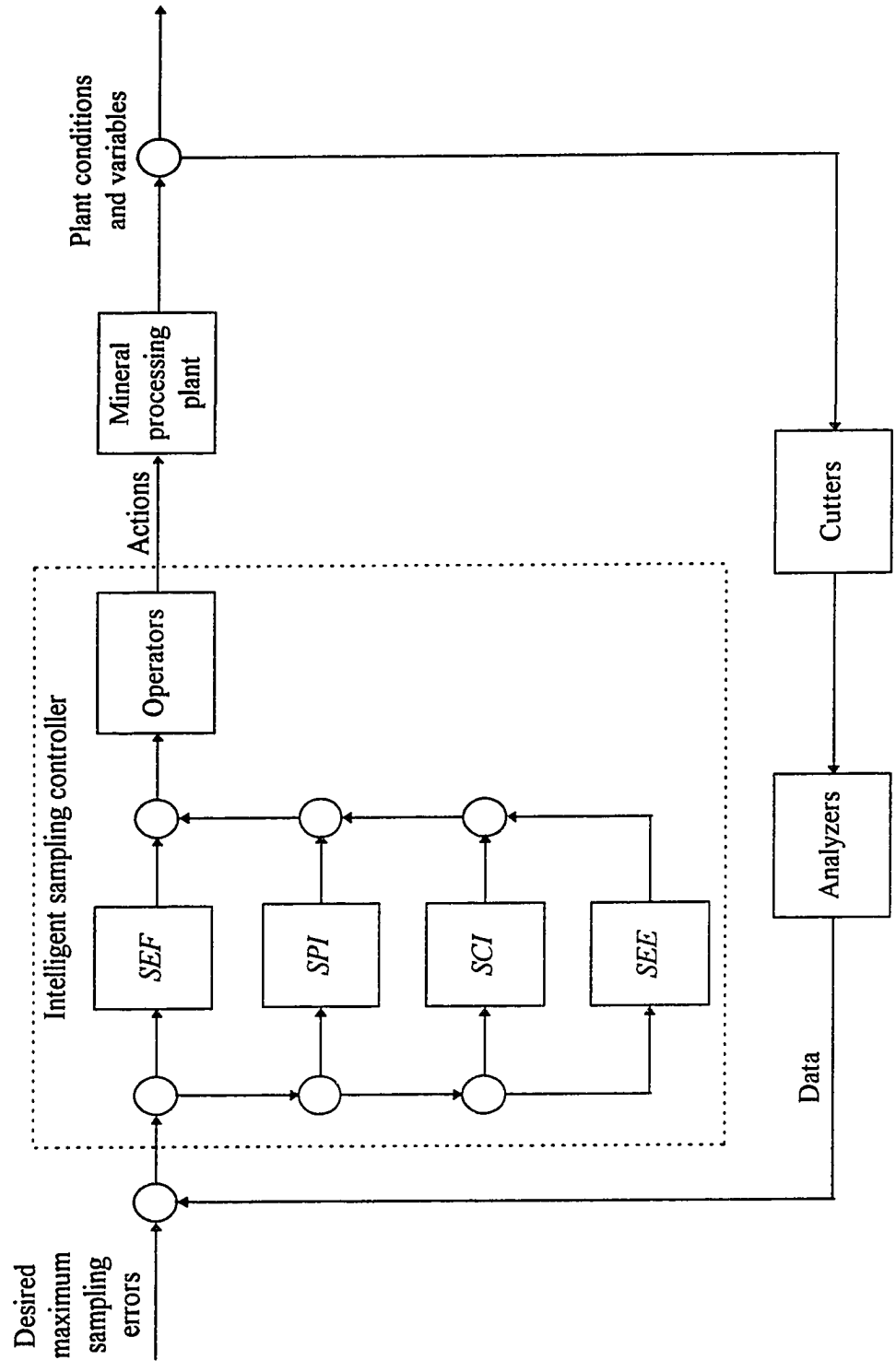


Figure 1.2. Block diagram of the Intelligent Sampling Control System (ISCS).

CHAPTER 2

STREAM SAMPLING

2.1. STREAM SAMPLING PROCEDURE

The objective of stream sampling is to collect a sample. The latter represents the average properties of the stream during a given time interval T (for example 1 hour, 1 shift of 8 hours, or 1 day). This sample is obtained by gathering and blending n sub-samples collected during the time interval T . The sample to be analyzed for its composition is called “composite sample” and the sub-samples “sample increments” or simply “increments”.

Gy (1988) defines three main strategies for incremental sampling:

- 1) Random sampling: the n increments are extracted at n uniformly distributed random times over the interval T .
- 2) Systematic sampling: the increments are extracted from the stream every $t_s = T/n$ time units.
- 3) Stratified random sampling: the increments are randomly extracted within the n time intervals t_s .

Only the systematic sampling strategy, which is the usual one, will be studied here. A plan for incremental sampling, to assure the sample is representative, must take into consideration the degree of heterogeneity in the material stream. The preferred method for best accuracy is sampling from the stream discharge.

Figure 2.1 illustrates the four principal steps of the incremental sampling process:

- 1) Selection of punctual increments before the materialization.
- 2) Delimitation of extended increments.
- 3) Extraction of fragmental increments.
- 4) Reunion of fragmental increments making up the sample.

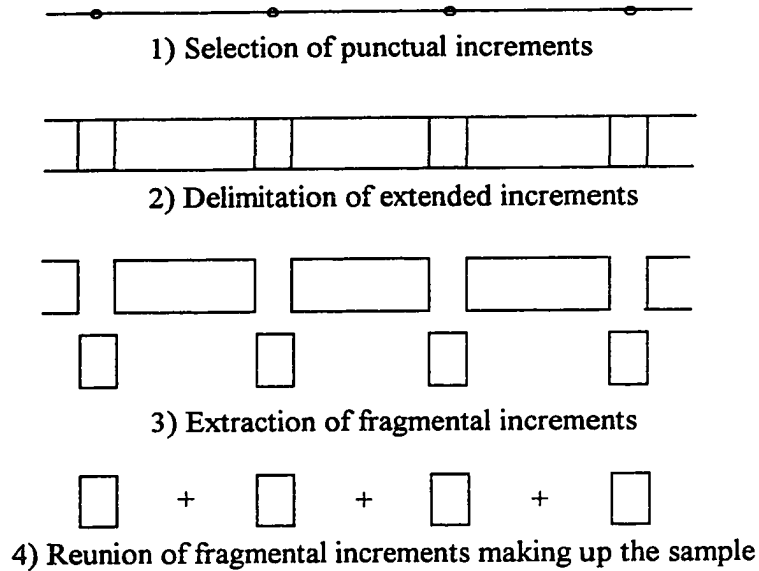


Figure 2.1. Four principal steps of the incremental sampling process.

The typical incremental sampling situation that is analyzed here is the one of a freely flowing stream containing size-distributed particles, possibly suspended in a fluid phase that is a slurry, as shown in Figure 2.2. The sample increments are extracted by a cross stream cutter. The moving increment extractor has an opening of width w , while the stream to be cut is assumed to have a width L (see Figure 2.2).

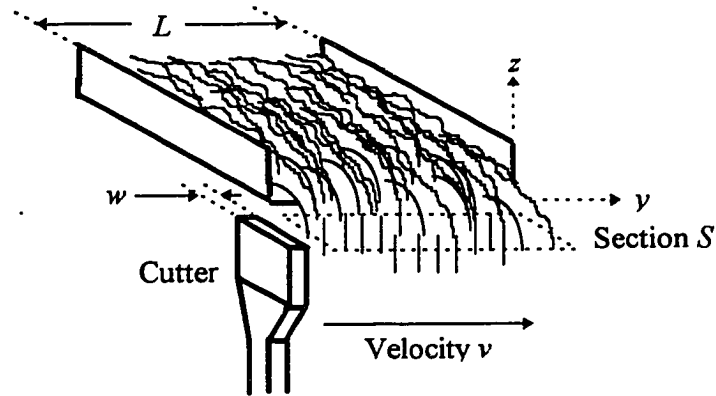


Figure 2.2. Sample increment extraction by a cross stream cutter.

2.2. SELECTION OF THE STREAM VARIABLES

A stream sample is collected for the determination of the average stream composition. Usually, the composition is defined, for a set of stream components, as the mass fractions of these components with respect to a reference phase. The latter could be either the overall stream material or a part of it. For example, the following composition definitions could be used:

- Mass fraction of solids in the slurry.
- Mass fraction of one metal in the solid phase.
- Mass fraction of one mineral in a given particle size interval.

If only one selected component of the stream is analyzed, the mass fraction of that component in the sample is:

$$a_s = \frac{\sum_{j=1}^n t_j m_j a_j}{\sum_{j=1}^n t_j m_j} \quad (2.1)$$

where:

- j is the index of the sample increment.
- m_j is the average mass flowrate of the slurry at the time of the extraction of the j th sample increment.
- a_j is the average mass fraction of the component of interest at the time of extraction of the j th sample increment.
- t_j is the time during which each part of the stream is sampled.

t_j is determined following the expression:

$$t_j = w / v \quad (2.2)$$

As the sampler transfer time from one side of the stream to the other one t_p has been assumed to be constant, t_j is independent of j and as a result:

$$a_s = \frac{\sum_{j=1}^n m_j a_j}{\sum_{j=1}^n m_j} = \frac{1}{\overline{m}} \sum_{j=1}^n m_j a_j \quad (2.3)$$

where \overline{m} is the average flowrate of the slurry over the n sample increments.

Therefore, the measured property of the sample is obtained by an averaging process of the flowrate ma of the component rather than by an averaging process of the mass fraction a of the component. For that reason the stream will be characterized by the component flowrate per unit of stream width: $x(y,t)$.

The variable x has a physical significance and its average value along the stream width can eventually be measured. For instance, using a magnetic flowmeter and a gamma gauge densimeter, and knowing the solid density, the ore flowrate can be continuously monitored. The ore flowrate is the product of the mass flowrate of the slurry with the solid mass fraction.

The variable $x(y,t)$ can be understood as a time variable continuously distributed along the y axis of the stream width (Figure 2.3). Because of the type of sampling device, the heterogeneity along the z axis (Figure 2.2) does not need to be described.

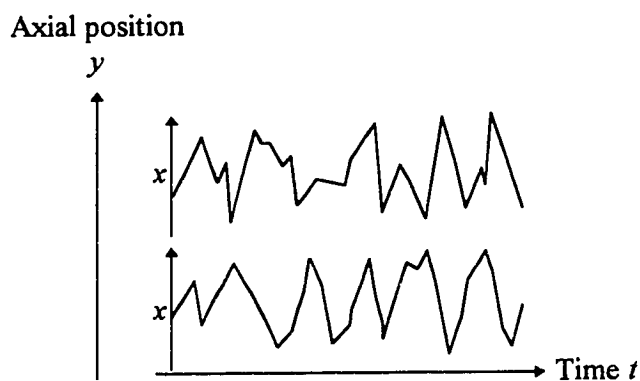


Figure 2.3. Component flowrate per unit of width $x(y,t)$ at two different y positions.

The question of the continuity of x can be discussed because of the discrete nature of the particulate material flowing through the section S (Figure 2.2). Both discrete and

continuous approaches are correct, although the continuous formalism is easier to manipulate. However, to calculate the sampling error the signal x must be discretized since the increment extraction from the stream is typically a discretization process. (Hodouin and Ketata, 1994)

In the following sections, it is assumed that the variables a_s and x obey normal distributions.

2.3. SAMPLING ERRORS

The qualitative problem in stream sampling is described in Figure 2.4. Particles drawn as white elements are matrix gangue materials, and smaller black elements are particles of the mineral grains of interest. Mineral grains are distributed throughout the matrix. Some grains are embedded in the gangue particles and some are free. Most are attached to the exterior of gangue particles. A sample of proper size will contain a statistically valid representation of the whole. (Cooper, 1985)

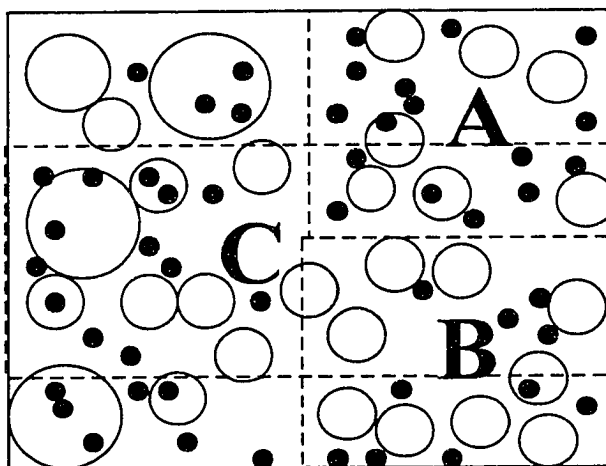


Figure 2.4. Stream sampling.

In Figure 2.4, a total of 50 visible mineral particles is counted. The concept of sample representation is illustrated by taking three sections of the material in Figure 2.4. Parts A and B are equal in area (one-quarter of the whole). However, part C is double in size, overlapping parts of A and B. The mineral grain count for A is 16; B, 10; and C, 24. Errors are respectively +28%, -20%, and -4% on the basis of the counts. The smaller samples deviate significantly from the true value and the results for the larger sample are relatively closer to the true value. This illustrates the relative effect of sample size or weight to characterize a material by samples, and also provides a qualitative example to state the problem of statistical representation.

A sample consists of several particles of different size and mineral content. Consequently, the between-sample variations tend to decrease with increasing sample size. Between-sample variation can be decreased to any desired value by taking larger sample sizes. It must also be considered that increasing sample size results in higher costs to carry out the sampling process because larger weights of sample are handled.

Several points emerge from this illustration. First, sample size is related to the desired level of between-sample variation. Second, to observe the variation in a specific case it is necessary to compare samples of the same weight. The greater the number of samples, the better the statistical estimate of the between-sample variation. Thirdly, to obtain a specified between-sample variation, it is necessary to either fix the sample size and vary the number of samples, or fix the number of samples and vary the sample size to obtain the objective sampling precision. A fourth point is that sample size is influenced by the abundance of the mineral. A high-grade ore stream will be adequately characterized by a smaller sample, all remaining factors being equal, by comparison to a low-grade ore stream.

Errors in sampling will be a function of nonuniformity in the flow stream both in cross section and along the length because of segregation by density and stratification by size. Slurry will classify during flow in a launder as a result of velocity variations, with larger and higher density particles settling to the bottom of the stream. Consequently, the sampling reliability is affected by this classification.

The mass fraction of the selected component in the sample, a_s , is different from the true average composition of the total amount of material considered in the time interval T , a_T . This infers a sampling error e :

$$e = a_s - a_T \quad (2.4)$$

Sampling is said to be correct when each class of particles in a flowing slurry has an equal probability of appearing in the sample. The sample will be composed of increments collected at different times. Choosing appropriate times at which to select these increments to ensure equiprobable sampling is called correct selection. Choosing sampler and sampler path geometry so that sampling remains equiprobable is named correct increment delimitation. Finally, using the sampling equipment in a manner that ensures that sampling remains equiprobable is labeled correct increment extraction.

Even when sampling is correct, however, the sample composition will not be identical to that of the stream sampled because of the probabilistic nature of the sampling process. Gy (1988) and Pitard (1993) considered the total sampling error SE as the sum of:

- 1) Continuous selection error CE .
- 2) Materialization error ME .

$$SE = CE + ME \quad (2.5)$$

Collecting the whole stream would eliminate sampling errors. Nevertheless, the size of the resulting sample would be impractical, and the absence of feed to following units would affect the experimental result. Correct increment materialization and adequate but not excessive sample masses are required to reduce the total sampling error.

2.4. CONTINUOUS SELECTION ERROR

The continuous selection error *CE* is generated by the immaterial selection process and is the result of material heterogeneity. Heterogeneity is defined as the condition of a lot under which all elements of the lot are not strictly identical. Therefore, homogeneity is the zero of heterogeneity and is an inaccessible limit. All particulate materials are essentially heterogeneous.

The definition of heterogeneity can be easily illustrated with a series of simple pictures. Assuming that a population of 36 fragments is represented by 36 small squares. A lot can be represented by a large square of 6×6 adjacent small squares.

Figure 2.5 shows 36 white squares that are strictly identical, corresponding to the definition of a lot with a homogeneous constitution. Furthermore, the distribution is strictly homogeneous. Any subset has exactly the same composition. Without constitution heterogeneity, any form of distribution heterogeneity cannot exist. The constitution heterogeneity is also called the composition heterogeneity.

Figure 2.6 shows 12 horizontal and 12 vertical modules. The constitution is heterogeneous since there are 3 different components represented by black squares, gray squares, and white squares. However, all groups described by the module 3 squares×1 square have strictly the same composition, either horizontally or vertically. Therefore, the distribution shows a form of homogeneity that can be defined as homogeneous modular distribution.

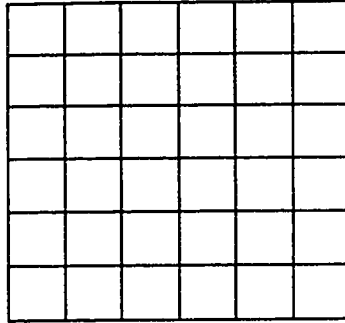


Figure 2.5. Homogeneous constitution and distribution.

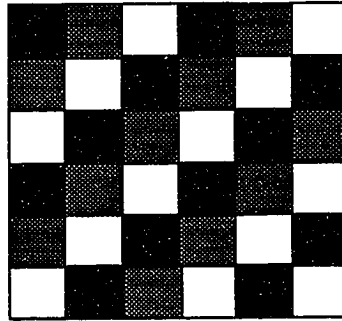


Figure 2.6. Heterogeneous constitution but homogeneous modular distribution.

Figure 2.7 shows 12 black squares, 12 gray squares, and 12 white squares. However, the 3 components are completely separated into three horizontal zones. The constitution is heterogeneous and the distribution heterogeneity is at its maximum, which corresponds to the perfect segregation.

Figure 2.8 shows 12 black squares, 12 gray squares, and 12 white squares, as in Figures 2.6 and 2.7. However, the three components are distributed randomly. It is the state of natural distribution homogeneity. For example, each square is selected at random, one by

one, before being placed inside the frame of the lot. Such a random selection is equivalent to mixing or homogenization, and tends to suppress any correlation between position and composition of the units. A few small clusters can still be detected, which clearly demonstrates that this state of natural distribution homogeneity is far from perfect.

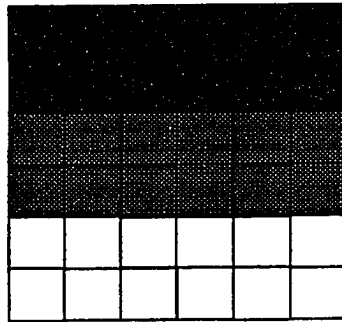


Figure 2.7. Heterogeneous constitution and maximum distribution heterogeneity.

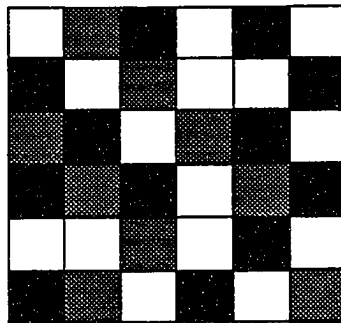


Figure 2.8. Heterogeneous constitution and natural distribution homogeneity.

The continuous selection error is regarded as the sum of three errors (Gy, 1988; Pitard, 1993):

- 1) Short range heterogeneity fluctuation error CE_1 , subdivided into fundamental error FE and grouping and segregation error GE .
- 2) Long range heterogeneity fluctuation error CE_2 .
- 3) Periodic heterogeneity fluctuation error CE_3 .

$$CE = CE_1 + CE_2 + CE_3 \quad (2.6)$$

with

$$CE_1 = FE + GE \quad (2.7)$$

Although short range heterogeneity fluctuation errors cannot be eliminated, they can be reduced by using samples composed of a large number of small increments. The degree to which the error can be reduced is limited by the need for a minimum increment mass, and a reasonable total sample mass.

Collecting the whole stream would eliminate continuous selection errors. However, the size of the resulting sample would be impractical, and the absence of feed to following units would affect the experimental result. Correct increment selection, delimitation, and extraction, together with adequate but not excessive sample masses, are required to reduce continuous selection errors.

Short range heterogeneity fluctuation errors are random variables, characterized by a given average, nil or not, and a given variance, never nil. Nevertheless, long range heterogeneity fluctuation error and periodic heterogeneity fluctuation error are nonrandom or systematic variables, distinguished by an assigned average, never nil, and an assigned variance, nil. In Table 2.1, the continuous selection errors, including their specific properties, are listed.

Table 2.1. Nature of continuous selection errors.

Continuous selection error	Nature	Can cancel out
Fundamental error FE	random	no
Grouping and segregation error GE	random	yes
Long-range heterogeneity fluctuation error CE ₂	nonrandom	yes
Periodic heterogeneity fluctuation error CE ₃	nonrandom	yes

2.4.1. Fundamental error

The fundamental error results from the constitution heterogeneity of the sampled material. The adjective “fundamental” is justified by the fact that, out of all the sampling errors, the fundamental error is the only one that can never cancel out. It is the error that remains when the sampling operation is perfect. The fundamental error is also the only sampling error that can be estimated beforehand.

Gy (1979) has demonstrated that the variance of the fundamental error can be expressed as follows:

$$\sigma^2 (FE) = \left[\frac{1}{M_s} - \frac{1}{M_L} \right] c f g d^3 \quad (2.8)$$

where

- M_s is the sample mass in grams (g).
- M_L is the lot mass in grams (g).
- c is the composition factor in grams per cubic centimeter (g/cm^3).

- l is the liberation factor.
- f is the particle shape factor.
- g is the granulometric factor.
- d is the maximum particle size in centimeters (cm).

Where units are not assigned, the factors are dimensionless.

In most sampling situations,

$$M_s \ll M_L \quad (2.9)$$

and Equation 2.9 can be approximated by

$$\sigma^2 (FE) = clfgd^3 \frac{1}{M_s} \quad (2.10)$$

Generally to obtain a suitably low fundamental error, relatively large samples are required for low grade streams, particularly if they are coarsely ground. As concentrate samples are rich and finely ground, they can be much smaller.

It is important to recognize that fundamental error calculations give only the smallest sample size required, assuming all other sampling operations are error free. In practice this is not true, and reducing other error sources may require larger samples.

2.4.1.1. Composition factor

The composition factor c is the maximum heterogeneity generated by the constituent of interest in the lot. This maximum is reached when the constituent of interest is completely liberated. This factor is shown by Gy (1979) to be given by:

$$c = \frac{1 - a_L}{a_L} [(1 - a_L)\rho_c + a_L\rho_g] \quad (2.11)$$

where a_L is the mass fraction of the critical component in the lot, ρ_c is the density of the critical component in g/cm^3 , and ρ_g is the mean density of the noncritical components in g/cm^3 .

This factor is the most variable in the fundamental error variance equation, ranging from 0.05 g/cm^3 for a very rich concentrate to as much as $2 \times 10^7 \text{ g/cm}^3$ for trace precious metals in ore slurries (Smith, 1985).

In many practical cases, Equation 2.11 can be simplified:

- 1) If the gangue and critical component are not widely different in density, or if $0.3 < a_L < 0.7$,

$$c = \frac{1 - a_L}{a_L} \rho \quad (2.12)$$

where ρ is average density of the ore.

- 2) For concentrates with $a_L > 0.9$,

$$c = (1 - a_L)\rho_g \quad (2.13)$$

3) For feeds and tailings with $a_L < 0.05$,

$$c = \frac{\rho_c}{a_L} \quad (2.14)$$

2.4.1.2. Liberation factor

The liberation factor l is defined as the correcting factor, taking into account the fact that the constituent of interest and gangue are not perfectly separated from one another. This factor varies from zero for perfectly homogeneous material, that is, when there is no liberation, to one for perfectly liberated material. In working with slurries, where the maximum particle size d will rarely exceed 3 mm, and where liberation sizes may be in the range of 100 μm , the liberation factor may vary from about 0.2 for coarse slurries to 1.0 for fine slurries.

Smith (1985) recommended for slurries a value of l equal to 1.0. This is always necessary if samples are to be sized, as size classes are clearly distinct, and there will rarely be enough data to justify a lower value.

Gy (1988) estimates the liberation factor l as a function of the liberation size d_l . When the actual size d is smaller than the liberation size d_l , the liberation is then complete and $l = 1$. However when $d > d_l$, the liberation is not complete and, as a rule of thumb, l can be estimated with the following formula:

$$l = \sqrt{d_l/d} \quad (2.15)$$

In addition, most materials can be classified according to their degree of heterogeneity (Pitard, 1993):

- 1) Very heterogeneous material, $l = 0.8$.
- 2) Heterogeneous material, $l = 0.4$.
- 3) Average material, $l = 0.2$.
- 4) Homogeneous material, $l = 0.1$.
- 5) Very homogeneous, $l = 0.05$.

2.4.1.3. Particle shape factor

The particle shape factor f is also called the coefficient of cubicity. It is a dimensionless factor that gives an idea of how different the shape of a particle is from the ideal cube. The shape factor of a cube is one. For spherical particles, the value of f is 0.524. Most minerals have a shape factor around 0.5. (Pitard, 1993)

2.4.1.4. Granulometric factor

The granulometric factor g is the particle size distribution factor. It is proportional to the third power of the maximum particle size d , defined as the opening of the square mesh retaining about 5% oversize. However, all particles do not have the same size, which is taken into account by a correcting factor g called the granulometric factor. Analysis of 114 materials with size distributions resulting from comminution shows that g ranges from 0.17 to 0.40, with a mean of 0.25 (Gy, 1979). For slurries, Smith (1985) recommended a value of 0.25 for g .

According to Pitard (1993), g would be equal to 1 if we were dealing with perfectly calibrated material. This means that the material is lying in a single screen range. He added that, in practice, the following values for g are mostly encountered:

- 1) With noncalibrated material, for example, material coming out of a jaw-crusher, g is around 0.25.
- 2) With calibrated material, for example, between two consecutive screen openings of a certain series, g is around 0.55.

2.4.1.5. Maximum particle size

The maximum particle size is defined as the mesh size of a standard screen that retains 5% of the material as oversize, which corresponds to a value around 0.25 for g (Gy, 1965). The size can usually be estimated, but can easily be measured if necessary in a preliminary study.

2.4.2. Grouping and segregation error

The grouping and segregation error results from the distribution heterogeneity of the sampled material, illustrated in Figures 2.5 through 2.8, and its variance is directly proportional to three factors:

- 1) Constitution heterogeneity.
- 2) Grouping factor.
- 3) Segregation factor.

The variance of the grouping and segregation error is (Gy, 1988; Pitard, 1993):

$$s^2(GE) = s^2(CE_1) - s^2(FE) = \gamma \xi s^2(FE) \quad (2.16)$$

Here, γ is a grouping factor, dimensionless, always positive, and characterizes the size of the increments making up a sample. In fact, γ is proportional to the size of the increments.

In addition, ξ is a segregation factor, dimensionless, whose value is always between 0 and 1. ξ characterizes the amount of segregation.

In practice, it would be extremely difficult to estimate the product $\gamma\xi s^2(FE)$ and it is never done. The estimation of the variance $s^2(GE)$ by difference as suggested in Equation 2.16 is not precise, especially when $s^2(GE)$ is small when compared to $s^2(CE_1)$ and $s^2(FE)$, which is often the case.

2.4.2.1. Constitution heterogeneity

The constitution heterogeneity of a lot is the heterogeneity that is inherent to the composition of each fragment or particle making up the lot. The greater the difference in composition between each fragment, the greater the constitution heterogeneity is. The constitution heterogeneity is also called the composition heterogeneity.

2.4.2.2. Grouping factor

The grouping factor γ is naturally introduced during the development of the notion of natural distribution homogeneity. Increments or groups making up a sample are most of the time composed of many neighboring particles. In other words, each particle of the lot does not have the same chance of being part of the sample. However, each group does. Therefore, an error is generated. This is taken into account by the grouping factor. The grouping factor is an increasing function of the average number of particles making up each increment in a sample, and is equal to zero when each increment is made up of only one particle.

2.4.2.3. Segregation factor

The distribution heterogeneity of a critical constituent has a natural range between the minimum distribution heterogeneity or distribution homogeneity and the maximum distribution heterogeneity. The maximum distribution heterogeneity is equal to the constitution heterogeneity. For the minimum distribution heterogeneity, the segregation factor ξ equals zero. For the maximum distribution heterogeneity, the segregation factor ξ equals 1. This factor ranges from 0 to 1.

2.4.3. Long-range heterogeneity fluctuation error

Long range heterogeneity fluctuation error is related to long-term drifts in flow rates. It may be associated, for example, with a slowly changing feed grade to a plant over a period of weeks or months. Long-range heterogeneity fluctuation error can generally be avoided if experiment durations are short, for instance, less than a shift of eight hours.

2.4.4. Periodic heterogeneity fluctuation error

Periodic heterogeneity fluctuation error is related to periodic disturbances affecting process streams. The time scale of these disturbances ranges from several seconds to many hours. For example, a stream leaving a level-controlled sump will quite frequently have flow fluctuations with a period of several seconds. In a multistage cleaner circuit with recycle flows, particularly when under current types of automatic control, disturbances have been observed with periods of 30 to 40 minutes (Merks, 1985). Shift periodic fluctuations with an eight-hour period have frequently been observed.

Theoretical schemes relating to selection, that is, timing of sample increment extraction, that provide some protection against periodic heterogeneity fluctuation error have been

proposed (Gy, 1979). For example, consider a stratified random sampling taking a sample increment at some random time in every interval of t_s minutes. This sampling strategy is superior to systematic sampling every t_s minutes if t_s is very close to an exact multiple of the period T_p of a periodic disturbance say within 5%. Otherwise, systematic sampling is better than stratified random sampling.

In practice, the timing of systematic sampling is rarely exact, and the sampling period t_s is rarely an exact multiple of T_p . Nevertheless, it is not safe to sample a process stream that is known to have periodic heterogeneity fluctuations unless this periodic behavior is taken into account. The most dangerous periodic disturbances are those that are unpredictable. They may be indicated during an experiment by cell level surges and similar phenomena. If this happens, the sampling campaign should be stopped until the cause of the periodic disturbance is identified and corrected.

If reasonable precautions are taken to avoid periodic heterogeneity fluctuation errors and samples are correctly delimited and extracted, grouping and segregation error will probably be small, and will certainly be no larger than the fundamental error (Gy, 1979). Therefore, determining sample size from the desired variance of fundamental error should generally control continuous selection error components to within reasonable limits.

2.5. MATERIALIZATION ERROR

So far, errors generated by the heterogeneity of the material making up the lot have been identified. However, the lot has been looked at as a continuous one-dimensional object. The sampling process has been based on the selection of imaginary points within the stream. Nevertheless, the real points are made of fragments or groups of fragments and the discrete nature of these units should be taken into account. The materialization of such groups of fragments provides the increments of a sample.

The materialization is achieved by first implementing an increment delimitation, then an increment extraction, and finally a sample preparation. These operations generate errors. Gy (1988) and Pitard (1993) define the materialization error ME as the sum of:

- 1) Delimitation error DE .
- 2) Extraction error EE .
- 3) Preparation error PE .

$$ME = DE + EE + PE \quad (2.17)$$

Delimitation error and extraction error are random variables, characterized by a given average, nil or not, and a given variance, never nil. Nevertheless, preparation errors are nonrandom or systematic variables, distinguished by an assigned average, never nil, and an assigned variance, nil. In Table 2.2, the materialization errors, including their specific properties, are listed.

Table 2.2. Nature of materialization errors.

Materialization error	Nature	Can cancel out
Delimitation error DE	random	yes
Extraction error EE	random	yes
Preparation error PE	nonrandom	yes

2.5.1. Increment delimitation error

Delimitation deals with the geometry of both the sampler and the sampling path that, ideally, must ensure that any single particle in a flowing stream has a uniform probability

of reporting to the sample. The increment delimitation error results from an incorrect shape of the volume delimiting the increment. This incorrect shape renders nonuniform the probability for a given material element to fall between the boundaries of the extended increment that is the ideal increment.

Gy (1979) notices that accurate delimitation is only possible when the sampled stream is one-dimensional, that is, a sample increment to be cut must represent a thin, ideally one-dimensional, cross-section of the stream. This is only possible when the stream falls vertically and the sampling device intercepts each portion of the stream in direct proportion to its area. This implies that, for cutters traversing a stream in a straight line, the sampler must:

- 1) Have straight and parallel edges.
- 2) Be moved so that, while in motion, its edges are perpendicular to the axis of the stream.
- 3) Move from well outside the stream crossing the entire stream cross-section with constant velocity until again well outside the stream.

These conditions cannot possibly be met exactly by hand sampling. However, if their importance is recognized, a good approximation can be obtained, and a sample obtained by a reasonable number of separate cuts will be essentially free from bias. In addition, physical inspection of the sampling situation will usually allow classification of samples into reliability ranges, for example, reliable, adequate, or doubtful.

2.5.2. Increment extraction error

The increment extraction error results from an incorrect extraction of the increment. The extraction is said to be correct if the rule of the center of gravity is respected. This means

that all particles with their center of gravity inside the boundaries of the correctly delimited increment belong to the increment.

Here every particle, whose center of gravity enters the sampler opening, is considered to enter the sampler and to be retained. The factors to be considered can be divided into factors related to cutter construction and to cutter use.

Cutter edges must be straight, parallel, of equal thickness, and symmetric about the axis of the cutter. The cutter must have enough volume to retain the sample without overflowing, and should be shaped to prevent splashing. The cutter width must be large enough so that coarse particles are not selectively rejected by rebounding from the cutter edge. Gy (1979) concludes, from experimental evidence, that the cutter width should be at least three times the diameter of the largest particles in the slurry, but in no case narrower than 1 cm.

In cutter use, the variable of interest is cutter velocity. At too high a cutter velocity, there is a risk of particles rebounding from the sampler edges over the sampler opening. Gy (1979) recommends a maximum cutting velocity equal to 0.6 m/s. Furthermore, he determines the minimum increment mass in grams (g) as equal to:

$$M_i = \frac{25mw}{9v} \quad (2.18)$$

where:

- m is the mass flow rate in tonnes per hour (t/h).
- w is the cutter width in centimeters (cm).
- v is the cutter velocity in meters per second (m/s).

Thus, the minimum increment mass is obtained:

$$M_I = \frac{125m}{27} \quad (2.19)$$

2.5.3. Sample preparation error

All of the nonselective operations performed on the sample are referred to as preparation stages. These preparation stages are necessary to convey the increments to a predetermined location and to modify them into an appropriate form for the ultimate analytical stages. Such analytical stages are assaying and estimation of the percent solids of a slurry. These preparation stages are classified into various categories such as:

- Transfer of the increment from the cutter to a conveying system.
- Transfer from the conveying system to the next sampling or preparation stage.
- Comminution stages, whose functions are to diminish the particle size d , and to increase the number of particles, for example, crushing, grinding, and pulverizing.
- Wet or dry screening, often in connection with a comminution.
- Filtration of pulps to separate the solid phase from the liquid phase.
- Drying of solids loaded with various amounts of moisture.

In a typical setting, slurry samples are filtered, dried, and weighed. A representative head sample will then be extracted for chemical analyses. These operations can selectively alter the composition of the sample. Consequently, they bias its critical content.

Preparation errors are strictly inherent to those who design, build, operate, and maintain sampling stations.

The preparation errors are introduced by:

- 1) Contamination.
- 2) Loss.
- 3) Chemical or physical alteration.
- 4) Human mistakes.
- 5) Fraud.

CHAPTER 3
STOCHASTIC MODELING

3.1. NORMAL PROBABILITY DISTRIBUTION

The normal probability distribution, also called Gaussian model, is the most common model in use in conventional statistics.

A random variable obeys a normal law of distribution when its probability density $f(x)$ can be written as follows:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]} \quad (3.1)$$

where μ is the true unknown mean, and σ is the true unknown standard deviation, in which the distances $\pm\sigma$ from the mean μ are defined as the x -coordinates of both inflection points of the above function $f(x)$.

However, most of the time, only the estimated average \bar{x} and standard deviation s are accessible. Then examining the normality of a distribution consists of comparing the characteristics of the estimated standard deviation s with those of an ideal model. For example, if the following probabilities are valid:

$$\begin{cases} P\{-s < (x - \bar{x}) < s\} \approx 68\% \\ P\{-2s < (x - \bar{x}) < 2s\} \approx 95\% \\ P\{-3s < (x - \bar{x}) < 3s\} \approx 99.7\% \end{cases}$$

then the distribution is approximately normal. An example of normal probability distribution is illustrated in Figure 3.1.

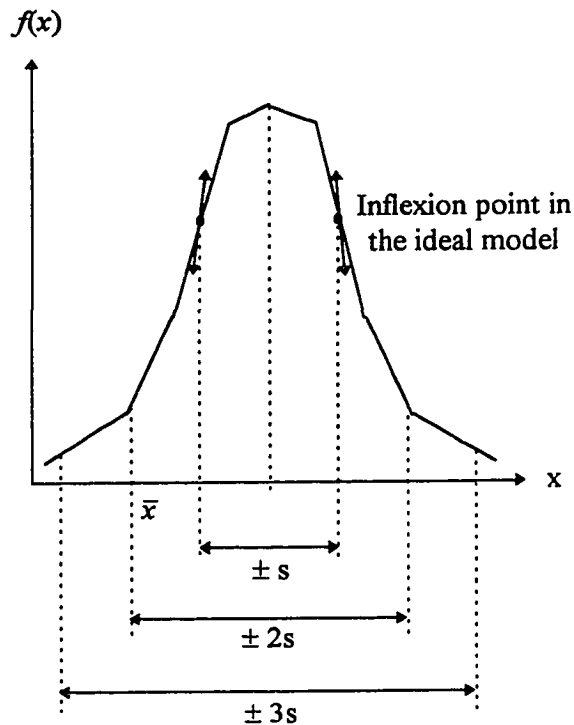


Figure 3.1. Example of normal probability distribution.

Pitard (1993) indicates that the theory of sampling is a preventive tool that may be compared to the possession of an “insurance policy” against major misunderstandings. Here, the mass fraction of a component in a sample of particulate material, a_s , and its corresponding component flowrate, x , are assumed to obey normal distributions following Merks (1985), Gy (1988), and Pitard (1993).

3.2. CENTRAL LIMIT THEOREM

The central limit theorem plays an important role in the theory of sampling (Merks, 1985). In addition, it is useful in simulating normal distributions. It is defined as follows.

Let $\{X_i\}$ be independent random variables that have the same distribution function and therefore the same mean μ and the same variance σ^2 . Let X be:

$$X = X_1 + \dots + X_n \quad (3.2)$$

Then as n tends to infinity

$$Y = \frac{X - n\mu}{\sigma\sqrt{n}} \quad (3.3)$$

tends to a normal random variable with mean zero and variance one, represented as $N(0,1)$.

Furthermore, the random variable X has the following properties:

- 1) X has the mean $n\mu$ and the variance $n\sigma^2$.
- 2) If the variables $X_1, \dots, \text{ and } X_n$ are normal then X is normal.

If the variables $X_1, \dots, \text{ and } X_n$ are not normal, then the second property fails to hold. However, if n is large, then X is approximately normal.

3.3. RANDOM AND SYSTEMATIC ERRORS

The accidental errors, such as long range heterogeneity fluctuation errors CE_2 , periodic heterogeneity fluctuation errors CE_3 , and preparation errors PE , affect the integrity of a sample. All other sampling errors are assumed to be random variables. These are characterized by a given mean, that is nil, and a given variance, that is never nil. Pitard (1993) points out that it is by exaggeration, often because it is convenient, that an error is supposed to be random, with mean nil and variance different from zero, or systematic, with variance nil and mean different from zero. Indeed, all sampling errors have two components:

- 1) A random component characterized by the variance only.
- 2) A nonrandom or systematic component characterized by the mean only.

The variance and the mean of an error are physically complementary, even if they are different properties. When several random variables such as sampling errors are independent in probability, they are cumulative. Thus, it is justified to write the following relationships (Gy, 1988; Pitard, 1993):

- If these errors occur separately:

$$SE = FE + GE + CE_2 + CE_3 + DE + EE + PE \quad (3.4)$$

where SE is the total sampling error.

- When the averages, that are the expected values, of the sampling errors are different from zero, they are additive and their respective sign should be taken into account during the summation:

$$E(SE) = E(FE) + E(GE) + E(CE_2) + E(CE_3) + E(DE) + E(EE) + E(PE) \quad (3.5)$$

- The standard deviations of the sampling errors are not additive, only their corresponding variances are:

$$s^2(SE) = s^2(FE) + s^2(GE) + s^2(CE_2) + s^2(CE_3) + s^2(DE) + s^2(EE) + s^2(PE) \quad (3.6)$$

3.3.1. Accuracy

Accuracy is independent from precision and a sampling selection is said to be accurate when the total sampling error has the absolute value of its mean μ smaller than a certain standard of accuracy μ_0 (Gy, 1988):

$$|\mu| < \mu_0 \quad (3.7)$$

Accuracy is a property of the mean of a given error exclusively. When the total sampling error has its average μ equal to zero, the sampling selection is said to be unbiased (Gy, 1988; Pitard, 1993). Even when carried out in an ideal way, sampling is always biased due to the particulate structure of the material to be sampled. The bias is never strictly zero. It may be negligible or very small but always different from zero. (Pitard, 1993)

3.3.2. Precision

Precision should not be confused with accuracy. A sampling selection is said precise when the total sampling error SE is a little dispersed around its average, regardless of the fact that sampling is biased or not. This means that the variance of the total sampling error is smaller than a certain standard of precision s_0^2 considered as acceptable (Gy, 1988):

$$s^2 \leq s_0^2 \quad (3.8)$$

Precision relates to measuring the variability of a sample around the average of the lot. This measurement is regularly expressed as the variance of the sampling error $s^2(SE)$.

Precision is used in reference to the magnitude of random variations between replicate measurements. It is a qualifying term, for example, low precision, or excellent precision characteristics. It is employed to describe the repeatability (Merks, 1985) or the reproducibility (Merks, 1985; Pitard, 1993) of sampling.

Furthermore, Merks (1985) indicates that the Pearson's coefficient of variation CV is a popular measure for precision. It is also defined as the relative standard deviation. It is a measure for random variations between series of measurements, and numerically equal to the standard deviation as a percentage of the average value for the parameter x :

$$CV = 100 \cdot s(x) / \bar{x} \quad (3.9)$$

3.3.3. Representativeness

A sampling selection is representative when the mean square of the total sampling error $r^2(SE)$, i.e., sum of $E^2(SE)$ and $s^2(SE)$, is smaller than a certain standard of representativeness $r_0^2(SE)$ considered as acceptable (Gy, 1988; Pitard, 1993):

$$r^2(SE) = E^2(SE) + s^2(SE) \leq r_0^2(SE) \quad (3.10)$$

If $r^2(SE) = 0$, $SE = 0$ and sampling is said to be exact (Gy, 1988), that is, unbiased and perfectly reproducible: a limit case never encountered (Pitard, 1993).

3.4. PROBABILISTIC SELECTING PROCESS

Pitard (1993) notes that if all the sampling errors are random, and if the distribution of the mass fraction of the selected component in the sample, α_s , is normally distributed, then α_s

accepts for the central value the mass fraction of the selected component in the lot, a_L . He adds that a systematic sampling error is prevented if the following conditions are fulfilled:

- 1) The entire lot is freely accessible to the sampling tool, so that an equal chance for each constituent of the lot to be part of the sample exists.
- 2) The sampling scheme is impartial, so that an equal chance for each constituent to be part of the sample exists.
- 3) The distribution of the mass fraction a_S obeys a normal distribution, which is an optimistic assumption in the case of trace elements.

3.5. STATIONARY TIME SERIES MODELS

Limited by a finite number of available observations, a finite order parametric model is often constructed to describe a time series process. In this section, the white noise processes and the autoregressive moving average (*ARMA*) models are introduced. The *ARMA* models include the autoregressive (*AR*) models and the moving average (*MA*) models as special cases.

3.5.1. White noise processes

Consider $\{\xi_t\}$ a white noise process. It is a sequence of uncorrelated random variables from a fixed distribution with constant mean μ_ξ equal to the expected value $E(\xi_t)$, usually assumed to be zero, and constant variance $\sigma_\xi^2 = E(\xi_t^2)$. By definition, it immediately follows that a white noise process ξ_t is stationary with the autocovariance function (Box and Jenkins, 1976; Abraham and Ledolter, 1983; Wei, 1990):

$$\gamma_k = E(\xi_t \xi_{t+k}) = \begin{cases} \sigma_\xi^2, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad (3.11)$$

The autocorrelation function described by

$$\rho_k = \gamma_k / \gamma_0 \quad (3.12)$$

equals

$$\rho_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad (3.13)$$

$\{\xi_t\}$ is always referred to as a zero mean Gaussian white noise process. Occasionally, these random variables are also called random shocks (Abraham and Ledolter, 1983).

3.5.2. Autoregressive processes

An autoregressive process or model x_t of order p , that is denoted as $AR(p)$, is given by (Box and Jenkins, 1976; Abraham and Ledolter, 1983; Wei, 1990):

$$\Phi_p(B)x_t = \xi_t \quad (3.14)$$

where x_t is the component flowrate at time t , B is a delay, and $\Phi_p(B)$ is the autoregressive polynomial of order p in B :

$$\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3.15)$$

Since $\sum_{j=1}^p |\phi_j| < \infty$, the process is always invertible. To be stationary, the roots of the polynomial $\Phi_p(B) = 0$ must lie outside of the unit circle. The AR processes are useful in describing situations in which the present value of a time series depends on its preceding values plus a random shock (Wei, 1990).

The p th order autoregressive process $AR(p)$ is

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)x_t = \xi_t \quad (3.16)$$

or

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \xi_t \quad (3.17)$$

To find the autocovariance function, multiply x_{t-k} on both sides of Equation 3.17

$$x_{t-k} x_t = \phi_1 x_{t-k} x_{t-1} + \dots + \phi_p x_{t-k} x_{t-p} + x_{t-k} \xi_t \quad (3.18)$$

and take the expected value

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p}, \quad k > 0 \quad (3.19)$$

where $E(\xi_t x_{t-k}) = 0$ for $k > 0$. Hence, the following recursive relationship for the autocorrelation function is obtained:

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k > 0 \quad (3.20)$$

For illustration, an $AR(2)$ process, $(1 - 0.7B + 0.5B^2)x_t = \xi_t$, is simulated (see Appendices A and H). The white noise series ξ_t are independent normal $N(0,1)$ random variables. Figure 3.2 shows the plot of a component flowrate series assuming that the component flowrate y_t , in tons per hour (t/h), equals the sum of x_t and 25. Figure 3.3 exhibits the autocorrelation function for the series.

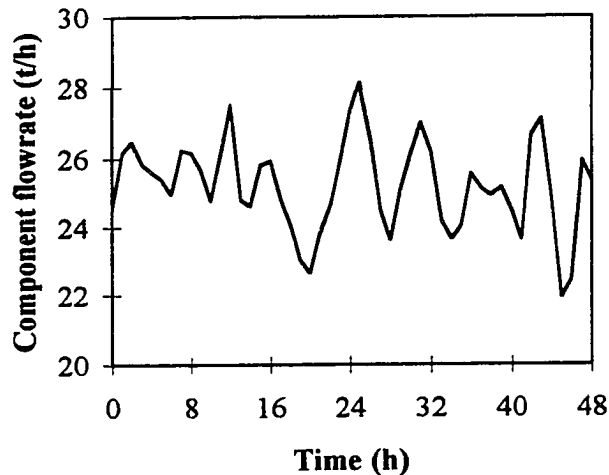


Figure 3.2. Component flowrate as a simulated $AR(2)$ series,

$$(1 - 0.7B + 0.5B^2)x_t = \xi_t, \text{ plus } 25 \text{ t/h.}$$

3.5.3. Moving average processes

A moving average process or model of order q , that is denoted as $MA(q)$, is given by (Box and Jenkins, 1976; Abraham and Ledolter, 1983; Wei, 1990):

$$x_t = \Theta_q(B)\xi_t \tag{3.21}$$

where Θ_q is the moving average polynomial of order q in B :

$$\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q \quad (3.22)$$

Consequently, x_t is defined by:

$$x_t = \xi_t - \theta_1 \xi_{t-1} - \dots - \theta_q \xi_{t-q} \quad (3.23)$$

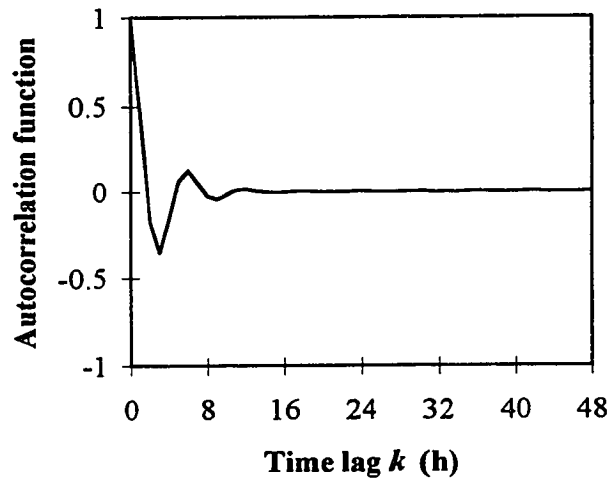


Figure 3.3. Autocorrelation function of the series $(1 - 0.7B + 0.5B^2)x_t = \xi_t$.

A finite moving average process is always stationary because $1 + \theta_1^2 + \dots + \theta_q^2 < \infty$. This moving average process is invertible if the roots of the polynomial $\Theta_q(B) = 0$ lie outside of the unit circle. Moving average processes are useful in describing phenomena in which events produce an immediate effect that only lasts for short periods of time (Wei, 1990).

The general q th order moving average process is:

$$x_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \xi_t \quad (3.24)$$

For this general $MA(q)$ process, the variance is:

$$\sigma_x^2 = \sigma_\xi^2 \sum_{j=0}^q \theta_j^2 = \gamma_0 \quad (3.25)$$

where $\theta_0 = 1$, and the other autocovariances are:

$$\gamma_k = \begin{cases} \sigma_\xi^2 (-\theta_k + \theta_1 \theta_{k-1} + \dots + \theta_{q-k} \theta_q), & k = 1, 2, \dots, q, \\ 0, & k > q. \end{cases} \quad (3.26)$$

Thus, the autocorrelation function becomes:

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k-1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & k = 1, 2, \dots, q \\ 0, & k > q \end{cases} \quad (3.27)$$

The autocorrelation function of an $MA(q)$ process cuts off after lag q . This important property enables the identification of a given time series that is generated by a moving average process.

For illustration, an $MA(2)$ process, $x_t = (1 - 0.7B + 0.5B^2) \xi_t$, is simulated (see Appendices A and H). The white noise series ξ_t are independent normal $N(0,1)$ random variables. Figure 3.4 shows the plot of a component flowrate series assuming that the component flowrate y_t , in t/h, equals the sum of x_t and 25. Figure 3.5 displays the autocorrelation function for the series.

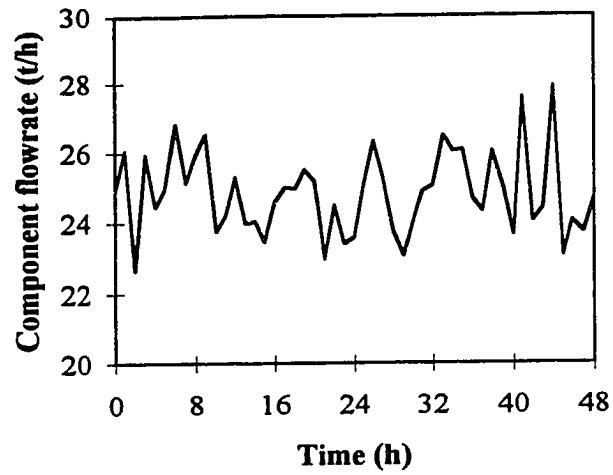


Figure 3.4. Component flowrate as a simulated $MA(2)$ series,

$$x_t = (1 - 0.7B + 0.5B^2)\xi_t, \text{ plus } 25 \text{ t/h.}$$

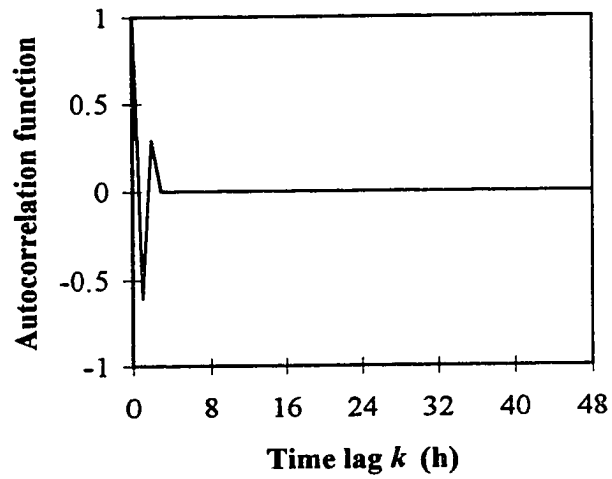


Figure 3.5. Autocorrelation function of the series $x_t = (1 - 0.7B + 0.5B^2)\xi_t$.

3.5.4. Dual relationship between $AR(p)$ and $MA(q)$ processes

For a stationary $AR(p)$ process given by Equation 3.17, the following can be written (Box and Jenkins, 1976; Abraham and Ledolter, 1983; Wei, 1990):

$$x_t = \frac{1}{\Phi_p(B)} \xi_t = \Psi(B) \xi_t \quad (3.28)$$

with

$$\Psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \quad (3.29)$$

such that

$$\Phi_p(B) \Psi(B) = 1 \quad (3.30)$$

The Ψ weights can be derived by equating the coefficients of B^j on both sides of (3.30).

Therefore, a finite order stationary AR process is equivalent to an infinite order MA process.

Now, given a general invertible $MA(q)$ process defined by Equation (3.24), it can be rewritten as (Box and Jenkins, 1976; Abraham and Ledolter, 1983; Wei, 1990):

$$\Pi(B)x_t = \frac{1}{\Theta_q(B)} x_t = \xi_t \quad (3.31)$$

where:

$$\Pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad (3.32)$$

In terms of the *AR* representation, a finite order invertible *MA* process is equivalent to an infinite order *AR* process.

3.5.5. Autoregressive moving average processes

In time series modeling, a useful class of parsimonious models is the autoregressive moving average *ARMA*(*p*,*q*) process x_t defined by (Box and Jenkins, 1976; Abraham and Ledolter, 1983; Wei, 1990):

$$\Phi_p(B)x_t = \Theta_q(B)\xi_t \quad (3.33)$$

$\Phi_p(B)x_t = \Theta_q(B)\xi_t$ will define a stationary process, provided that the characteristic equation $\Phi_p(B) = 0$ has all its roots lying outside the unit circle. Similarly, $\Theta_q(B) = 0$ has all its roots lying outside the unit circle if the process is to be invertible.

The stationary and invertible *ARMA* process can be written in a pure autoregressive (*AR*) representation:

$$\Pi(B)x_t = \xi_t \quad (3.34)$$

where

$$\Pi(B) = \frac{\Phi_p(B)}{\Theta_q(B)} = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad (3.35)$$

This process can also be written as a pure moving average (*MA*) representation:

$$x_t = \Psi(B)\xi_t \quad (3.36)$$

where:

$$\Psi(B) = \frac{\Theta_q(B)}{\Phi_p(B)} = 1 + \psi_1 B + \psi_2 B^2 + \dots \quad (3.37)$$

Therefore, the *ARMA* process x_t can be represented as the output from a linear filter, whose input is white noise ξ_t , that is:

$$x_t = \xi_t + \psi_1 \xi_{t-1} + \psi_2 \xi_{t-2} + \dots = \xi_t + \sum_{j=1}^{\infty} \psi_j \xi_{t-j} \quad (3.38)$$

or:

$$x_t = \sum_{j=0}^{\infty} \psi_j \xi_{t-j}, \text{ with } \psi_0 = 1 \quad (3.39)$$

For illustration, an *ARMA*(1,1) process, $(1 - 0.5B)x_t = (1 - 0.8B)\xi_t$, is simulated (see Appendices A and H). The white noise series ξ_t are independent normal $N(0,1)$ random variables. Figure 3.6 shows the plot of a component flowrate series supposing that the component flowrate y_t , in t/h, equals the sum of x_t and 25. Figure 3.7 displays the autocorrelation function for the series.

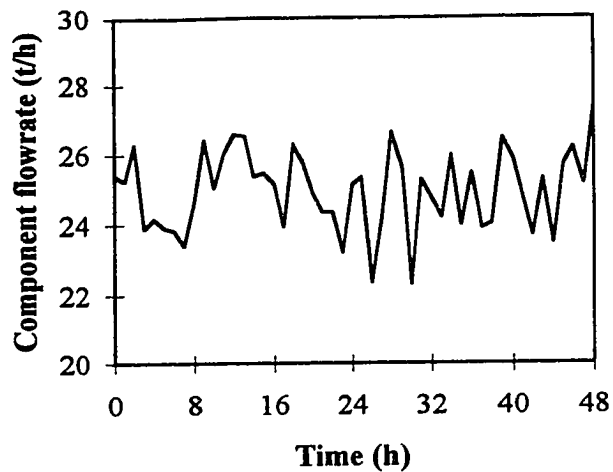


Figure 3.6. Component flowrate as a simulated $ARMA(1,1)$ series,

$$(1 - 0.5B)x_t = (1 - 0.8B)\xi_t, \text{ plus } 25 \text{ t/h.}$$

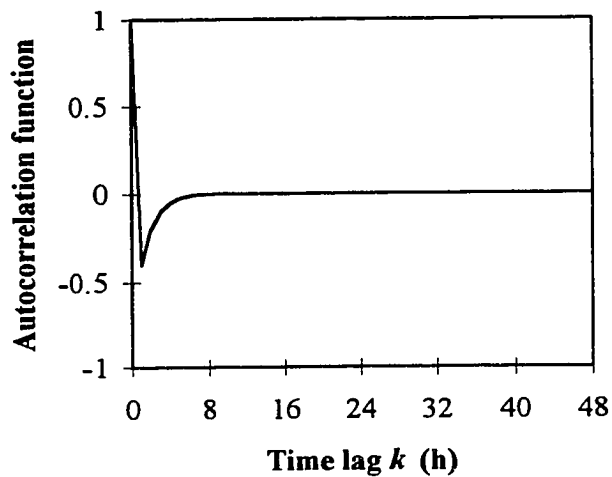


Figure 3.7. Autocorrelation function of the series $(1 - 0.5B)x_t = (1 - 0.8B)\xi_t$.

Note that an $ARMA(1,1)$ model, $(1 - \phi_1 B)x_t = (1 - \theta_1 B)\xi_t$, has the following autocovariance function:

$$\gamma_k = \begin{cases} \frac{(1 + \theta_1^2 - 2\phi_1\theta_1)}{(1 - \phi_1^2)} \sigma_\xi^2, & k = 0 \\ \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{(1 - \phi_1^2)} \sigma_\xi^2, & k = 1 \\ \phi_1 \gamma_{k-1}, & k \geq 2 \end{cases} \quad (3.40)$$

and the following autocorrelation function:

$$\rho_k = \begin{cases} 1, & k = 0 \\ \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{(1 + \theta_1^2 - 2\phi_1\theta_1)}, & k = 1 \\ \phi_1 \rho_{k-1}, & k \geq 2 \end{cases} \quad (3.41)$$

3.6. STATE SPACE MODELS

The state space representation of a system is a fundamental concept in modern control theory (Wei, 1990). The state of a system is defined to be a minimum set of information from the present and past such that the future behavior of the system can be completely described by the knowledge of the present state and the present input. In this way, the state space representation is based on the Markov property. Therefore, given the present state of a system, its future is independent of its past. The state space representation of a system is also called the Markovian representation of the system.

Let $X_{1,t}$ and $X_{2,t}$ be the outputs of a system to the inputs $U_{1,t}$ and $U_{2,t}$, respectively. A system is said to be linear if a linear combination of the inputs, $aU_{1,t} + bU_{2,t}$, produces the same linear combination of the outputs, $aX_{1,t} + bX_{2,t}$, for any constants a and b . A system is said to be time-invariant if the characteristics of a system do not change with time so that if the input U_t produces the output X_t , then the input U_{t-t_0} will produce the output X_{t-t_0} .

A system is linear time-invariant if it is both linear and time-invariant. The mineral processes can be modeled by linear time-invariant systems that contain stationary processes or time-series models that are discussed previously. For a linear time-invariant system, its state space representation is described by the state equation:

$$X_{1,t+1} = A_1 X_{1,t} + B_1 U_{1,t} \quad (3.42)$$

and the output equation:

$$Y_{1,t} = C_1 X_{1,t} \quad (3.43)$$

where $X_{1,t}$ is a state vector of dimension k , A_1 is a transition matrix of dimension $k \times k$, B_1 is an input matrix of dimension $k \times l$, $U_{1,t}$ is an input vector of dimension l , $Y_{1,t}$ is an output vector of dimension m , and C_1 is an output or observation matrix of dimension $m \times k$. Furthermore, k represents the number of streams in the mineral processing system, l the number of inputs to the system, and m the number of measured variables in the system. Here, the variables are component flowrates. Besides, $X_{1,t}$ is composed of the measured variables and the unmeasured variables, $Y_{1,t}$ is made up of the measured variables, and $U_{1,t}$ contains the input variables.

In this way, the stochastic process $Y_{1,t}$ is the output of a time-invariant linear system driven by the stochastic input $U_{1,t}$. The $X_{1,t}$ is known as the state of the process. The state equation is also called the system equation or the transition equation, and the output equation is also referred to as the measurement equation or the observation equation.

Equations 3.42 and 3.43 define a mineral processing system for only one selected component in the streams. Indeed, many components are present in the streams. Suppose that the number of components in each stream equals n . Consequently, the state space representation for the whole system taking into account all the components is described by the state equation:

$$X_{t+1} = AX_t + BU_t, \quad (3.44)$$

and the output equation:

$$Y_t = CX_t, \quad (3.45)$$

where X_t is a state vector of dimension kn :

$$X_t = [X_{1,t}, \dots, X_{n,t}]' \quad (3.46)$$

U_t is an input vector of dimension ln :

$$U_t = [U_{1,t}, \dots, U_{n,t}]' \quad (3.47)$$

Y_t is an output vector of dimension mn :

$$Y_t = [Y_{1,t}, \dots, Y_{n,t}]' \quad (3.48)$$

A is a transition matrix of dimension $kn \times kn$:

$$A = \begin{bmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_n \end{bmatrix} \quad (3.49)$$

B is an input matrix of dimension $kn \times ln$:

$$B = \begin{bmatrix} B_1 & 0 & 0 & \dots & 0 \\ 0 & B_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & B_n \end{bmatrix} \quad (3.50)$$

and C is an output or observation matrix of dimension $mn \times kn$:

$$C = \begin{bmatrix} C_1 & 0 & 0 & \dots & 0 \\ 0 & C_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_n \end{bmatrix} \quad (3.51)$$

3.7. STATE REPRESENTATION OF INPUT VARIABLES

The input variables are assumed to be first order autoregressive process $AR(1)$:

$$(1 - \phi B)u_t = \xi_t \quad (3.52)$$

As mentioned in Section 3.5.2, these processes are always invertible. To be stationary, the root of $(1 - \phi B) = 0$ must be outside of the unit circle. That is, for a stationary process, $|\phi| < 1$. The $AR(1)$ process is also called the Markov process because the value of x_{t+1} is completely determined by the knowledge of x_t .

For only one selected component in the streams, the state space representation of the $AR(1)$ model is given by:

$$U_{1,t+1} = F_1 U_{1,t} + G_1 \xi_{1,t+1} \quad (3.53)$$

where F_1 is a transition matrix of dimension $l \times l$, G_1 is an input matrix of dimension $l \times l$, and $\xi_{1,t+1}$ is a white noise input vector of dimension l .

Equation 3.53 defines a mineral processing system for only one selected component in the streams. Indeed, many components are present in the streams. Suppose that the number of components in each stream equals n . Consequently, the state space representation for the whole system input vector taking into account all the components is described by the state equation:

$$U_{t+1} = F U_t + G \xi_{t+1} \quad (3.54)$$

where F is a transition matrix of dimension $ln \times ln$:

$$F = \begin{bmatrix} F_1 & 0 & 0 & \dots & 0 \\ 0 & F_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & F_n \end{bmatrix} \quad (3.55)$$

G is an input matrix of dimension $ln \times ln$:

$$G = \begin{bmatrix} G_1 & 0 & 0 & \dots & 0 \\ 0 & G_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & G_n \end{bmatrix} \quad (3.56)$$

and ξ_{t+1} is a white noise vector of dimension ln :

$$\xi_{t+1} = [\xi_{1,t+1}, \dots, \xi_{n,t+1}]' \quad (3.57)$$

3.8. TRANSFER FROM COMPONENT FLOWRATES TO COMPOSITIONS

The composition of a component c in a stream j is given by:

$$a_j^c = \frac{x_j^c}{\sum_{l=1}^n x_j^l} \quad (3.58)$$

where x_j^c is the component flowrate of a component c in a stream j , and n the number of components in each stream.

Equation 3.58 shows that the composition a_j^c is not linear with respect to the component flowrates x_j^i , $i = 1, \dots, n$. To apply linear statistics, the composition a_j^c must be linearized around its true value as follows:

$$a_j^c = a_j^{c*} + \sum_{i=1}^n \left(\frac{\partial a_j^c}{\partial x_j^i} \right)_{x_j^*} (x_j^i - x_j^{i*}) \quad (3.59)$$

If $i \neq c$, $\frac{\partial a_j^c}{\partial x_j^i}$ equals:

$$\frac{\partial a_j^c}{\partial x_j^i} = \frac{(x_j^1 + \dots + x_j^c + \dots + x_j^n) \times 0 - x_j^c \times 1}{(x_j^1 + \dots + x_j^c + \dots + x_j^n)^2} \quad (3.60)$$

which is:

$$\frac{\partial a_j^c}{\partial x_j^i} = \frac{-x_j^c}{(x_j^1 + \dots + x_j^c + \dots + x_j^n)^2} \quad (3.61)$$

Otherwise, $\frac{\partial a_j^c}{\partial x_j^i}$ equals:

$$\frac{\partial a_j^c}{\partial x_j^i} = \frac{(x_j^1 + \dots + x_j^c + \dots + x_j^n) \times 1 - x_j^c \times 1}{(x_j^1 + \dots + x_j^c + \dots + x_j^n)^2} \quad (3.62)$$

that is:

$$\frac{\partial a_j^c}{\partial x_j^i} = \frac{(x_j^1 + \dots + x_j^c + \dots + x_j^n) - x_j^c}{(x_j^1 + \dots + x_j^c + \dots + x_j^n)^2} \quad (3.63)$$

Consequently, the following state equation is obtained:

$$Z_t = HX_t \quad (3.64)$$

where Z_t is a composition state vector of dimension kn :

$$Z_t = [Z_{1,t}, \dots, Z_{n,t}]' \quad (3.65)$$

and H is the transfer matrix from component flowrates to compositions:

$$H = \frac{1}{\left(\sum_{i=1}^n x_j^{i*}\right)^2} \begin{bmatrix} H_1^1 & H_1^2 & \dots & H_1^n \\ H_2^1 & H_2^2 & \dots & H_2^n \\ \vdots & \vdots & \ddots & \vdots \\ H_n^1 & H_n^2 & \dots & H_n^n \end{bmatrix} \quad (3.66)$$

Here, the following diagonal matrices of dimension $k \times k$ are defined:

$$H_1^1 = \left(\sum_{i=1}^n x_j^{i*} - x_j^{1*}\right) I_k \quad (3.67)$$

$$H_1^2 = H_3^2 = \dots = H_n^2 = -x_j^{2*} I_k \quad (3.68)$$

$$H_2^1 = H_3^1 = \dots = H_n^1 = -x_j^{1*} I_k \quad (3.69)$$

$$H_1^n = H_2^n = \dots = H_{n-1}^n = -x_j^{n*} I_k \quad (3.70)$$

and:

$$H_n^n = \left(\sum_{i=1}^n x_j^{i*} - x_j^{n*} \right) I_k \quad (3.71)$$

I_k is a $k \times k$ identity matrix.

3.9. VARIANCE MATRICES

Usually, the variance matrices are supposed to be diagonal. This means that the correlations between the component flowrates in the whole mineral processing system are neglected. Nevertheless, the contribution of the correlation between the streams component flowrates should be taken into account. Therefore, the variance matrices should not be diagonal.

3.9.1. Variance matrix of input vector

Using Equation 3.54, the variance matrix of the input vector is:

$$V_U = E \left[(FU_{t-1} + G\xi_t)(FU_{t-1} + G\xi_t)' \right] \quad (3.72)$$

that is:

$$V_U = FE \left[U_{t-1} U_{t-1}' \right] F' + GE \left[\xi_t \xi_t' \right] G' + FE \left[U_{t-1} \xi_t' \right] G' + GE \left[\xi_t U_{t-1}' \right] F' \quad (3.73)$$

However, the input vector variables and the white noise vector variables are not correlated. Therefore, the following expression is obtained:

$$V_U = FV_U F' + G V_\xi G' \quad (3.74)$$

where:

$$V_\xi = E \left[\xi_t \xi_t' \right] \quad (3.75)$$

3.9.2. Variance matrix of state vector

From the expression developed for X_t (Equation 3.44), the variance matrix V_X of the state vector X_t can be calculated by:

$$V_X = E \left[X_t X_t' \right] \quad (3.76)$$

where E stands for the mathematical expectation. Thus:

$$V_X = E \left[(AX_t + BU_t)(AX_t + BU_t)' \right] \quad (3.77)$$

which becomes:

$$V_X = E \left[AX_t X_t' A' + AX_t U_t' B' + BU_t X_t' A' + BU_t U_t' B' \right] \quad (3.78)$$

Finally, the variance matrix V_X equals:

$$V_X = AV_X A' + AV_{XU} B' + BV_{UX} A' + BV_U B' \quad (3.79)$$

where:

$$V_{XU} = E\left[X_t U_t'\right] \quad (3.80)$$

$$V_{UX} = E\left[U_t X_t'\right] \quad (3.81)$$

and:

$$V_U = E\left[U_T U_T'\right] \quad (3.82)$$

3.9.3. Variance matrix of output vector

From the output equation (Equation 3.45), it follows that:

$$V_Y = C V_X C' \quad (3.83)$$

3.9.4. Variance matrix of composition vector

Using Equation 3.64, the variance matrix of the composition vector Z_t is obtained:

$$V_Z = H V_X H' \quad (3.84)$$

3.10. COVARIANCE MATRICES

Similar to the variance matrices, the covariance matrices are supposed to be diagonal. This means that the correlations between the component flowrates in the whole mineral

processing system are neglected. Nevertheless, the contribution of the correlation between the streams component flowrates should be taken into account. Therefore, the covariance matrices should not be diagonal.

3.10.1. Covariance matrix for input vector

The covariance matrix of the input vector U_t is defined by:

$$V_{UU}(k) = E[U_t U_{t+k}'] \quad (3.85)$$

Using Equation 3.54, the following is established:

$$U_{t+2} = F^2 U_t + FG\xi_{t+1} + G\xi_{t+2} \quad (3.86)$$

and:

$$U_{t+3} = F^3 U_t + F^2 G\xi_{t+1} + FG\xi_{t+2} + G\xi_{t+3} \quad (3.87)$$

Thus, the resulting recurrence formula is obtained:

$$U_{t+k} = F^k U_t + \sum_{i=1}^{k-1} F^{k-i} G\xi_{t+i} + G\xi_{t+k} \quad (3.88)$$

Hence, the covariance matrix $V_{UU}(k)$ equals:

$$V_{UU}(k) = E \left[U_t \left(F^k U_t + \sum_{i=1}^{k-1} F^{k-i} G \xi_{t+i} + G \xi_{t+k} \right)' \right] \quad (3.89)$$

which becomes:

$$V_{UU}(k) = V_U F'^k + \sum_{i=1}^{k-1} E \left[U_t \xi_{t+i}' \right] G' F'^{k-i} + E \left[U_t \xi_{t+k}' \right] G' \quad (3.90)$$

However, U_t and ξ_{t+i} , for $i \geq 1$, are not correlated. As a result, the covariance matrix $V_{UU}(k)$ is:

$$V_{UU}(k) = V_U F'^k \quad (3.91)$$

where $V_{UU}(k)$ is not necessarily symmetric. Its transpose equals:

$$V_{UU}'(k) = E \left[U_{t+k} U_t' \right] = V_{UU}(-k) \quad (3.92)$$

3.10.2. Covariance matrix for state vector

The covariance matrix of the state vector X_t is given by:

$$V_{XX}(k) = E \left[X_t X_{t+k}' \right] \quad (3.93)$$

Now, from the expression obtained for X_t (Equation 3.44), X_{t+2} is:

$$X_{t+2} = AX_{t+1} + BU_{t+1} \quad (3.94)$$

Thus:

$$X_{t+2} = A^2 X_t + ABU_t + BU_{t+1} \quad (3.95)$$

and:

$$X_{t+3} = AX_{t+2} + BU_{t+2} \quad (3.96)$$

that is:

$$X_{t+3} = A^3 X_t + A^2 BU_t + ABU_{t+1} + BU_{t+2} \quad (3.97)$$

Consequently, the following recurrence formula is obtained:

$$X_{t+k} = A^k X_t + \sum_{i=1}^{k-1} A^{k-i} BU_{t+i-1} + BU_{t+k-1} \quad (3.98)$$

Therefore, the covariance matrix of the state vector X_t is:

$$V_{XX}(k) = E \left[X_t \left(A^k X_t + \sum_{i=1}^{k-1} A^{k-i} BU_{t+i-1} + BU_{t+k-1} \right)' \right] \quad (3.99)$$

which is expressed as follows:

$$V_{XX}(k) = V_X A'^k + \sum_{i=1}^{k-1} V_{XU}(i-1) B' A'^{k-i} + V_{XU}(k-1) B' \quad (3.100)$$

where $V_{xx}(k)$ is not necessarily symmetric. Its transpose equals:

$$V_{xx}'(k) = E[X_{t+k}X_t'] = V_{xx}(-k) \quad (3.101)$$

3.10.3. Covariance matrix for output vector

The covariance matrix of the output vector Y_t is defined by:

$$V_{yy}(k) = E[Y_t Y_{t+k}'] \quad (3.102)$$

which is equal to:

$$V_{yy}(k) = CE[X_t X_{t+k}']C' \quad (3.103)$$

Finally, the covariance matrix $V_{yy}(k)$ is expressed as:

$$V_{yy}(k) = CV_{xx}(k)C' \quad (3.104)$$

where $V_{yy}(k)$ is not necessarily symmetric. Its transpose equals:

$$V_{yy}'(k) = E[Y_{t+k}Y_t'] = V_{yy}(-k) \quad (3.105)$$

3.10.4. Covariance matrix for composition vector

The covariance matrix of the composition vector Z_t is defined by:

$$V_{zz}(k) = E\left[Z_t Z_{t+k}'\right] = HV_{xx}(k)H' \quad (3.106)$$

where $V_{zz}(k)$ is not necessarily symmetric. Its transpose equals:

$$V_{zz}'(k) = E\left[Z_{t+k} Z_t'\right] = V_{zz}(-k) \quad (3.107)$$

3.10.5. Cross-covariance matrix between input and state vectors

The covariance matrix $V_{ux}(k)$ is defined as follows:

$$V_{ux}(k) = E\left[U_t X_{t+k}'\right] \quad (3.108)$$

that is:

$$V_{ux}(k) = E\left[U_t \left(A^k X_t + \sum_{i=1}^{k-1} A^{k-i} B U_{t+i-1} + B U_{t+k-1}\right)'\right] \quad (3.109)$$

and is written:

$$V_{ux}(k) = V_{ux}(0)A'^k + \sum_{i=1}^{k-1} V_{uu}(i-1)B'A'^{k-i} + V_{uu}(k-1)B' \quad (3.110)$$

Using Equation 3.91, Equation 3.110 becomes:

$$V_{UX}(k) = V_{UX}(0)A'^k + \sum_{i=1}^{k-1} V_U F'^{i-1} B' A'^{k-i} + V_U F'^{k-1} B' \quad (3.111)$$

which is equivalent to:

$$V_{UX}(k) = V_{UX}(0)A'^k + V_U \left(\sum_{i=1}^{k-1} F'^{i-1} B' A'^{k-i} + F'^{k-1} B' \right) \quad (3.112)$$

where $V_{UX}(k)$ is not necessarily symmetric. Its transpose equals:

$$V_{UX}'(k) = E \left[U_{t+k} X_t' \right] = V_{UX}(-k) \quad (3.113)$$

3.10.6. Cross-covariance matrix between state and input vectors

The covariance matrix $V_{XU}(k)$ is:

$$V_{XU}(k) = E \left[X_t U_{t+k}' \right] \quad (3.114)$$

which equals:

$$V_{XU}(k) = E \left[X_t \left(F^k U_t + \sum_{i=1}^{k-1} F^{k-i} G \xi_{t+i} + G \xi_{t+k} \right)' \right] \quad (3.115)$$

and is written:

$$V_{XU}(k) = V_{XU}(0)F'^k \quad (3.116)$$

where $V_{xu}(k)$ is not necessarily symmetric. Its transpose equals:

$$V_{xu}'(k) = E[X_{t+k}U_t'] = V_{xu}(-k) \quad (3.117)$$

Thus, Equation 3.100 becomes:

$$V_{xx}(k) = V_x A'^k + \sum_{i=1}^{k-1} V_{xu}(0) F'^{i-1} B' A'^{k-i} + V_{xu}(0) F'^{k-1} \quad (3.118)$$

CHAPTER 4
STOCHASTIC EVALUATION OF SAMPLING ERROR

4.1. SAMPLING PROCEDURE

Gy (1979, 1988, 1992) defines three main strategies for incremental sampling:

- Random sampling: the n increments are extracted at n uniformly distributed random times over the interval T .
- Systematic sampling: the increments are extracted from the stream every $t_s = T/n$ time units.
- Stratified random sampling: the increments are randomly extracted within the n time intervals t_s .

Only the systematic sampling strategy, which is the usual one, is studied here.

The typical incremental sampling situation that is analyzed here is the one of a freely flowing stream containing size-distributed particles in a slurry as shown in Figure 2.2. It is assumed that the sample increments are extracted by a constant-speed linearly moving cutter that crosses the stream. Let w be the cutter width and L the stream width (see Figure 2.2). Consequently, the sampler is moving at a speed:

$$v = (L + 2w) / t_p \tag{4.1}$$

where t_p is the sampler transfer time from one side of the stream to the other one.

The sampling correctness rules as described by Gy are assumed to be perfectly respected. A sample is correct when all the particles have the same probability of falling into the

cutter. Assuming sampling correctness is equivalent to neglecting the errors of long-range and periodic heterogeneity fluctuations, increment delimitation, extraction, and preparation as defined by Gy.

4.2. SELECTION OF STREAM VARIABLES

A stream sample is collected for the determination of the average stream composition. Usually, the composition is defined, for a set of stream components, as the mass fractions of these components with respect to a reference phase. The latter could be either the overall stream material or a part of it. For example, the following composition definitions could be used:

- Mass fraction of solids in the slurry.
- Mass fraction of one metal in the solid phase.
- Mass fraction of one mineral in a given particle size interval.

If only one selected component of the stream is analyzed, the mass fraction of that component in the sample is:

$$a_s = \frac{\sum_{i=1}^n t(i)m(i)a(i)}{\sum_{i=1}^n t(i)m(i)} \quad (4.2)$$

where:

- i is the index of the sample increment.

- $m(i)$ is the average mass flowrate of the reference phase at the time of the extraction of the i th sample increment.
- $a(i)$ is the average mass fraction of the component of interest at the time of the extraction of the i th sample increment.
- $t(i)$ is the time during which each part of the stream is sampled.

$t(i)$ is given by:

$$t(i) = w / v \quad (4.3)$$

As the transfer time t_p of the sampler has been assumed to be constant, $t(i)$ is independent of i and as a result:

$$a_s = \frac{\sum_{i=1}^n m(i)a(i)}{\sum_{i=1}^n m(i)} = \frac{1}{n\bar{m}} \sum_{i=1}^n m(i)a(i) \quad (4.4)$$

where \bar{m} is the average flowrate of the reference phase over the n sample increments.

Therefore, the measured property of the sample is obtained by an averaging process of the flowrate ma of the component rather than by an averaging process of the mass fraction a of the component. For that reason the stream is characterized in the following development by the component flowrate: $x(y,t)$.

The variable $x(y,t)$ has a concrete sense and its average value along the stream width can eventually be measured. For instance, using a magnetic flowmeter and a gamma gauge densimeter, and knowing the solid density, the ore flowrate can be continuously

monitored. The ore flowrate is the product of the mass flowrate of the slurry with the solid mass fraction.

The variable $x(y,t)$ can be understood as a time variable distributed along the y axis of the stream width (Figure 2.3). Because of the type of sampling device, the heterogeneity along the z -axis (Figure 2.2) does not need to be described. The question of the continuity of x can be discussed because of the discrete nature of the particulate material flowing through the section S (Figure 2.2). Both discrete and continuous approaches are correct, although the continuous formalism is easier to manipulate. However, to calculate the sampling error the signal x must be discretized since the increment extraction from the stream is typically a discretization process. (Hodouin and Ketata, 1994)

If the mass flowrate $m(i)$ is constant, then the measured property of the sample is obtained by an averaging process of the composition α . This case is studied in Appendix B.

In the following sections it is assumed that the variables $m(i)$, $\alpha(i)$, and $x(i)$ obey normal distributions.

4.3. DISCRETIZATION OF THE COMPONENT FLOWRATE

The discretization operator that is used is a mathematical transformation that represents the sampling of a stream by a moving cutter as previously described. The time is discretized using a time period t_d such that $t_d \geq t_p$. Within each time interval $[(i-1)t_d, it_d]$ of index i , the continuous function $x(y,t)$ is replaced by the value:

$$x(i) = \frac{v}{w} \int_{y=0}^L \int_{t=\tau}^{\tau'} x(y,t) dt dy \quad (4.5)$$

with:

$$\tau = t_0 + y/v + (i-1)t_d \quad (4.6)$$

and:

$$\tau' = \tau + w/v \quad (4.7)$$

where t_0 is the elapsed time since the beginning of the interval when the cutter is started. The discretization operator is depicted in Figure 4.1. This discretizer has a double function. First it is a filter that smoothes the continuous function $x(y,t)$ in both y and t directions through the double integral. Second it is a discretizer in the time direction.

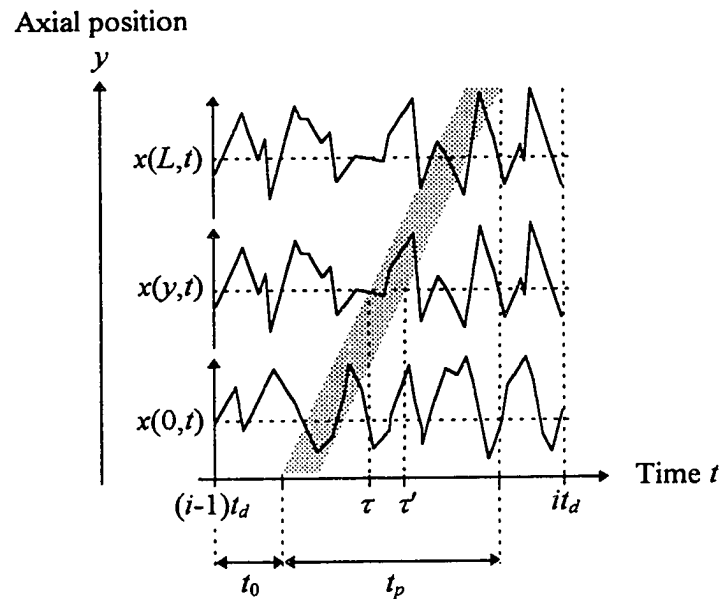


Figure 4.1. Discretization operator.

It must be emphasized that the physical discretization operation by a moving cutter converges towards the mathematical operator of Equation 4.5 only when the number of particles entering the samples tends to infinity. The difference between the two operators is due to the particulate structure of the flowing material. The mathematical integration is not perfectly performed by the sample increment since particles cannot be cut by the edges of the flow cutter. Statistically this difference is vanishing when the number of particles considered in the integral is large.

4.4. DISCRETIZATION ERROR

The smoothing or discretization procedure necessarily produces a loss of information, which is characterized here by the discretization error. This error is the difference between the discretized value $x(i)$ and the true mean value $x^*(i)$ within the time interval $[(i-1)t_d, it_d]$. It can be evaluated by:

$$e_d(i) = x(i) - x^*(i) \quad (4.8)$$

with:

$$x^*(i) = \frac{1}{t_d} \int_0^{t_d} \int_{(i-1)t_d}^{it_d} x(y, t) dt dy \quad (4.9)$$

It is assumed in the following that x is a randomly distributed value, so that e_d is also a random variable. Due to the assumption of sampling correctness its mean value is zero. Its variance σ_d^2 depends on the local variance of x , that is, the variance in the frequency range

$\left[\frac{2\pi}{t_d}, \frac{2\pi\nu}{w} \right]$. The larger it is, the larger is the discretization error variance. It also depends on the ratio t_p/t_d : the closer to one, the smaller is the discretization error variance.

Although a continuous formalism for x is used, which embeds the constitution and distribution heterogeneities, the discretization error variance includes both the fundamental and the grouping and segregation errors defined by Gy. When the number of particles entering the cutter is too small, the difference between the mathematical and the physical discretizers is an additional contribution to the variance of e_d .

Finally it is assumed that the discretization error has a constant variance σ_d^2 and is uncorrelated from discretization interval to discretization interval. These properties are quite reasonable, since they are based on the assumption that the variable x is locally stationary, that is, stationary within a window of width t_d .

When the mean value of x is not constant within a t_d interval, the function x can be replaced by its variation around its moving average. This will filter out the frequencies lower than $2\pi/t_d$. The variance σ_d^2 is the variance of this variation. This is a local variance, or in other words, the x variance for frequencies higher than $2\pi/t_d$. These assumptions mean that the function x can be viewed as the sum of two functions:

$$x(t) = \mu(t) + \varepsilon(t) \quad (4.10)$$

$\mu(t)$ contains the low frequencies of $x(t)$ and $\varepsilon(t)$ the high frequencies. The time-discrete version of this equation is:

$$x(i) = x^*(i) + e_d(i) \quad (4.11)$$

with

$$E(e_d) = 0 \quad (4.12)$$

$$E(e_d^2) = \sigma_d^2 \quad (4.13)$$

$$E[e_d(i)e_d(i+j)] = 0, \quad \forall j \neq i \quad (4.14)$$

and:

$$E[x^*(i)e_d(i+j)] = 0, \quad \forall i, j \quad (4.15)$$

This decomposition can be graphically understood as depicted in Figure 4.2.

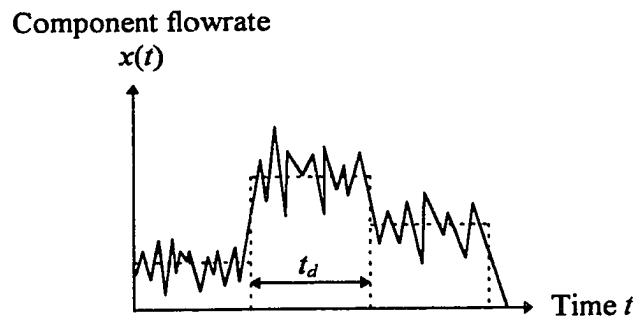


Figure 4.2. Decomposition of $x(t)$ into a sum of a moving average and a stationary noise.

4.5. SAMPLING ERROR

Equation 4.11 shows that the component flowrate x is the sum of two time-discrete functions where $x^*(i)$ is the real mean value of the flowrate in the time interval $[(i-1)t_d, it_d]$ and $e_d(i)$ the discretization error considered as a white sequence. Now it is assumed that $x^*(i)$ is a stationary random signal with a mean value:

$$E[x^*(i)] = E[x(i)] = x^* = m^* \alpha^* \quad (4.16)$$

and a variance:

$$E[(x^*(i) - x^*)^2] = \sigma_*^2 \quad (4.17)$$

As a result the variance of $x(i)$ is:

$$\sigma^2 = \sigma_*^2 + \sigma_d^2 \quad (4.18)$$

The autocovariance of $x^*(i)$ is:

$$\sigma_*^2 \rho^*(k) = E[(x^*(i) - x^*)(x^*(i+k) - x^*)] \quad (4.19)$$

where ρ^* is the autocorrelation function of $x^*(i)$. Consequently, the autocovariance of $x(i)$ is:

$$\sigma^2 \rho(k) = E[(x(i) - x^*)(x(i+k) - x^*)] = \sigma_*^2 \rho^*(k) \quad (4.20)$$

where ρ is the autocorrelation function of $x(i)$. The sampling error e (see Figure 4.3) is the difference between the component mass fraction a_S in the composite sample and the component mass fraction a_T in the total amount of reference phase. The latter flows through the S section (Figure 2.2) during the time interval T .

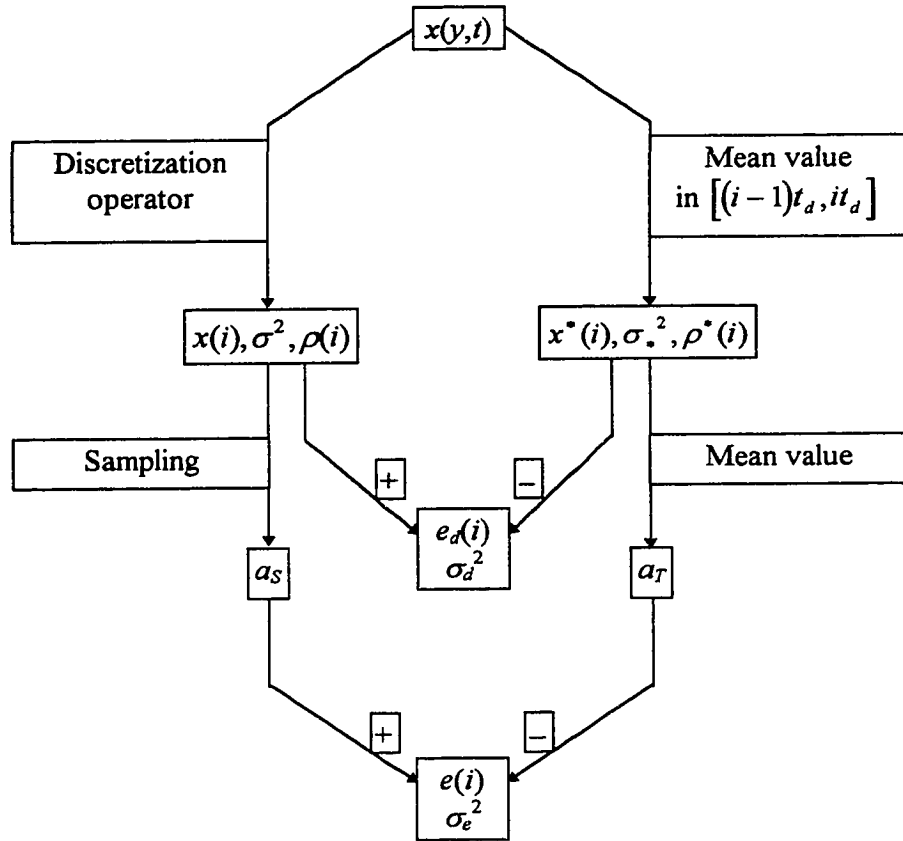


Figure 4.3. Scheme of the process leading to the sampling error variance.

The expression for a_S is obtained from Equation 4.4:

$$a_s = \frac{1}{mn} \sum_{i=0}^{n-1} x(l+ik) = \frac{1}{mn} \sum_{i=0}^{n-1} x^*(l+ik) + \frac{1}{mn} \sum_{i=0}^{n-1} e_d(l+ik) \quad (4.21)$$

where:

- n is the number of increments in the sample.
- k is the number of discretization intervals t_d in the sampling period t_s .

k is an integer given by:

$$k = \frac{t_s}{t_d} \quad (4.22)$$

l is the time index of the first sample increment. For even values of k , l equals:

$$l = \frac{k}{2} \quad (4.23)$$

and for odd values of k , it is:

$$l = \frac{k+1}{2} \quad (4.24)$$

The true value a_T of the mass fraction in the time interval T , that is, from $i = 1$ to nk , is:

$$a_T = \frac{1}{m_T nk} \sum_{i=1}^{nk} x^*(i) \quad (4.25)$$

and by definition the sampling error is:

$$e = a_s - a_T = SE \quad (4.26)$$

which can also be written:

$$e = \frac{1}{n\bar{m}} \sum_{i=1}^n m(i) [a(i) - a_T] \quad (4.27)$$

Appendix C examines the case of variable \bar{m} and m_T .

Because of the sampling correctness assumption, the sampling error has a zero mean value:

$$E(e) = 0 \quad (4.28)$$

and a variance equal to:

$$\sigma_e^2 = E(e^2) = E[(a_s - a^*)^2] + E[(a_T - a^*)^2] - 2E[(a_s - a^*)(a_T - a^*)] \quad (4.29)$$

4.6. SAMPLING ERROR VARIANCE

The sampling error is given by: $e = a_s - a_T$, where a_s and a_T are:

$$a_s = \frac{1}{\bar{m}n} \left[\sum_{i=0}^{n-1} x^*(l+ik) + \sum_{i=0}^{n-1} e_d(l+ik) \right] \quad (4.30)$$

$$a_T = \frac{1}{m_T n k} \sum_{i=1}^{nk} x^*(i) \quad (4.31)$$

The variance of e is then:

$$\begin{aligned} \sigma_e^2 = & \frac{1}{(\bar{m}n)^2} \text{var} \left[\sum_{i=0}^{n-1} x^*(l+ik) \right] + \frac{1}{(\bar{m}n)^2} n \sigma_d^2 + \frac{1}{(m_T nk)^2} \text{var} \left[\sum_{i=1}^{nk} x^*(i) \right] \\ & - \frac{2}{\bar{m} m_T n^2 k} \text{cov} \left[\sum_{j=0}^{n-1} x^*(l+jk); \sum_{i=1}^{nk} x^*(i) \right] \end{aligned} \quad (4.32)$$

where $\text{var}[\quad]$ and $\text{cov}[\quad]$ stand for the variance and covariance of the term within the brackets. Each term of Equation 4.32 can now be evaluated:

$$\text{var} \left[\sum_{i=0}^{n-1} x^*(l+ik) \right] = E \left[\left(\sum_{i=0}^{n-1} (x^*(l+ik) - x^*) \right)^2 \right] = \sigma_s^2 \left[n + 2 \sum_{i=1}^{n-1} (n-i) \rho^*(ik) \right] \quad (4.33)$$

$$\text{var} \left[\sum_{i=1}^{nk} x^*(i) \right] = E \left[\left(\sum_{i=1}^{nk} (x^*(i) - x^*) \right)^2 \right] = \sigma_s^2 \left[nk + 2 \sum_{i=1}^{nk-1} (nk-i) \rho^*(i) \right] \quad (4.34)$$

and:

$$\begin{aligned} \text{cov} \left[\sum_{j=0}^{n-1} x^*(l+jk); \sum_{i=1}^{nk} x^*(i) \right] &= E \left[\left(\sum_{j=0}^{n-1} (x^*(l+jk) - x^*) \right) \left(\sum_{i=1}^{nk} (x^*(i) - x^*) \right) \right] \\ &= \sigma_s^2 \left[\sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho^*(l+jk-i) \right] \end{aligned} \quad (4.35)$$

Using Equations 4.32, 4.33, 4.34, and 4.35, the sampling error variance becomes:

$$\sigma_e^2 = \frac{\sigma_d^2}{\bar{m}^2 n} + \left(\frac{1}{\bar{m}^2} + \frac{1}{m_T^2 k} \right) \frac{\sigma_s^2}{n} + \frac{2\sigma_s^2}{n^2} \left[\frac{1}{\bar{m}^2} \sum_{i=1}^{n-1} (n-i) \rho^*(ik) + \frac{1}{m_T^2 k^2} \sum_{i=1}^{nk-1} (nk-i) \rho^*(i) - \frac{1}{\bar{m} m_T k} \sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho^*(l+jk-i) \right] \quad (4.36)$$

Using Equations 4.18 and 4.20, Equation 4.36 can also be written:

$$\sigma_e^2 = \frac{-\sigma_d^2}{m_T^2 nk} + \frac{(m_T^2 k + \bar{m}^2) \sigma^2}{m_T^2 \bar{m}^2 nk} + \frac{2\sigma^2}{n^2} \left[\frac{1}{\bar{m}^2} \sum_{i=1}^{n-1} (n-i) \rho(ik) + \frac{1}{m_T^2 k^2} \sum_{i=1}^{nk-1} (nk-i) \rho(i) - \frac{1}{\bar{m} m_T k} \sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho(l+jk-i) \right] \quad (4.37)$$

It is impossible to measure the difference between \bar{m} and m_T . Thus, they are assumed to have a common estimate m :

$$\bar{m} = m_T = m \quad (4.38)$$

which finally leads to the following expression of the variance of the sampling error:

$$\sigma_e^2(n, k) = \frac{1}{nkm^2} \left((k+1) \sigma^2 - \sigma_d^2 \right) + \frac{2\sigma^2}{n^2 m^2} \left[\sum_{i=1}^{n-1} (n-i) \rho(ik) + \frac{1}{k^2} \sum_{i=1}^{nk-1} (nk-i) \rho(i) - \frac{1}{k} \sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho(l+jk-i) \right] \quad (4.39)$$

σ_e^2 can be expressed as a function of σ_d^2 , σ_s^2 , and ρ^* (see Equation 4.36). However, in practice only σ_d^2 , σ^2 , and ρ are experimentally accessible so that Equation 4.39 is more convenient.

The general formula given by Equation 4.39 has particular applications. First, in the case where there is no autocorrelation of the signal $x(i)$, the sampling error variance becomes:

$$\sigma_e^2 = \frac{1}{nkm^2} \left((k+1)\sigma^2 - \sigma_d^2 \right) \quad (4.40)$$

or:

$$\sigma_e^2 = \frac{1}{nkm^2} \left(k\sigma^2 + \sigma_e^2 \right) \quad (4.41)$$

The second case is when only one sample increment is extracted ($n = 1$). The sampling error variance is then:

$$\sigma_e^2(1, k) = \frac{-\sigma_d^2}{m^2k} + \frac{\sigma^2}{m^2} \left[1 + \frac{1}{k} - \frac{2}{k} \sum_{i=1}^k \rho(l-i) + \frac{2}{k^2} \sum_{i=1}^{k-1} (k-i)\rho(i) \right] \quad (4.42)$$

Finally, when all the possible increments are extracted, $t_s = t_d$. This means that $k = 1$, and $l = 1$. In addition, n tends to infinity. Consequently, the sampling error is zero, which is a logical result since the sample is the lot itself.

4.7. CORRELATION BETWEEN SAMPLE INCREMENTS

The assumption that:

$$\bar{m} = m_T = m$$

is initially made before calculating of the correlation between the sample increments. Equations 4.26, 4.30, and 4.31 can be rearranged as follows:

$$e = \frac{1}{nm} \sum_{j=0}^{n-1} \left[x^*(l+jk) + e_d(l+jk) - \frac{1}{k} \sum_{i=1}^k x^*(jk+i) \right] \quad (4.43)$$

or, in a more compact form:

$$e = \frac{1}{nm} \sum_{j=0}^{n-1} e_1(j) \quad (4.44)$$

where $e_1(j)$ is the x error associated with the j th increment (incremental error), that is, the x sampling error when taking a single increment ($n = 1$) in a time interval containing k discrete values of the stream property:

$$e_1(j) = x(l+jk) - \frac{1}{k} \sum_{i=1}^k x^*(jk+i) \quad (4.45)$$

Now the variance of e is:

$$\sigma_e^2(n, k) = \frac{1}{n} \sigma_e^2(1, k) + \frac{2}{n^2 m^2} \sum_{i=1}^{n-1} (n-i) \sigma_1(i) \quad (4.46)$$

where:

- $\sigma_e^2(1, k)$ is given by Equation 4.42.
- $\sigma_1(i)$ is the covariance between the x errors of increments separated by ik time intervals t_d .

The covariance term $\sigma_1(i)$ is then:

$$\sigma_1(i) = E[e_1(j)e_1(j+i)] \quad (4.47)$$

which is independent of j because of the stationarity of the series. It equals:

$$\sigma_1(i) = E\left[\left(x(l+jk) - \frac{1}{k} \sum_{p=1}^k x^*(jk+p)\right)\left(x(l+(j+i)k) - \frac{1}{k} \sum_{q=1}^k x^*((j+i)k+q)\right)\right] \quad (4.48)$$

Hence, the covariance term is:

$$\begin{aligned} \sigma_1(i) = \sigma^2 & \left[\rho(ki) - \frac{1}{k} \sum_{j=1}^k \rho(ki+j-l) - \frac{1}{k} \sum_{j=1}^k \rho(ki-j+l) \right. \\ & \left. + \frac{1}{k^2} \sum_{j=1}^k j \rho((i+1)k-j) + \frac{1}{k^2} \sum_{j=1}^{k-1} j \rho((i-1)k+j) \right] \end{aligned} \quad (4.49)$$

Finally, from Equation 4.46, the covariance term between the incremental errors is obtained:

$$\begin{aligned} \sigma_e^2(n,k) - \frac{1}{n} \sigma_e^2(1,k) = \frac{2\sigma^2}{n^2 m^2} & \left[\sum_{i=1}^{n-1} (n-i) \rho(ki) \right. \\ & \left. - \frac{1}{k} \left(\sum_{i=1}^{n-1} \sum_{j=1}^k (n-i) \rho(ki+j-l) + \sum_{i=1}^{n-1} \sum_{j=1}^k (n-i) \rho(ki-j+l) \right) \right. \\ & \left. + \frac{1}{k^2} \left(\sum_{i=1}^{n-1} \sum_{j=1}^k j(n-i) \rho((i+1)k-j) + \sum_{i=1}^{n-1} \sum_{j=1}^{k-1} j(n-i) \rho((i-1)k+j) \right) \right] \end{aligned} \quad (4.50)$$

This term represents the contribution of the correlation between the incremental errors. The same result could have been obtained by directly calculating the difference $\sigma_e^2(n, k) - (1/n)\sigma_e^2(1, k)$ using Equations 4.39 and 4.42.

4.8. COMPARISON WITH GY'S APPROACH

Conceptually the approach above is very close to Gy's analysis of sampling errors. Here, however, signal processing tools are used which give a different light to Gy's theory.

Gy's decomposition of the sampling error into two contributions (the continuous selection and the materialization errors) is followed here. Using the sampling correctness assumption, the long-range heterogeneity fluctuation, the periodic heterogeneity fluctuation, and the materialization errors are assumed to be negligible, so that the sampling error that is studied here is only the short-range heterogeneity fluctuation error of Gy's theory.

To evaluate the sampling error, Gy uses the heterogeneity contribution $h(i)$ as a stream descriptor:

$$h(i) = \frac{m(i)}{\bar{m}} \frac{a(i) - a_T}{a_T} \quad (4.51)$$

where:

- $m(i)$ symbolizes the mass flowrate of the reference phase at the time of extraction of the i th increment.
- \bar{m} represents the average of $m(i)$.

From Equation 4.27, the relative sampling error can be expressed as a function of $h(i)$:

$$e/a_T = \frac{1}{n} \sum_{i=1}^n \frac{m(i)}{\bar{m}} \frac{a(i) - a_T}{a_T} = \frac{1}{n} \sum_{i=1}^n h(i) \quad (4.52)$$

Instead of $h(i)$, the component flowrate $x(i)$ is used as a stream descriptor in this study, which has a concrete sense and its average value along the stream width can eventually be measured. For instance, using a magnetic flowmeter and a gamma gauge densimeter, and knowing the solid density, the ore flowrate can be continuously monitored.

Instead of using the geostatistical tool: the variogram $\nu(k)$, the signal processing tool: the autocorrelation function $\rho(k)$ is used here. The relationship between these tools is determined in the following development.

The autocorrelation function for a zero mean Gaussian variable $x(i)$ is given by (Box and Jenkins, 1973; Abraham and Ledolter, 1983; Wei, 1990):

$$\rho(k) = \frac{E[x(i)x(i+k)]}{E[x^2(i)]} = \frac{1}{(n-k)\sigma^2} \sum_{i=1}^{n-k} x(i)x(i+k) \quad (4.53)$$

Thus, the next expression is obtained:

$$(n-k)\sigma^2 \rho(k) = \sum_{i=1}^{n-k} x(i)x(i+k) \quad (4.54)$$

Now, the variogram $\nu(k)$ is given by (Gy, 1988):

$$v(k) = \frac{1}{2(n-k)} \sum_{i=1}^{n-k} (x(i+k) - x(i))^2 \quad (4.55)$$

which is equivalent to:

$$2(n-k)v(k) = \sum_{i=1}^{n-k} (x^2(i+k) + x^2(i) - 2x(i)x(i+k)) \quad (4.56)$$

Therefore, the variogram $v(k)$ is related to the variance σ^2 and the autocorrelation function $\rho(k)$ as follows:

$$2(n-k) \cdot v(k) = 2(n-k)\sigma^2 - 2(n-k)\sigma^2 \rho(k) \quad (4.57)$$

Finally, the relationship between the variogram $v(k)$ and the autocorrelation function $\rho(k)$ is:

$$v(k) = \sigma^2(1 - \rho(k)) \quad (4.58)$$

The variance of the sampling error σ_e^2 while introduced from a different point of view, is linked to the value v_0 of Gy (the discontinuity component), that is obtained from the extrapolation of $v(k)$ down to $k = 0$, in the following manner:

$$\sigma_e^2 = a_T^2 v_0 \quad (4.59)$$

Usually, v_0 is not $v(0)$ that is zero.

In his evaluation of the continuous selection error, Gy neglects the covariance terms between the sample increments. In other words, he assumes that:

$$\sigma_e^2(n, k) = \frac{1}{na_\tau^2} \sigma_e^2(1, k) \quad (4.60)$$

Equations 4.39 and 4.42 show that this is not true. However, when t_s is largely greater than the correlation length, that is the time lag k at which the autocorrelation function $\rho(k)$ becomes negligible, or for some shapes of the autocorrelation function $\rho(k)$, the approximation of Equation 4.60 is acceptable. The previous section calculates the difference between the two formulae. Equation 4.50 shows the term representing the contribution of the correlation between the incremental errors which is neglected in Gy's formula.

Gy (1988) defines an error generator that is a function of the sampling period k only:

$$W(k) = m^2 \sigma_e^2(1, k) / \sigma^2 a_\tau^2 \quad (4.61)$$

An error generator can also be defined from Equation 4.39:

$$W(n, k) = nm^2 \sigma_e^2(n, k) / \sigma^2 \quad (4.62)$$

that is a function of both the number of sample increments n and the sampling period k . This is different from Gy's expression.

The error generator of Equation 4.61 is calculated by Gy using the simple (w) and double (w') integrals of the variogram $v(k)$ of $h(i)$:

$$W = 2w(l) - w'(k)$$

Equation 4.63 gives essentially the same results as Equation 4.42 when the component flowrate is considered to be the stream descriptor. Some negligible discrepancies may appear depending on the methods of numerical approximation of w and w' .

Table 4.1 summarizes the differences between the proposed method and Gy's procedure for the evaluation of the sampling error.

Table 4.1. Comparison of the proposed method with Gy's procedure.

	Proposed method	Gy's method
Technique	Signal processing	Geostatistics
Stream descriptor	Composition $x(i)$	Heterogeneity $h(i)$
Data variation descriptor	Autocorrelation function $\rho(k)$	Variogram $\nu(k)$
Sampling error	e	e/a_T
Sampling error variance	$\sigma_e^2(n, k)$	$\sigma_e^2(1, k)/na_T^2$
Error generator	$W(n, k)$	$W(k)$

4.9. NUMERICAL EXAMPLES

For the purpose of illustrating the above formulae, eight different dynamic behaviors of component flowrates are considered (see Appendices A and H). These dynamic behaviors are modeled using *ARMA* generators of stochastic process variables (Box and Jenkins,

1973; Abraham and Ledolter, 1983; Wei, 1990). The corresponding autocorrelation functions are exhibited in Figures 4.4 through 4.6.

Figure 4.4 presents the autocorrelation functions of the simulated component flowrates following the *AR* selected models. The first autocorrelation function is obtained through a first order *AR* model, $(1 + 0.9B)x_i = \xi_i$, allowing cyclic variations. The second one, typical of a signal exhibiting an exponential decay of its correlation, is generated by an aperiodic third order *AR* model: $(1 - 0.3B - 0.3B^2 - 0.3B^3)x_i = \xi_i$.

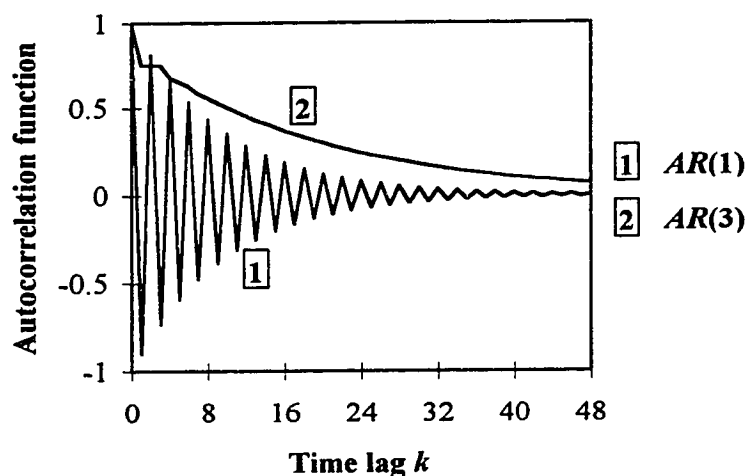


Figure 4.4. Autocorrelation functions of the *AR* series.

Figure 4.5 illustrates the autocorrelation functions of the simulated component flowrates following the *MA* selected models. The autocorrelation functions conform to the streams for which the correlation is significant up to lags of 5, 15, 30, and 45 sampling periods; they are obtained by *MA*(5): $x_i = (1 - 0.9B - \dots - 0.9B^5)\xi_i$, *MA*(15):

$x_i = (1 - 0.9B - \dots - 0.9B^{15})\xi_i$, $MA(30)$: $x_i = (1 - 0.9B - \dots - 0.9B^{30})\xi_i$, and $MA(45)$: $x_i = (1 - 0.9B - \dots - 0.9B^{45})\xi_i$ respectively.

Figure 4.6 shows the autocorrelation functions of the simulated component flowrates following the *ARMA* selected models. The first autocorrelation function is produced through an *ARMA*(3,3), $(1 - 0.9B - 0.9B^2 + 0.9B^3)x_i = (1 - 0.9B - 0.9B^2 - 0.9B^3)\xi_i$, showing cyclic variations. The second one, typical of a signal manifesting an exponential decay of its correlation, is generated by an aperiodic *ARMA*(1,5), $(1 - 0.9B)x_i = (1 - 0.9B - \dots - 0.9B^5)\xi_i$.

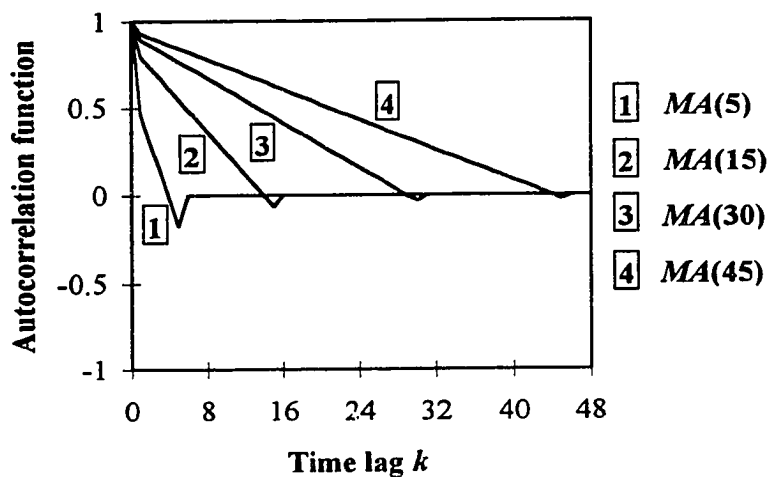


Figure 4.5. Autocorrelation functions of the four *MA* series.

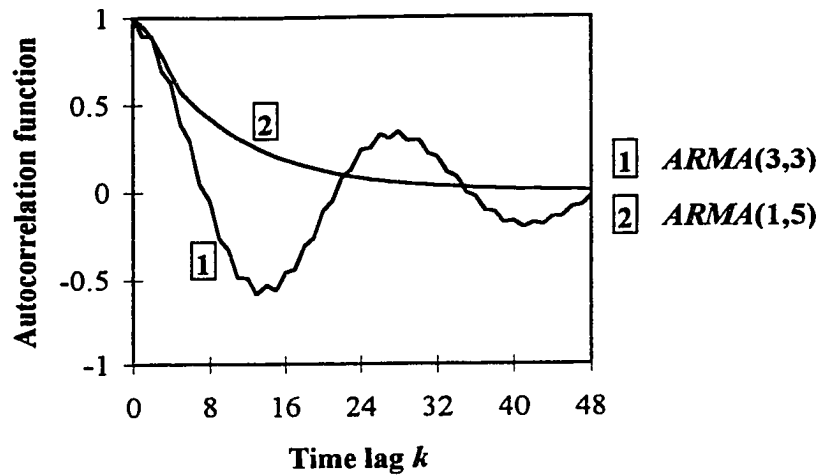


Figure 4.6. Autocorrelation functions of the two *ARMA* series.

Figures 4.7 through 4.9 show the values of the relative sampling error variance:

$$\sigma_R^2 = m^2 \sigma_e^2(n, k) / \sigma^2 \quad (4.64)$$

when the sample consists of only one increment ($n = 1$), and assuming $\sigma_d^2 = 0$. These figures indicate that the longer the correlation length the lower the sampling error variance for the same sampling period. Therefore, signals that are strongly autocorrelated lead to more representative samples. Thus, they correspond to more homogeneous material constitutions and distributions.

Furthermore, the sampling error variance tends towards the ratio σ^2/m^2 for large values of T . This is logical since when the time window T is large compared to the correlation length, the signal can be considered as stationary so that any increment in the window has a variance close to the ratio σ^2/m^2 .

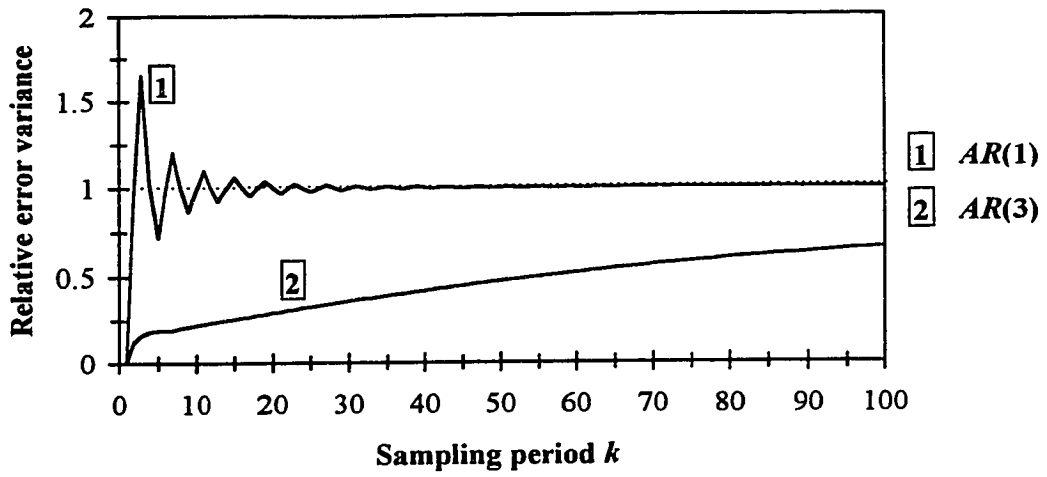


Figure 4.7. Relative sampling error variance of the *AR* series for a single-increment sample.

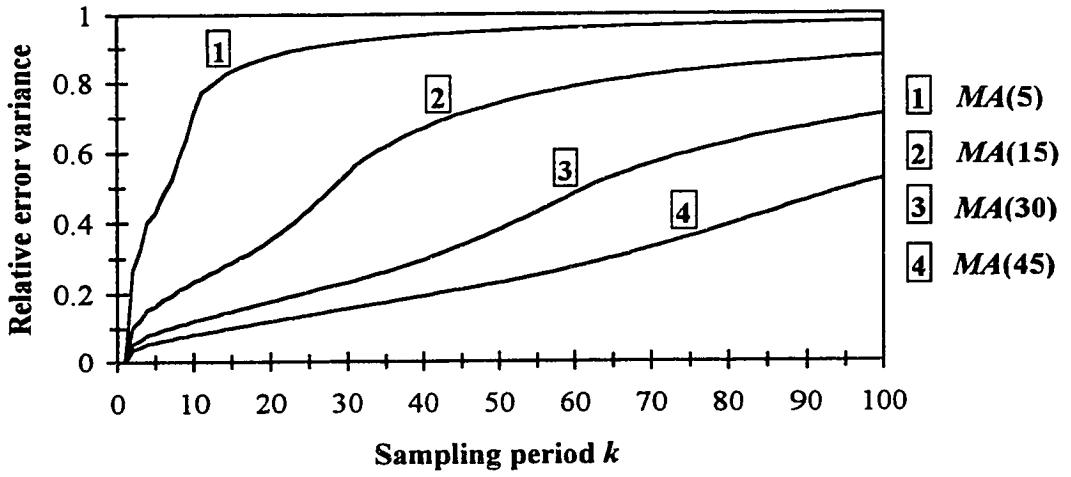


Figure 4.8. Relative sampling error variance of the *MA* series for a single-increment sample.

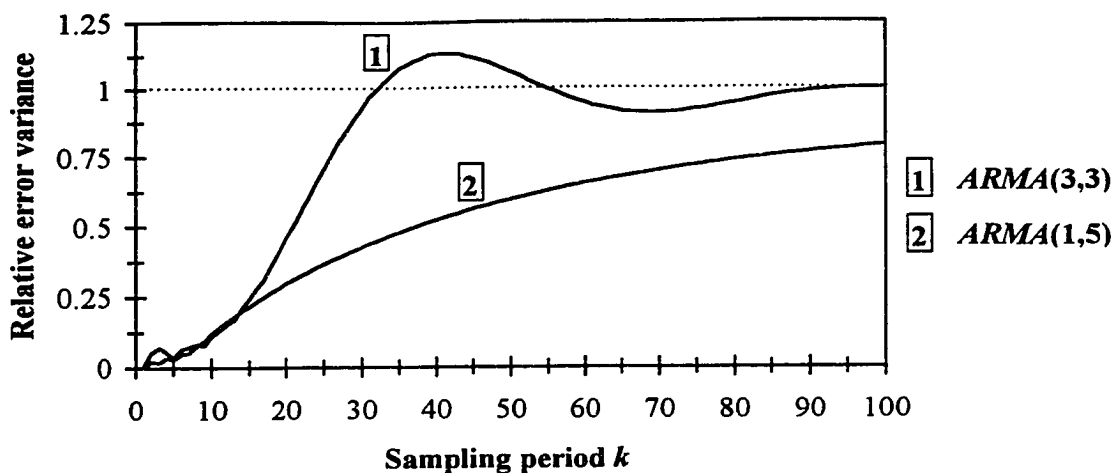


Figure 4.9. Relative sampling error variance of the *ARMA* series for a single-increment sample.

It is noticed that the stream component flowrates exhibiting periodicity in their variations, such as *AR*(1) and *ARMA*(3,3) (see Figures 4.7 and 4.9), lead to sampling error variances that can be greater than the ratio σ^2/m^2 .

Figures 4:10 through 4.12 and 4.13 through 4.15 are analogous to Figures 4.7 through 4.9 in the case of five and ten sample increments respectively. The behavior of the sampling error variances of the eight models is roughly the same as for one single increment. However oscillatory phenomena appear and become even stronger as the number of sample increments increases. They are related to the periodicity of the systematic sampling procedure. For practical purposes these oscillations can be a drawback or an advantage, depending whether the sampling period t_s corresponds to a maximum or a minimum of the variance curve. These figures demonstrate that for large values of k the variance tends towards the ratio σ^2/nm^2 .

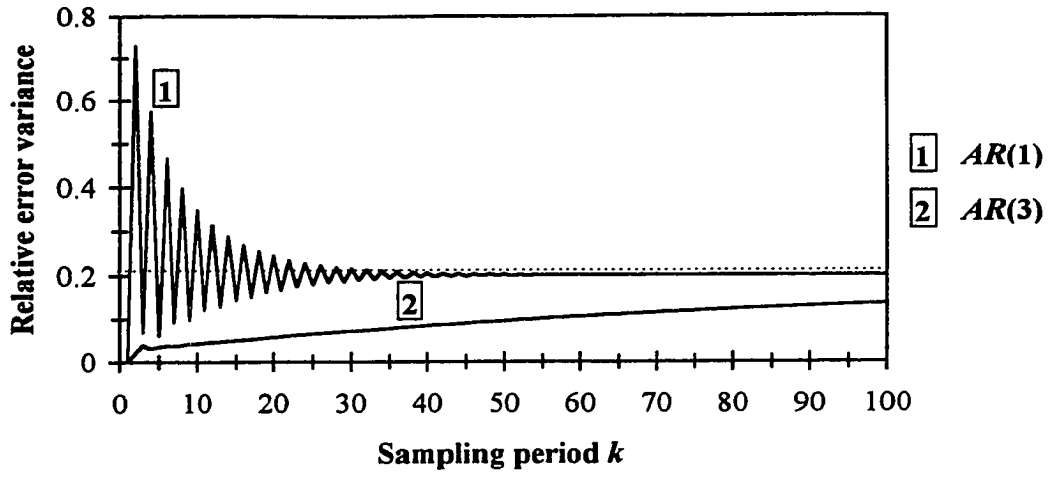


Figure 4.10. Relative sampling error variance of the AR series for a five-increment sample.

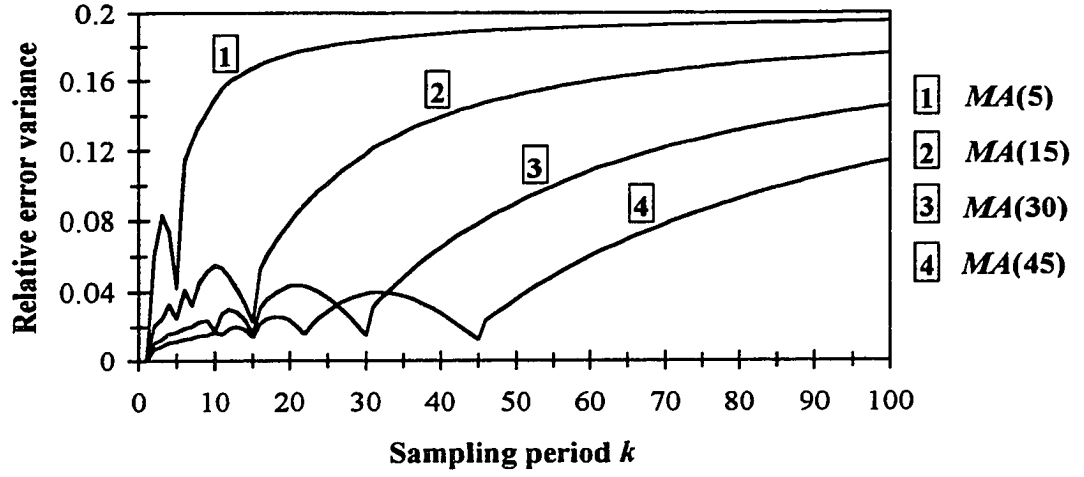


Figure 4.11. Relative sampling error variance of the MA series for a five-increment sample.

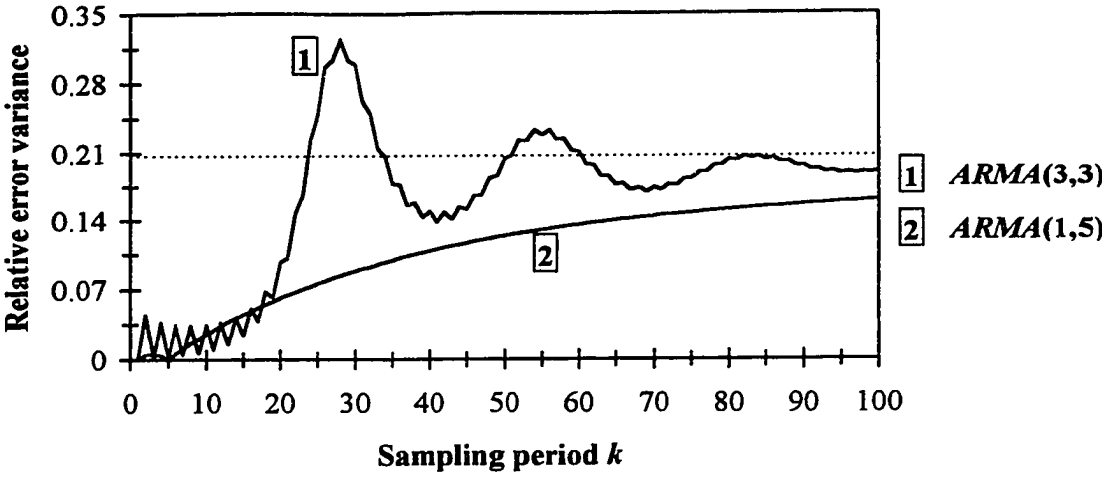


Figure 4.12. Relative sampling error variance of the ARMA series for a five-increment sample.

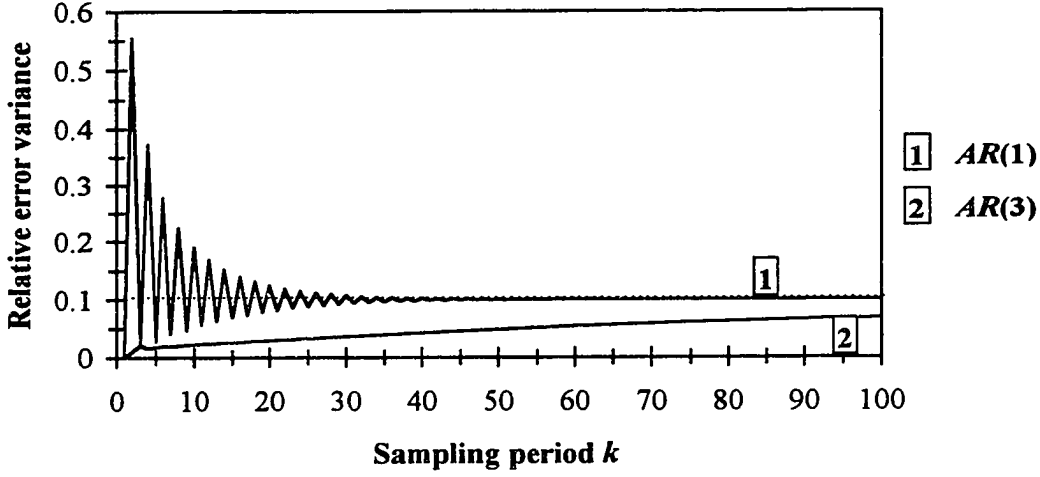


Figure 4.13. Relative sampling error variance of the AR series for a ten-increment sample.

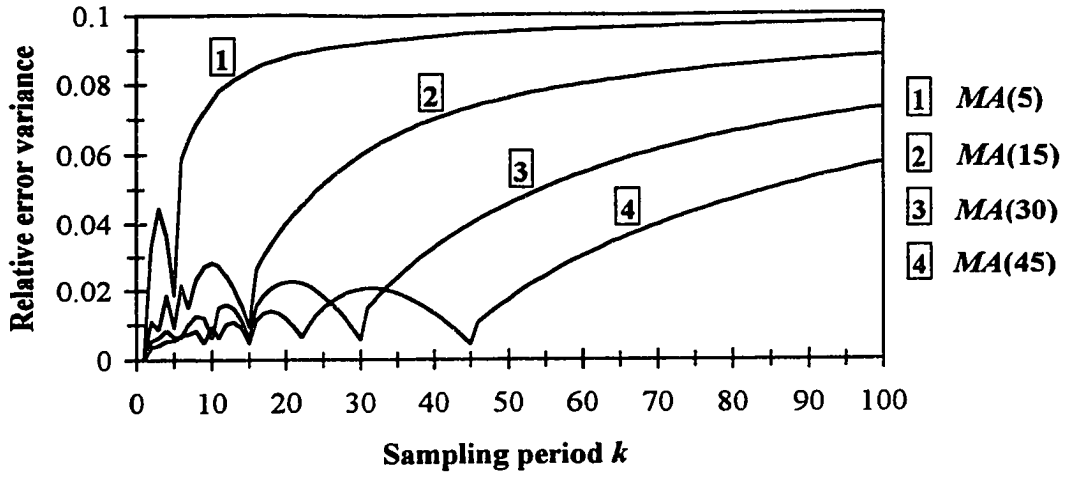


Figure 4.14. Relative sampling error variance of the *MA* series for a ten-increment sample.

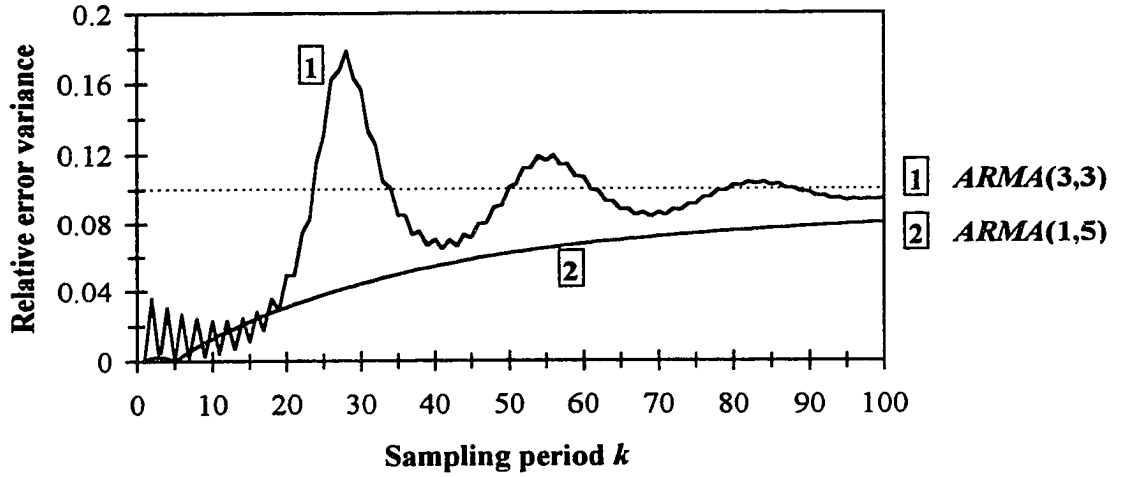


Figure 4.15. Relative sampling error variance of the *ARMA* series for a ten-increment sample.

These simulated examples can also be used to illustrate the differences that exist between Equation 4.39 and Equation 4.60 that does not take into account the correlation between the incremental errors. Figures 4.16 through 4.19 present the relative error:

$$e_R = \frac{\sigma_e^2(n, k) - \frac{\sigma_e^2(1, k)}{n}}{\sigma_e^2(n, k)} \tag{4.65}$$

for the *MA* models. Obviously the relative error vanishes when the correlation between incremental errors decreases, that is when *k* becomes large. However, for some low values of *k* compared to the correlation length the relative error is important. This means that Gy's approximation is inaccurate.

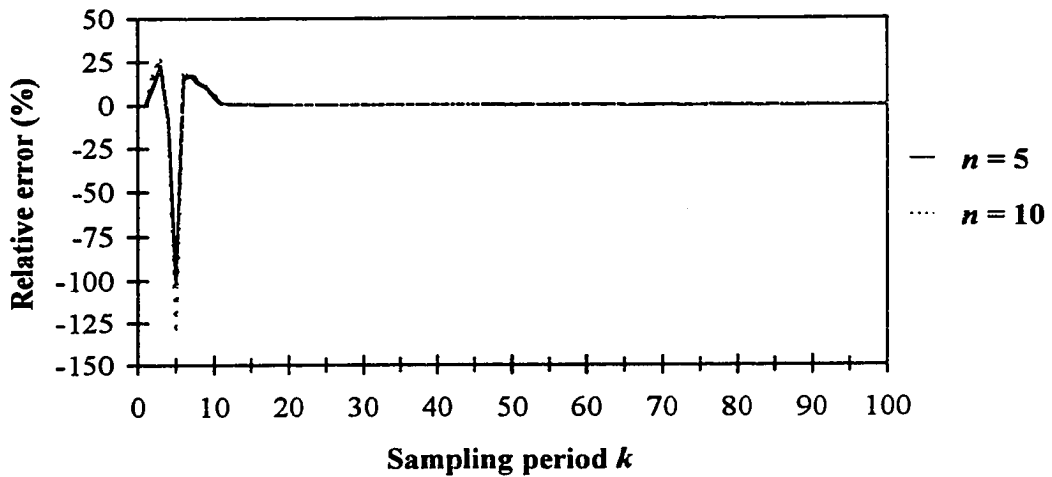


Figure 4.16. Relative error for the *MA*(5) series, $x_i = (1 - 0.9B - \dots - 0.9B^5)\xi_i$.

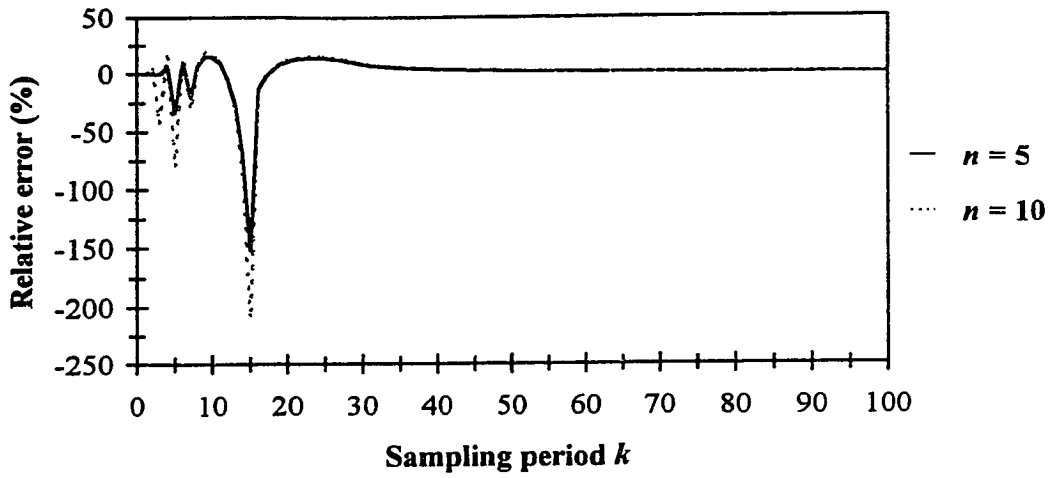


Figure 4.17. Relative error for the $MA(15)$ series, $x_i = (1 - 0.9B - \dots - 0.9B^{15})\xi_i$.

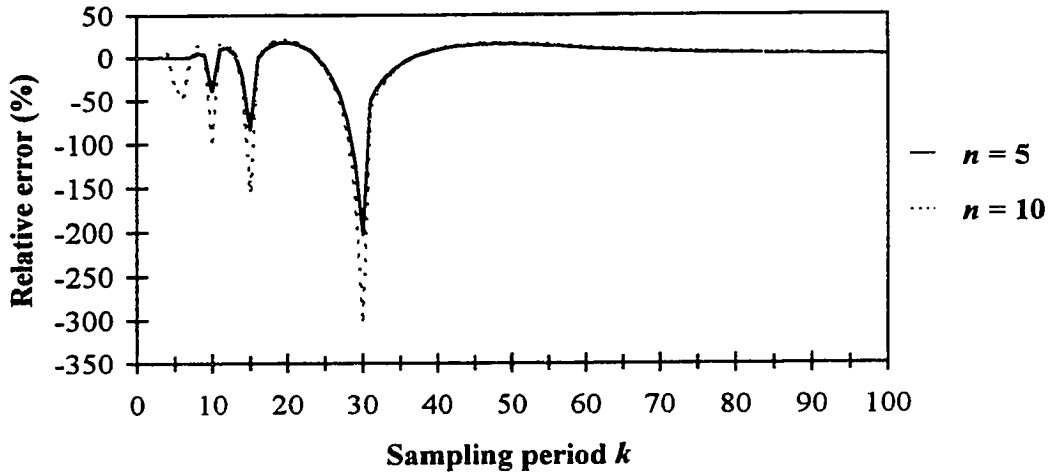


Figure 4.18. Relative error for the $MA(30)$ series, $x_i = (1 - 0.9B - \dots - 0.9B^{30})\xi_i$.

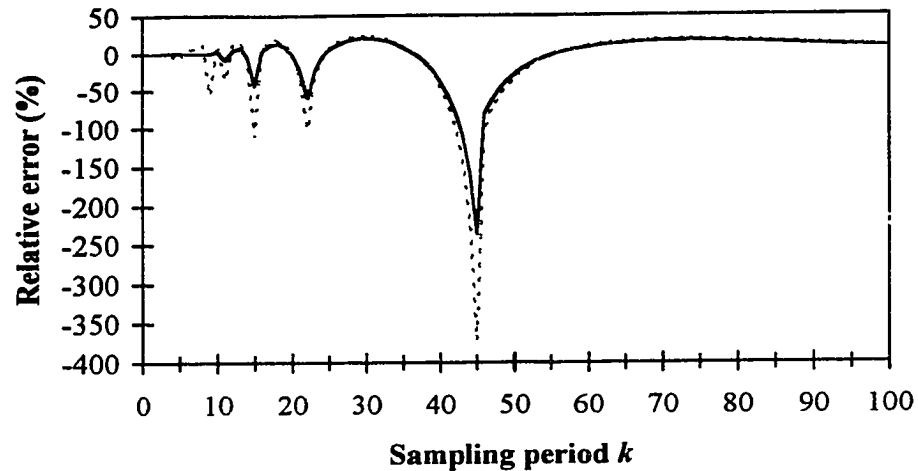


Figure 4.19. Relative error for the $MA(45)$ series, $x_i = (1 - 0.9B - \dots - 0.9B^{45})\xi_i$.

The highest relative error in absolute value is attained when the sampling period equals the correlation length. In this case, the sampling error variance is overestimated by Gy. In addition it is demonstrated that the relative error in absolute value increases with the number of sample increments and the correlation length.

The evaluation of sampling error variance is useful for the selection of the number of increments that is required to obtain a specified accuracy of the mean composition of a stream. Moreover, it can help for the evaluation of the weighting factors in data reconciliation techniques (Hodouin et al., 1989) used for process performance evaluation.

As a prerequisite to any variance calculation, the parameters σ_d^2 , σ^2 , and $\rho(k)$ must be estimated. This problem may appear quite contradictory, since to calculate these parameters, the complete signal must be known. In this case there is no reason why the average composition should be evaluated by sampling. Reciprocally when a stream has to be sampled, this is probably because no continuous monitoring sensor is available.

However, the in-plant measurement circumstances are more complex. First, it is common that a stream instrumented for control purposes with, for instance, an X-ray fluorescence analyzer, is also equipped for the collection of a composite sample. The latter is analyzed subsequently by the laboratory for the metallurgical inventory and the assessment of the on-stream analyzer correctness.

Secondly, a stream that is routinely sampled as described here can be exceptionally instrumented, using available sensors in the plant, for the purpose of determining the autocorrelation function. Finally a stream can be exceptionally sampled with a much more detailed test in order to determine the variance and covariance parameters. In the two last situations, the variance properties may be inconsistent and change from shift to shift or day to day.

An example is now considered to find the mean composition of a stream for a period T of an eight-hour shift. The discretization time t_d is selected as 10 minutes, so that $nk = 48$. A first test is run to evaluate σ_d^2 : it consists of taking 10 or 20 samples during 10 minutes and analyzing them. Then samples are taken every ten minutes during eight hours, and analyzed. This second test leads to σ^2 and $\rho(k)$. Now it is possible to calculate the sampling error variance as a function of the number of increments in such a way that nk is equal to 48.

- The results can be presented by plotting $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ as a function of the number of sample increments n . Examples are given in Figures 4.20 through 4.22 for the eight models.

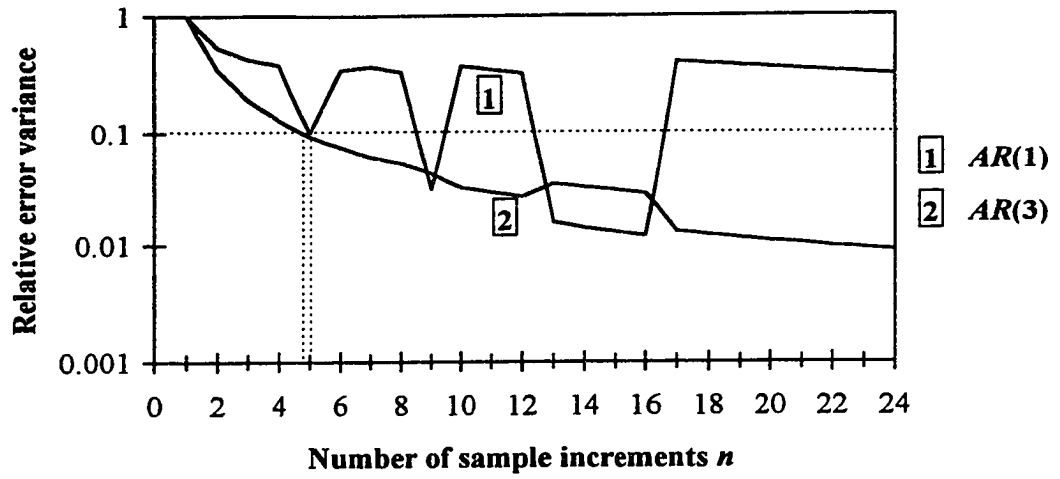


Figure 4.20. Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ where $nk = 48$ for the AR series.

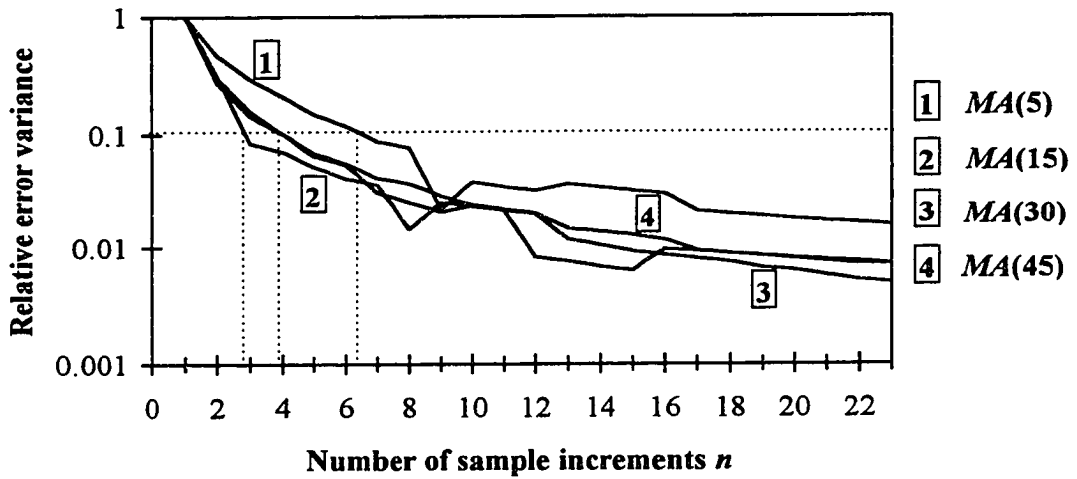


Figure 4.21. Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ where $nk = 48$ for the MA series.

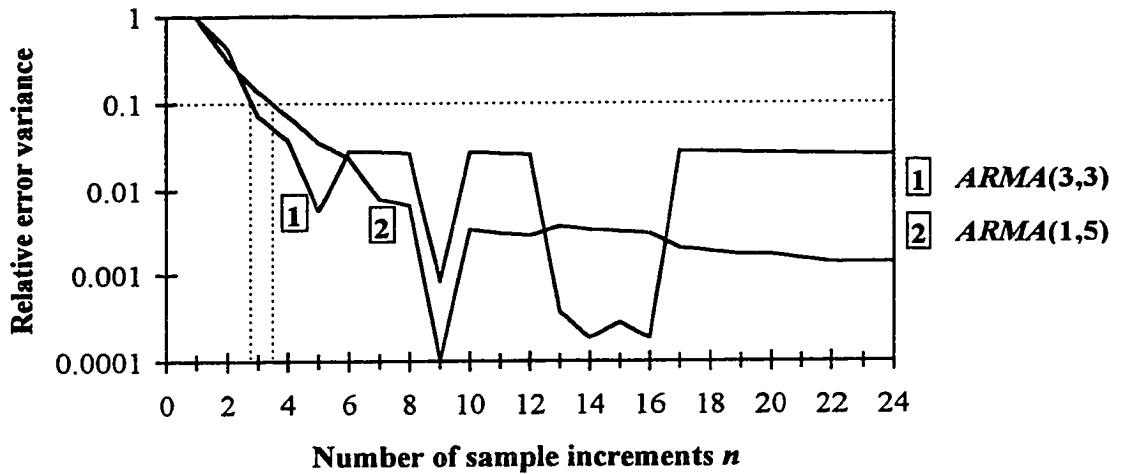


Figure 4.22. Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ where $nk = 48$ for the *MA* series.

If, for example, it is specified that the variance of the sample is aimed at being one tenth of the variance corresponding to a single increment, the graphs indicate that the recommended numbers of sample increments are:

- 3 for *MA*(15) and *ARMA*(3,3).
- 4 for *MA*(30), *MA*(45), and *ARMA*(1,5).
- 5 for *AR*(1) and *AR*(3).
- 7 for *MA*(5).

CHAPTER 5

SAMPLING ERROR FILTERING BY MATERIAL BALANCE

5.1. MATERIAL BALANCE

In a mineral processing plant, it is necessary to account for the products in terms of material and contained component weights. This is done to assess plant performance, and to control the operation using the evaluated results. Consequently, the material balance of a process flowsheet is needed.

Material balance is also named “mass balance”, “metallurgical inventory”, or “process audit”. The mass balance principle is that the total mass of material undergoing a physical or chemical transformation remains constant.

Most concentrators produce a material balance showing the performance of each shift. The shift results are cumulated over a longer period, for instance, daily, monthly, or annually, to show the overall performance.

The optimal operation of a mineral processing plant requires an accurate assessment of its performance under various conditions. The plant behavior is evaluated from material balance calculations using the measured values of process variables such as component flowrates. A component flowrate is the product of a flowrate with a composition. The measured values of these variables are obtained by stream sampling. Samples are taken periodically of the feeds, concentrates, and tailings streams. The composite samples are collected and assayed regularly, for example, at the end of each shift.

The accomplishment of a reliable material balance is difficult because of the measurement errors, inherent to the heterogeneous nature of the flowing material, the dynamic

disturbances of the processes, and the imperfection of the sensors. Mathematical filtering and estimation are powerful approaches to overcome these problems and to extract the correct information from the noisy experimental data of the plant (Hodouin et al., 1989).

Significant research work has been performed to find systematic methods of simultaneously balancing all the components in all the streams of a process flowsheet. The proposed methods are generally based on least-squares procedures (Hodouin and Flament, 1985; Laguitton, 1985; Laguitton and Hodouin, 1985; Hodouin et al., 1989; Hodouin and Flament, 1991). They correct the raw data for measurement errors in order to make them consistent and simultaneously calculate the unknown variables.

Such data-adjustment techniques are utilized for global inventory of the inputs and outputs of a plant. More frequently, they are used to obtain the best representation of the state of a process. They are also powerful tools to evaluate the performance of processes since they give accurate values of the various component flowrates. Besides, they are a preliminary step to upgrade raw data obtained after sampling around various process units. Furthermore, they are employed to give a reliable database for the building of unit models that are the basic elements of a process or plant computer simulator.

5.2. FLOTATION PROCESS UNIT MODELS

A plant flowsheet can be reduced to a series of nodes, where process streams either join or separate. A node can have:

- An input and an output in case of a delay (see Figure 5.1) or a mixer (see Figure 5.2).
- An input and two outputs in case of a separator (see Figure 5.3).
- Two inputs and one output in case of a junction (see Figure 5.4).

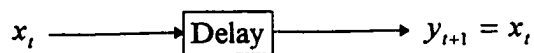


Figure 5.1. Delay node.



Figure 5.2. Mixer node.

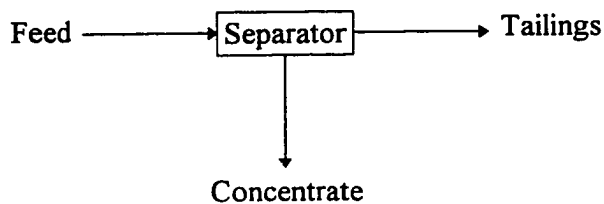


Figure 5.3. Separator node.

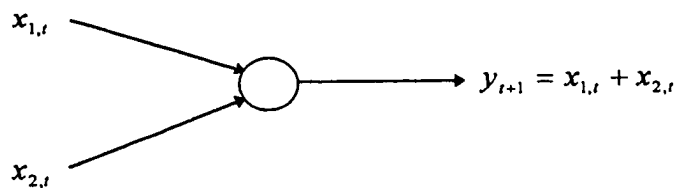


Figure 5.4. Junction node.

A flotation cell (see Figure 5.5) can be represented by a set of a mixer node and a separator node as shown in Figure 5.6.

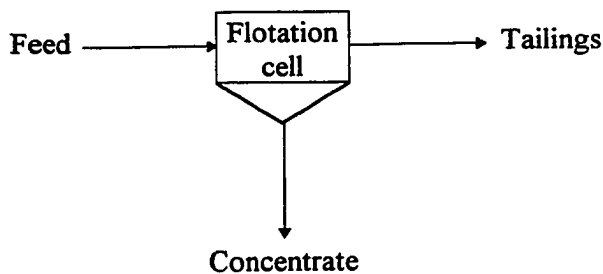


Figure 5.5. Flotation cell in flowsheet form.

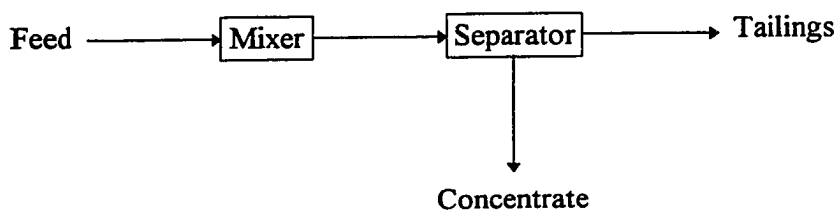


Figure 5.6. Flotation cell in node form.

5.2.1. Mixer model

A mixer is assumed to be described by a first-order differential equation for the component flowrates, with x and y being its input and output component flowrates respectively:

$$\tau \dot{y} + y = x \quad (5.1)$$

where τ represents the time constant of the mixer. It is the time necessary for the mixer to adjust to a change in its input flowrate (Stephanopoulos, 1984).

Equation 5.1 can be written in a discrete form:

$$\tau \frac{y_{t+1} - y_t}{\Delta t} + y_t = x_t \quad (5.2)$$

because Δt , that is the time interval between two time indices, is negligible compared to the sampling period T . As a result, it follows that:

$$y_{t+1} = \left(1 - \frac{\Delta t}{\tau}\right) y_t + \frac{\Delta t}{\tau} x_t \quad (5.3)$$

or in a more compact form:

$$y_{t+1} = K y_t + (1 - K) x_t \quad (5.4)$$

where:

$$K = 1 - \frac{\Delta t}{\tau} \quad (5.5)$$

5.2.2. Separator model

A separator is characterized by its separation efficiency. The latter is supposed to be equal to the recovery R of the ore concentration:

$$R = \frac{x_C}{x_F} \quad (5.6)$$

where x_F and x_C are the component flowrates in the feed and the concentrate streams respectively.

Thus, the tailings component flowrate x_T is given by:

$$x_T = (1 - R)x_F \quad (5.7)$$

5.3. MASS CONSERVATION EQUATIONS

In all mineral processes the mass of consumed elements is equal to that of produced elements. For example, the mass of copper entering a flotation cell per time unit is equal to that leaving the same unit per time unit provided the steady-state conditions have been reached.

In a single unit such as shown in Figure 5.7, if x_i represents the component flowrate of stream i , the following mass conservation equation is valid:

$$x_1^* = x_2^* \quad (5.8)$$

where the asterisk denotes the true value of the variables.

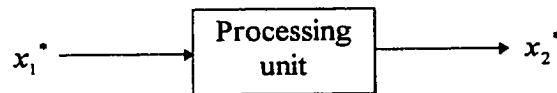


Figure 5.7. One product unit.

In practice, only an experimental estimate of x_i can be measured. Obviously, the better the experimental precision, the more valid the assumption that it is the best estimate. Better estimates than x_i can be calculated. They are represented by \hat{x}_i , the maximum likelihood estimates.

The mass conservation equations, for different components flowing in a mineral processing plant can be written using the concept of networks. A network is a graph where the branches are the streams and the nodes the process units. A network is associated with each component of the material flowing in the plant.

A network is characterized by an incidence matrix and each branch by its component flowrates. Each line of the incidence matrix corresponds to a node of the network and each column to a stream. If a stream converges to or diverges from a node, a value +1 or -1 respectively is entered at the corresponding position in the matrix. If a stream is not connected to a node a zero value is entered.

Each node corresponds to a mass conservation equation. For a steady state operation, the sum of the input component flowrates is equal to the sum of the output component flowrates for the same component. The mass conservation equations are written at each node of the network. Thus, the system of mass conservation equations is formulated as:

$$M_i X_i = 0 \quad (5.9)$$

where M_i is the incidence matrix of the process network.

5.4. DATA REDUNDANCY

From the number of mass conservation equations, the number of process variables and the number of measurements, a redundancy degree (Hodouin and Flament, 1989) is determined. The redundancy degree is zero when the measurement design is minimal, that is, when the data set is not sufficient to allow the estimation of all process variables of interest. On the opposite, the redundancy degree equals one when all the process variables are measured.

Hodouin et al. (1989) demonstrates that, for a specific data reconciliation procedure based on mass balance filtering, an increasing redundancy degree results in a higher reduction of the variance of measured variables and lower estimated variances for unmeasured variables.

A degree of redundancy μ_r is defined by (Hodouin et al., 1989):

$$\mu_r = \frac{n_m - n_i}{n_v - n_i} \quad (5.10)$$

where n_v is the number of process variables, n_i and n_m are respectively the numbers of independent and measured variables. It follows that μ_r is one when all the process variables are measured.

The number of independent variables n_i is given by (Hodouin et al., 1989):

$$n_i = n_v - n_e \quad (5.11)$$

where n_e is the number of equations in the process network.

The mass balance situation that is considered in this chapter is the redundant case. Hence, all the component flowrate variables are assumed to be measured.

5.5. LEAST-SQUARES ESTIMATION

The observed values of the measured variables are redundant and erroneous. Then the system $M_j X_j = 0$ does not have a solution. To solve it the first step is to correct the

values of the measured variables contained in X_t^m , that is Y_t , replacing them by their estimates in \hat{X}_t^m such that:

$$J = 0.5(X_t^m - \hat{X}_t^m)' V_x^{-1} (X_t^m - \hat{X}_t^m) \quad (5.12)$$

is minimum under the constraints:

$$M_t(\hat{X}_t - X_t^*) = 0 \quad (5.13)$$

In equation 5.12, V_x symbolizes the variance matrix of X_t^m , $(X_t^m - \hat{X}_t^m)'$ the transpose of $(X_t^m - \hat{X}_t^m)$. J is a least-squares criterion whose terms are weighted by the inverse of the measurement variances or covariances.

Thus, a least-squares problem is obtained. In this way, the following should be minimized:

$$0.5r'V_x^{-1}r \quad (5.14)$$

that is subject to:

$$M_t(e - r) = 0 \quad (5.15)$$

where r is the residual $(X_t - \hat{X}_t)$, e the measurement error $(X_t - X_t^*)$, and $(e-r)$ is the estimation error ε .

The constrained minimum is attained by forming the Lagrangian:

$$L = 0.5r'V_x^{-1}r + \lambda'(M_1(e-r)) \quad (5.16)$$

where λ is the Lagrange multiplier, and establishing the stationarity conditions:

$$\frac{\partial L}{\partial r} = V_x^{-1}r - M_1' \lambda = 0 \quad (5.17)$$

and:

$$\frac{\partial L}{\partial \lambda} = M_1(e-r) = 0 \quad (5.18)$$

Solving Equation 5.17 for the Lagrange multiplier λ , the latter is obtained:

$$\lambda = (M_1V_xM_1')^{-1} M_1r \quad (5.19)$$

Using Equations 5.18 and 5.19, the residual r is found:

$$r = V_xM_1' \alpha M_1e \quad (5.20)$$

where:

$$\alpha = (M_1V_xM_1')^{-1} \quad (5.21)$$

5.6. VARIANCE ANALYSIS

Using the expression developed for r , the variance matrix V_r of the residuals is given by:

$$V_r = E[rr'] = V_x M_1' \alpha M_1 V_x \quad (5.22)$$

where E stands for the mathematical expectation.

The variance matrix V_ε of the estimates is acquired:

$$V_\varepsilon = E[\varepsilon\varepsilon'] = E[(e-r)(e-r)'] = V_x + V_r - 2V_{re} \quad (5.23)$$

where V_{re} is the covariance between the residual and the measurement errors:

$$V_{re} = E[re'] = E[V_x M_1' \alpha M_1 ee'] = V_x M_1' \alpha M_1 V_x = V_r \quad (5.24)$$

Therefore, the variance matrix V_ε of the estimates becomes:

$$V_\varepsilon = V_x - V_r \quad (5.25)$$

The variance elements that are the diagonals of the matrices V_x , V_r , and V_ε should be positive.

5.7. STREAM VARIABLES

A stream is sampled for the determination of its average compositions. Usually, the composition is defined, for a set of stream components, as the mass fractions of these components with respect to a reference phase. The latter could be either the overall stream material or a part of it. For example, the following composition definitions could be used:

- Mass fraction of solids in the slurry.
- Mass fraction of one metal in the solid phase.
- Mass fraction of one mineral in a given particle size interval.

In the following development, all the streams are assumed to be sampled simultaneously. Now, if only one selected component of the streams is analyzed, the mass fraction of that component in the samples is represented by the composition vector $Z_{1,S}$:

$$Z_{1,S} = \frac{\sum_{i=1}^{n_i} t(i)M_1(i)Z_1(i)}{\sum_{i=1}^{n_i} t(i)M_1(i)} \quad (5.26)$$

where:

- i is the index of the samples increments.
- n_i is the number of sample increments.
- $M_1(i)$ is the $k \times k$ diagonal matrix of average mass flowrates of the reference phases at the time of the extraction of the i th sample increments:

$$M_1(i) = \begin{bmatrix} m_1(i) & 0 & \dots & 0 \\ 0 & m_2(i) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & m_k(i) \end{bmatrix} \quad (5.27)$$

where $m_1(i)$, ..., and $m_k(i)$ denote the average mass flowrates of the reference phases 1, ..., and k respectively.

- $Z_1(i)$ is the k vector of average mass fractions of the component of interest at the time of the extraction of the i th sample increments.
- $t(i)$ is the time during which each part of the stream is sampled.

$t(i)$ is given by Equation 4.3. As the transfer time t_p of the samplers has been assumed to be constant, $t(i)$ is independent of i and as a result:

$$Z_{1,S} = \frac{\sum_{i=1}^n M_1(i)Z_1(i)}{\sum_{i=1}^n M_1(i)} = \frac{1}{n_i \bar{M}_1} \sum_{i=1}^n M_1(i)Z_1(i) \quad (5.28)$$

where \bar{M}_1 is the average flowrate of the reference phase over the n sample increments.

Equations 5.26 through 5.28 define a mineral processing system for only one selected component in the streams. Indeed, many components are present in the streams. Suppose that the number of components in each stream equals n . As a result, a new state equation is generated:

$$Z_S = [Z_{1,S}, \dots, Z_{n,S}]' = \frac{1}{n_i \bar{M}} \sum_{i=1}^n M(i)Z(i) \quad (5.29)$$

where Z_S is a state vector of dimension kn , $M(i)$ the mass flowrate matrix of dimension $kn \times kn$ given by:

$$M(i) = \begin{bmatrix} M_1(i) & 0 & \dots & 0 \\ 0 & M_2(i) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & M_n(i) \end{bmatrix} \quad (5.30)$$

and \bar{M} the average of the matrix M for the n_i sample increments.

Therefore, the measured property of the sample is obtained by an averaging process of the component flowrate vector Y rather than by an averaging process of the composition vector Z . For that reason the stream is characterized in the following development by the component flowrate vector $Y(y,t)$.

The variable $Y(y,t)$ has a concrete sense and its average value along the streams widths can eventually be measured. For instance, using a magnetic flowmeter and a gamma gauge densimeter, and knowing the solid density, the ore flowrate can be continuously monitored. The ore flowrate is the product of the mass flowrate of the slurry with the solid mass fraction.

The variable $Y(y,t)$ can be understood as a time variable distributed along the y axis of the streams widths (Figure 2.3). Because of the type of sampling device, the heterogeneity along the z -axis (Figure 2.2) does not need to be described. The continuity of Y can be discussed because of the discrete nature of the particulate material flowing through the sections S (Figure 2.2). Both discrete and continuous approaches are correct, although the continuous formalism is easier to manipulate. However, to calculate the sampling error the signals of the vector Y must be discretized since the increment extraction from the stream is typically a discretization process.

In the following sections it is assumed that the variables $M(i)$, $Z(i)$, and $Y(i)$ obey normal distributions.

5.8. DISCRETIZATION OF THE COMPONENT FLOWRATES

The discretization operator that is used is a mathematical transformation that represents the sampling of streams by moving cutters as previously described. The time is discretized using a time period t_d such that $t_d \geq t_p$. Within each time interval $[(i-1)t_d, it_d]$ of index i , the continuous function $Y(y,t)$ is replaced by the value:

$$Y(i) = \frac{v}{w} \int_{y=0}^L \int_{t=\tau}^{\tau'} Y(y,t) dt dy \quad (5.31)$$

with:

$$\tau = t_0 + y/v + (i-1)t_d \quad (5.32)$$

and:

$$\tau' = \tau + w/v \quad (5.33)$$

where t_0 is the elapsed time since the beginning of the interval when the cutters are started. The discretization operator is depicted in Figure 4.1. This discretizer has a double function. First it is a filter that smoothes the continuous function $Y(y,t)$ in both y and t directions through the double integral. Second it is a discretizer in the time direction.

The physical discretization operation by moving cutters converges towards the mathematical operator of Equation 5.31 only when the number of particles entering each

sample tends to infinity. The difference between the two operators is due to the particulate structure of the flowing material. The mathematical integration is not perfectly performed by the sample increment since particles cannot be cut by the edges of the cutter. Statistically this difference is vanishing when the number of particles considered in the integral is large.

5.9. DISCRETIZATION ERRORS

The smoothing or discretization procedure necessarily produces a loss of information, which is characterized here by the discretization error. This error is the difference between the discretized value $Y(i)$ and the true mean value $Y^*(i)$ within the time interval $[(i-1)t_d, it_d]$. It can be evaluated by:

$$Y_d(i) = Y(i) - Y^*(i) \quad (5.34)$$

with:

$$Y^*(i) = \frac{1}{t_d} \int_0^{it_d} \int_{(i-1)t_d}^{it_d} Y(y, t) dt dy \quad (5.35)$$

It is assumed that the variables making up Y are randomly distributed, so that the variables of Y_d are also random. Due to the assumption of sampling correctness the mean value of Y_d is zero. Its variance matrix V_d depends on the local variances of Y variables that are the variances in the frequency range $\left[\frac{2\pi}{t_d}, \frac{2\pi\nu}{w} \right]$. The larger they are, the larger are the discretization errors variances. It also depends on the ratio t_p/t_d : the closer to one, the smaller are the discretization errors variances.

Although a continuous formalism for Y is used, which embeds the constitution and distribution heterogeneities, the discretization error variance matrix includes both the fundamental and the grouping and segregation errors defined by G_y . When the number of particles entering the cutter is too small, the difference between the mathematical and the physical discretizers is an additional contribution to the variance matrix of Y_d .

Finally, it is assumed that the discretization error Y_d has a constant variance V_d and is uncorrelated from discretization interval to discretization interval. These properties are quite reasonable, since they are based on the assumption that the Y variables are locally stationary, that is, stationary within a window of width t_d .

When the mean value of Y is not constant within a t_d interval, the function Y can be replaced by its variations around its moving averages. This will filter out the frequencies lower than $2\pi/t_d$. The variance matrix V_d is the variance matrix of these variations. This is defined by local variances, or in other words, the Y variance matrix for frequencies higher than $2\pi/t_d$. These assumptions mean that the function Y can be viewed as the sum of two functions:

$$Y(t) = Y_\mu(t) + Y_\varepsilon(t) \quad (5.36)$$

$Y_\mu(t)$ contains the low frequencies of $Y(t)$ and $Y_\varepsilon(t)$ the high frequencies. The time-discrete version of this equation is:

$$Y(i) = Y^*(i) + Y_d(i) \quad (5.37)$$

with:

$$E(Y_d) = 0 \quad (5.38)$$

$$E(Y_d Y_d') = V_d \quad (5.39)$$

$$E[Y_d(i) Y_d(i+j)] = 0, \quad \forall j \neq i \quad (5.40)$$

and:

$$E[Y^*(i) Y_d(i+j)] = 0, \quad \forall i, j \quad (5.41)$$

This decomposition can be graphically understood as depicted for a variable $x(t)$ in Figure 4.2.

5.10. SAMPLING ERRORS

Equation 5.36 shows that the component flowrate vector Y is the sum of two time-discrete functions where $Y^*(i)$ is the real mean value of the component flowrate in the time interval $[(i-1)t_d, it_d]$, and $Y_d(i)$ the discretization error considered as a vector of white sequences. Now it is assumed that $Y^*(i)$ contains stationary random signals with a mean value:

$$E[Y^*(i)] = E[Y(i)] = Y^* = M^* Z^* \quad (5.42)$$

and a variance:

$$E\left[(Y^*(i) - Y^*)(Y^*(i) - Y^*)'\right] = V^* \quad (5.43)$$

where Y^* , M^* , and Z^* represent the overall mean matrices of Y , M , and Z respectively, and V^* the variance of $Y^*(i)$.

As a result the variance of $Y(i)$ is:

$$V_Y = V^* + V_d \quad (5.44)$$

The autocovariance of $Y^*(i)$ is:

$$V_{YY^*} = E \left[\left(Y^*(i) - Y^* \right) \left(Y^*(i+k) - Y^* \right)' \right] \quad (5.45)$$

Consequently, the autocovariance of $Y(i)$ is:

$$V_{YY}(k) = E \left[\left(Y(i) - Y^* \right) \left(Y(i+k) - Y^* \right)' \right] = V_{YY^*}(k) \quad (5.46)$$

The sampling error Z_e (see Figure 5.8) is the difference between the composition vector Z_S in the composite sample and the composition vector Z_T in the total amounts of reference phases.

The expression for Z_S is obtained from Equation 5.29:

$$Z_S = \frac{1}{n\bar{M}} \sum_{i=0}^{n-1} Y(l+ik) = \frac{1}{n\bar{M}} \sum_{i=0}^{n-1} Y^*(l+ik) + \frac{1}{n\bar{M}} \sum_{i=0}^{n-1} Y_d(l+ik) \quad (5.47)$$

where:

- n is the number of increments in the sample.
- k is the number of discretization intervals t_d in the sampling period t_s . k is an integer given by Equation 4.22.
- l is the time index of the first samples increments. It is calculated following Equation 4.23 or 4.24.

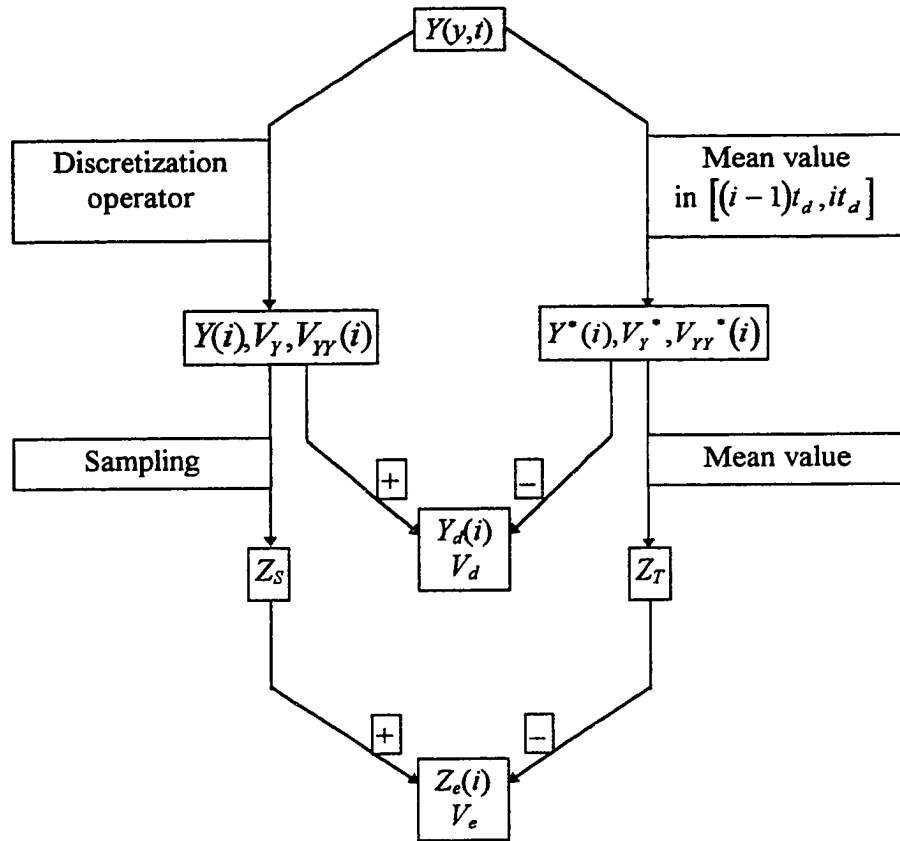


Figure 5.8. Process leading to the sampling error variance matrix.

The true composition matrix Z_T in the time interval T , that is, from $i = 1$ to nk , is:

$$Z_T = \frac{1}{M_T nk} \sum_{i=1}^{nk} Y^*(i) \quad (5.48)$$

and by definition the sampling error is:

$$Z_e = Z_S - Z_T \quad (5.49)$$

which is equivalent to:

$$Z_e = \frac{1}{nM} \sum_{i=1}^n M(i) [Z(i) - Z_T] \quad (5.50)$$

Because of the sampling correctness assumption, the sampling error has a zero mean value:

$$E[Z_e] = 0 \quad (5.51)$$

and a variance equal to:

$$V_e = E[Z_e Z_e'] = E[(Z_S - Z^*)(Z_S - Z^*)'] + E[(Z_T - Z^*)(Z_T - Z^*)'] - E[(Z_S - Z^*)(Z_T - Z^*)'] - E[(Z_T - Z^*)(Z_S - Z^*)'] \quad (5.52)$$

5.11. SAMPLING ERROR VARIANCE

The variance matrix of sampling errors is calculated before and after material balancing in order to know the influence of this technique on the sampling errors and therefore on the accuracy of the estimates.

Using Equation 5.51, the variance of Z_e is then:

$$\begin{aligned}
 V_e &= \frac{1}{n^2} \bar{M}^{-1} \text{var} \left[\sum_{i=0}^{n-1} Y^*(l+ik) \right] \bar{M}'^{-1} + \frac{1}{n} \bar{M}^{-1} V_d \bar{M}'^{-1} \\
 &+ \frac{1}{(nk)^2} M_T^{-1} \text{var} \left[\sum_{i=1}^{nk} Y^*(i) \right] M_T'{}^{-1} - \frac{1}{n^2 k} \bar{M}^{-1} \text{cov} \left[\sum_{i=0}^{n-1} Y^*(l+ik); \sum_{j=1}^{nk} Y^*(j) \right] M_T'{}^{-1} \\
 &- \frac{1}{n^2 k} \bar{M}^{-1} \text{cov} \left[\sum_{i=1}^{nk} Y^*(i); \sum_{j=1}^{n-1} Y^*(l+jk) \right] M_T'{}^{-1}
 \end{aligned} \tag{5.53}$$

where $\text{var}[\]$ and $\text{cov}[\]$ stand for respectively the variance and covariance of the term within the brackets. Each term of Equation 5.53 can now be evaluated:

$$\text{var} \left[\sum_{i=0}^{n-1} Y^*(l+ik) \right] = E \left[\left(\sum_{i=0}^{n-1} (Y^*(l+ik) - Y^*) \right) \left(\sum_{j=0}^{n-1} (Y^*(l+jk) - Y^*) \right)' \right] \tag{5.54}$$

which equals:

$$\text{var} \left[\sum_{i=0}^{n-1} Y^*(l+ik) \right] = nV_Y^* + \sum_{i=1}^{n-1} (n-i) (V_{YY^*}(ik) + V_{YY^*}(-ik)) \tag{5.55}$$

$$\text{var} \left[\sum_{i=1}^{nk} Y^*(i) \right] = E \left[\left(\sum_{i=1}^{nk} (Y^*(i) - Y^*) \right) \left(\sum_{j=1}^{nk} (Y^*(j) - Y^*) \right)' \right] \tag{5.56}$$

that is:

$$\text{var} \left[\sum_{i=1}^{nk} Y^*(i) \right] = nkV_Y^* + \sum_{i=1}^{nk} (nk-i) (V_{YY^*}(i) + V_{YY^*}(-i)) \tag{5.57}$$

$$\text{cov} \left[\sum_{i=0}^{n-1} Y^*(l+ik); \sum_{j=1}^{nk} Y^*(j) \right] = E \left[\left(\sum_{i=0}^{n-1} (Y^*(l+ik) - Y^*) \right) \left(\sum_{j=1}^{nk} (Y^*(j) - Y^*) \right)' \right] \quad (5.58)$$

which is:

$$\text{cov} \left[\sum_{i=0}^{n-1} Y^*(l+ik); \sum_{j=1}^{nk} Y^*(j) \right] = \sum_{j=0}^{n-1} \sum_{i=1}^{nk} V_{YY^*}(l+jk-i) \quad (5.59)$$

and:

$$\text{cov} \left[\sum_{i=1}^{nk} Y^*(i); \sum_{j=0}^{n-1} Y^*(l+jk) \right] = E \left[\left(\sum_{i=1}^{nk} (Y^*(i) - Y^*) \right) \left(\sum_{j=0}^{n-1} (Y^*(l+jk) - Y^*) \right)' \right] \quad (5.60)$$

that implies:

$$\text{cov} \left[\sum_{i=1}^{nk} Y^*(i); \sum_{j=0}^{n-1} Y^*(l+jk) \right] = \sum_{i=1}^{nk} \sum_{j=0}^{n-1} V_{YY^*}(i-l-jk) \quad (5.61)$$

Using Equations 5.55, 5.57, 5.59, 5.61, and 5.53, the sampling error variance matrix becomes:

$$\begin{aligned} V_e(n, k) &= \frac{1}{n^2} \bar{M}^{-1} \left[nV_Y^* + \sum_{i=1}^{n-1} (n-i)(V_{YY^*}(ik) + V_{YY^*}(-ik)) \right] \bar{M}'^{-1} + \frac{1}{n} \bar{M}^{-1} V_d \bar{M}'^{-1} \\ &+ \frac{1}{(nk)^2} M_T^{-1} \left[nkV_Y^* + \sum_{i=1}^{nk} (nk-i)(V_{YY^*}(i) + V_{YY^*}(-i)) \right] M_T'^{-1} \\ &- \frac{1}{n^2 k} \bar{M}^{-1} \left[\sum_{j=0}^{n-1} \sum_{i=1}^{nk} V_{YY^*}(l+jk-i) \right] M_T'^{-1} - \frac{1}{n^2 k} \bar{M}^{-1} \left[\sum_{i=1}^{nk} \sum_{j=0}^{n-1} V_{YY^*}(i-l-jk) \right] M_T'^{-1} \end{aligned} \quad (5.62)$$

It is impossible to measure the difference between \bar{M} and M_T . Thus, they are assumed to have a common estimate M :

$$\bar{M} = M_T = M \quad (5.63)$$

which finally leads to the following expression of the variance of the sampling error matrix:

$$\begin{aligned} V_e(n, k) = & \frac{1}{n^2} M^{-1} \left[nV_Y^* + \sum_{i=1}^{n-1} (n-i)(V_{YY}^*(ik) + V_{YY}^*(-ik)) \right] M'^{-1} + \frac{1}{n} M^{-1} V_d M'^{-1} \\ & + \frac{1}{(nk)^2} M^{-1} \left[nkV_Y^* + \sum_{i=1}^{nk} (nk-i)(V_{YY}^*(i) + V_{YY}^*(-i)) \right] M'^{-1} \\ & - \frac{1}{n^2 k} M^{-1} \left[\sum_{j=0}^{n-1} \sum_{i=1}^{nk} V_{YY}^*(l+jk-i) \right] M'^{-1} - \frac{1}{n^2 k} M^{-1} \left[\sum_{i=1}^{nk} \sum_{j=0}^{n-1} V_{YY}^*(i-l-jk) \right] M'^{-1} \end{aligned} \quad (5.64)$$

Using Equations 5.44 and 5.46, Equation 5.64 can be written as follows:

$$\begin{aligned} V_e(n, k) = & \frac{1}{n^2} M^{-1} \left[nV_Y + \sum_{i=1}^{n-1} (n-i)(V_{YY}(ik) + V_{YY}(-ik)) \right] M'^{-1} - \frac{1}{nk} M^{-1} V_d M'^{-1} \\ & + \frac{1}{(nk)^2} M^{-1} \left[nkV_Y + \sum_{i=1}^{nk} (nk-i)(V_{YY}(i) + V_{YY}(-i)) \right] M'^{-1} \\ & - \frac{1}{n^2 k} M^{-1} \left[\sum_{j=0}^{n-1} \sum_{i=1}^{nk} V_{YY}(l+jk-i) \right] M'^{-1} - \frac{1}{n^2 k} M^{-1} \left[\sum_{i=1}^{nk} \sum_{j=0}^{n-1} V_{YY}(i-l-jk) \right] M'^{-1} \end{aligned} \quad (5.65)$$

It is interesting to consider particular applications of the general formula (Equation 5.65). First, in the case where there is no autocorrelation of the signals making up $Y(i)$, the sampling error variance becomes:

$$V_e(n, k) = \frac{1}{n^2} M^{-1} [nV_Y] M'^{-1} - \frac{1}{nk} M^{-1} V_d M'^{-1} + \frac{1}{(nk)^2} M^{-1} [nkV_Y] M'^{-1} \quad (5.66)$$

or:

$$V_e(n, k) = \frac{1}{nk} M^{-1} [(k+1)V_Y - V_d] M'^{-1} \quad (5.67)$$

The second particular case is when only one sample increment is extracted ($n = 1$). The sampling error variance is then:

$$\begin{aligned} V_e(1, k) &= M^{-1} V_Y M'^{-1} - \frac{1}{k} M^{-1} V_d M'^{-1} \\ &+ \frac{1}{k^2} M^{-1} \left[kV_Y + \sum_{i=1}^k (k-i)(V_{YY}(i) + V_{YY}(-i)) \right] M'^{-1} \\ &- \frac{1}{k} M^{-1} \left[\sum_{i=1}^k V_{YY}(l-i) \right] M'^{-1} - \frac{1}{k} M^{-1} \left[\sum_{i=1}^k V_{YY}(i-l) \right] M'^{-1} \end{aligned} \quad (5.68)$$

or:

$$\begin{aligned} V_e(1, k) &= M^{-1} \left[\left(1 + \frac{1}{k} \right) V_Y - \frac{1}{k} V_d \right. \\ &\left. + \frac{1}{k} \sum_{i=1}^k \left(\left(1 - \frac{i}{k} \right) (V_{YY}(i) + V_{YY}(-i)) - (V_{YY}(l-i) + V_{YY}(i-l)) \right) \right] M'^{-1} \end{aligned} \quad (5.69)$$

Finally, a last special case is when all the possible increments are extracted. Therefore $t_s = t_d$. This means that $k = 1$ and $l = 1$. In addition, the number of sample increments n tends to infinity. Consequently, the sampling error is zero, which is a logical result since the sample is the lot itself.

5.12. CORRELATION BETWEEN SAMPLES INCREMENTS

Using Equations 5.47, 5.48, and 5.63, Equation 5.49 can be rearranged as follows:

$$Z_e = \frac{1}{nM} \sum_{j=0}^{n-1} \left[Y^*(l+jk) + Y_d(l+jk) - \frac{1}{k} \sum_{i=1}^k Y^*(jk+i) \right] \quad (5.70)$$

or in a more compact form:

$$Z_e = \frac{1}{nM} \sum_{j=0}^{n-1} Y_{e1}(j) \quad (5.71)$$

where $Y_{e1}(j)$ is the Y error associated with the j th increments (incremental errors). It is the Y sampling error when taking a single increment ($n = 1$) in a time interval containing k discrete values of the streams property:

$$Y_{e1}(j) = Y(l+jk) - \frac{1}{k} \sum_{i=1}^k Y^*(jk+i) \quad (5.72)$$

Now the variance of e is:

$$V_e(n, k) = \frac{1}{n} V_e(1, k) + \frac{1}{n^2 M} \left[\sum_{i=1}^{n-1} (n-i) V_1(i) \right] \frac{1}{M'} + \frac{1}{n^2 M} \left[\sum_{i=1}^{n-1} (n-i) V_1(-i) \right] \frac{1}{M'} \quad (5.73)$$

where:

- $V_e(1, k)$ is given by Equation 5.68.
- $V_1(i)$ is the covariance between the Y incremental errors of increments separated by ik time intervals t_d .

The covariance term $V_1(i)$ is then:

$$V_1(i) = E\left[Y_{s1}(j)Y_{s1}'(j+i)\right] \quad (5.74)$$

which is independent of j because of the stationarity of the series. It equals:

$$V_1(i) = E\left[\left(Y(l+jk) - \frac{1}{k} \sum_{p=1}^k Y^*(jk+p)\right) \left(Y(l+(j+i)k) - \frac{1}{k} \sum_{q=1}^k Y^*((j+i)k+q)\right)'\right] \quad (5.75)$$

Hence, the covariance term is:

$$V_1(i) = \left[V_{YY}(ki) - \frac{1}{k} \sum_{j=1}^k V_{YY}(ki+j-l) - \frac{1}{k} \sum_{j=1}^k V_{YY}(ki-j+l) \right. \\ \left. + \frac{1}{k^2} \sum_{j=1}^k j V_{YY}[(i+1)k-j] + \frac{1}{k^2} \sum_{j=1}^{k-1} j V_{YY}((i-1)k+j) \right] \quad (5.76)$$

Finally, from Equation 5.73, the covariance term between the incremental errors is obtained:

$$V_e(n,k) - \frac{1}{n} V_e(1,k) = \frac{1}{n^2 M} \left[\sum_{i=1}^{n-1} (n-i) (V_{YY}(ki) + V_{YY}(-ki)) \right. \\ \left. - \frac{1}{k} \left(\sum_{i=1}^{n-1} \sum_{j=1}^k (n-i) (V_{YY}(ki+j-l) + V_{YY}(l-ki-j)) \right) \right. \\ \left. + \sum_{i=1}^{n-1} \sum_{j=1}^k (n-i) (V_{YY}(ki-j+l) + V_{YY}(l-ki+j)) \right) \\ \left. + \frac{1}{k^2} \left(\sum_{i=1}^{n-1} \sum_{j=1}^k j(n-i) (V_{YY}((i+1)k-j) + V_{YY}(j-(i+1)k)) \right) \right. \\ \left. + \sum_{i=1}^{n-1} \sum_{j=1}^{k-1} j(n-i) (V_{YY}((i-1)k+j) + V_{YY}((1-i)k-j)) \right) \Big] \frac{1}{M'} \quad (5.77)$$

This term represents the contribution of the correlation between the incremental errors. The same result could have been attained by directly calculating the difference $V_s(n,k) - (1/n)V_s(1,k)$ using Equations 5.64 and 5.68.

5.13. NUMERICAL EXAMPLES

To illustrate the above formulae, the two-stage flotation circuit of Figure 5.9 is considered (see Appendices A and H). The network for mass conservation equations is given in Figure 5.10. This network contains eleven branches and eight nodes. The basic parameters of the network are the component flowrates. Each branch or stream is supposed to be made up of four components. All the component flowrate variables are assumed to be measured.

The time constants of the mixers are respectively $2.5\Delta t$ and $4\Delta t$. The separation efficiencies are 0.7, 0.5, 0.4, and 0.2 for the four components of the first separator. They are 0.9, 0.7, 0.5, and 0.3 for the four components of the second separator.

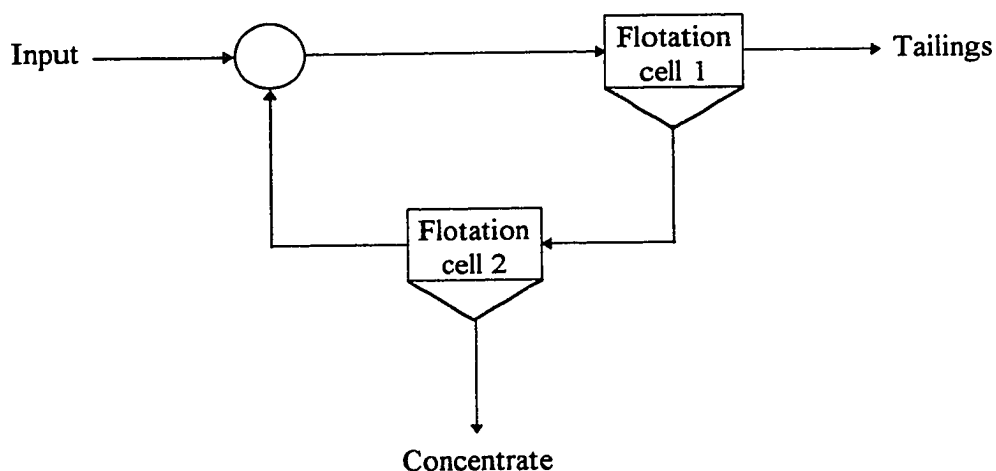


Figure 5.9. Flotation circuit in flowsheet form.

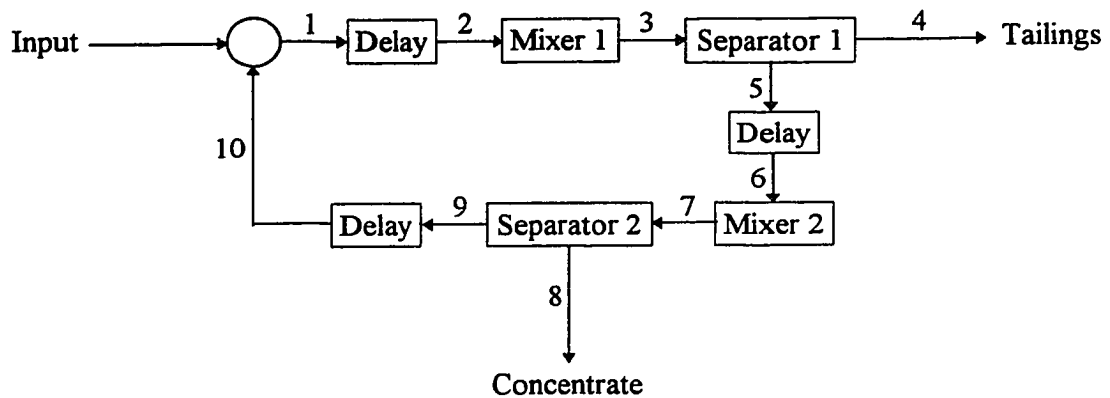


Figure 5.10. Flotation circuit in node form.

The input component flowrates are supposed to follow the same dynamic behavior. Three dynamic behaviors of the input component flowrates are considered. These dynamic behaviors are modeled using *AR* generators of stochastic process variables (Box and Jenkins, 1973; Abraham and Ledolter, 1983; Wei, 1990).

The corresponding autocorrelation functions are illustrated in Figure 5.11. This graph presents the autocorrelation functions of the simulated component flowrates following the *AR* selected models. The three autocorrelation functions are obtained through aperiodic first-order *AR* models: $(1 - 0.2B)x_i = \xi_i$, $(1 - 0.5B)x_i = \xi_i$, and $(1 - 0.9B)x_i = \xi_i$ respectively. These autocorrelation functions exhibit exponential decays of their correlations.

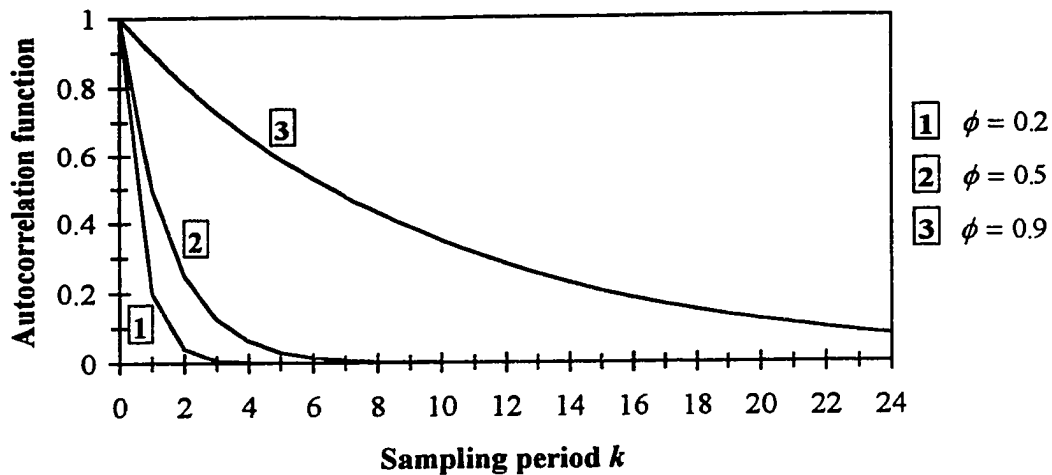


Figure 5.11. Autocorrelation functions of the *AR* series.

The flotation circuit of Figure 5.10 is used to find the mean composition of the process streams for a period T of an eight-hour shift. The discretization time t_d is selected as 10 minutes, so that $nk = 48$. For each stream, a first test is run to evaluate σ_d^2 : it consists of taking 10 or 20 samples during 10 minutes and analyzing them. Then, for each stream, samples are taken every 10 minutes during eight hours, and analyzed. This second test leads to σ^2 and $\rho(k)$. Now it is possible to calculate the sampling error variance as a function of the number of increments in such a way that nk equals 48.

The results can be presented by plotting $\sigma_e^2(n, k) / \sigma_e^2(1, 48)$ as a function of the number of sample increments n . Examples are given in Figures 5.12 through 5.14. In this section, it is supposed that the variance of the sample is aimed at being one tenth of the variance corresponding to a single increment. For instance, Figure 5.12 indicates that, before material balancing, the recommended numbers of sample increments are:

- 4 for *MA*(1): $(1 - 0.9B)x_i = \xi_i$.

- 7 for $MA(1)$: $(1 - 0.5B)x_i = \xi_i$.
- 8 for $MA(1)$: $(1 - 0.2B)x_i = \xi_i$.

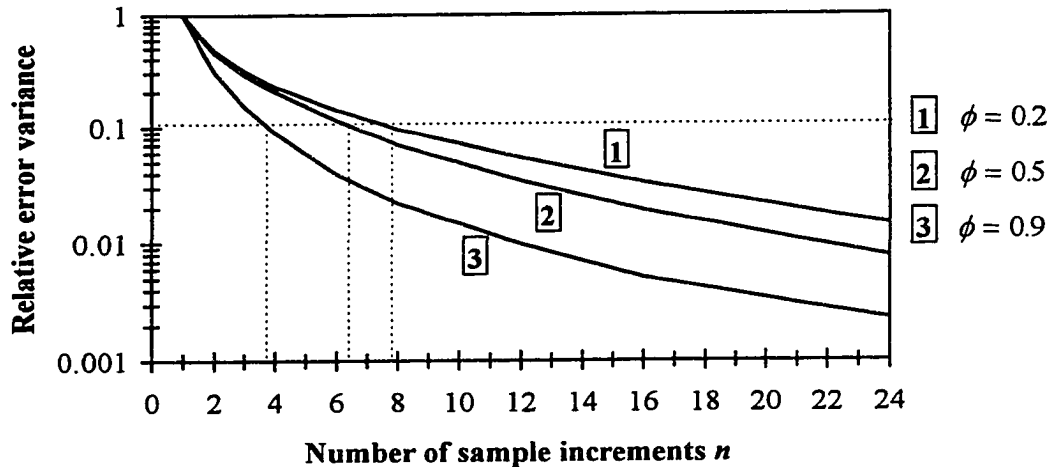


Figure 5.12. Ratio $\sigma_e^2(n, k)/\sigma_e^2(1, 48)$ of the AR series for the first component of the stream 1 before material balancing.

After material balancing, two cases are considered. The first case does not take into account the correlations between the components and between the streams. However, in the second case, these correlations are considered. Figure 5.13 corresponds to the first case. As a result, the recommended numbers of sample increments are:

- 3 for $MA(1)$: $(1 - 0.9B)x_i = \xi_i$.
- 5 for $MA(1)$: $(1 - 0.5B)x_i = \xi_i$.
- 6 for $MA(1)$: $(1 - 0.2B)x_i = \xi_i$.

that are lower than the respective numbers obtained before material balancing.

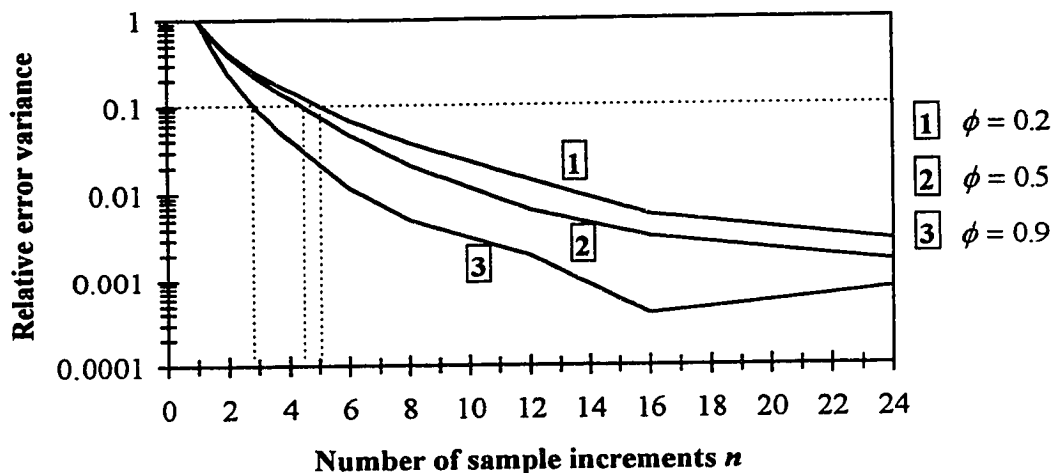


Figure 5.13. Ratio $\sigma_e^2(n, k)/\sigma_e^2(1, 48)$ of the AR series

for the first component of the stream 1 after material balancing without covariances.

Figure 5.14 corresponds to the second case. Consequently, the recommended numbers of sample increments are:

- 2 for $MA(1)$: $(1 - 0.9B)x_i = \xi_i$.
- 3 for $MA(1)$: $(1 - 0.5B)x_i = \xi_i$, and $MA(1)$: $(1 - 0.2B)x_i = \xi_i$.

that are lower than the respective numbers obtained before material balancing and after material balancing without taking into account the streams covariances.

Therefore, Figures 5.12 through 5.14 demonstrate that material balancing reduce the number of sample increments n necessary to evaluate the average stream composition according to the proposed accuracy. Besides, taking into account the correlations existing between the different components in a stream and between the different streams reduce the number of sample increments n .

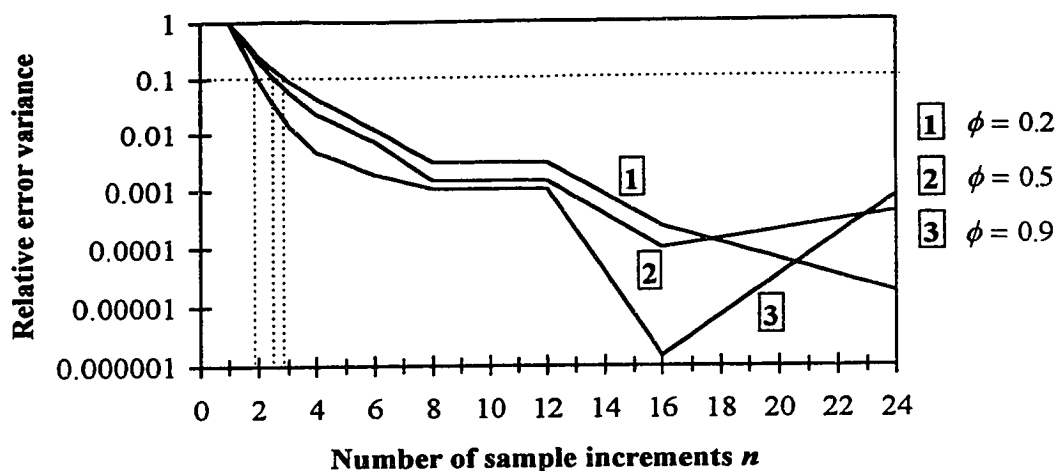


Figure 5.14. Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ of the *AR* series for the first component of the stream 1 after material balancing with covariances.

In the remaining illustrations, the input component flowrates are assumed to follow the *AR* model: $(1 - 0.9B)x_i = \xi_i$. Figure 5.15 illustrates the relative error variance, which is the ratio of $\sigma_e^2(n,k)$ to $\sigma_e^2(1,48)$ that is computed using uncorrected data, for the first component of the stream 1 (see Figure 5.10) before and after material balancing. This figure shows that the sampling error variances of the corrected data are lower than those of the experimental data. This justifies the use of material balance technique for upgrading data.

In addition, the sampling error variances of the corrected data, when the components and streams correlations are taken into account, are lower than when the correlation contribution is not considered in the interval of sample increments number]1,24[. This justifies considering the covariances between the components of a stream and between the streams of a mineral processing circuit.

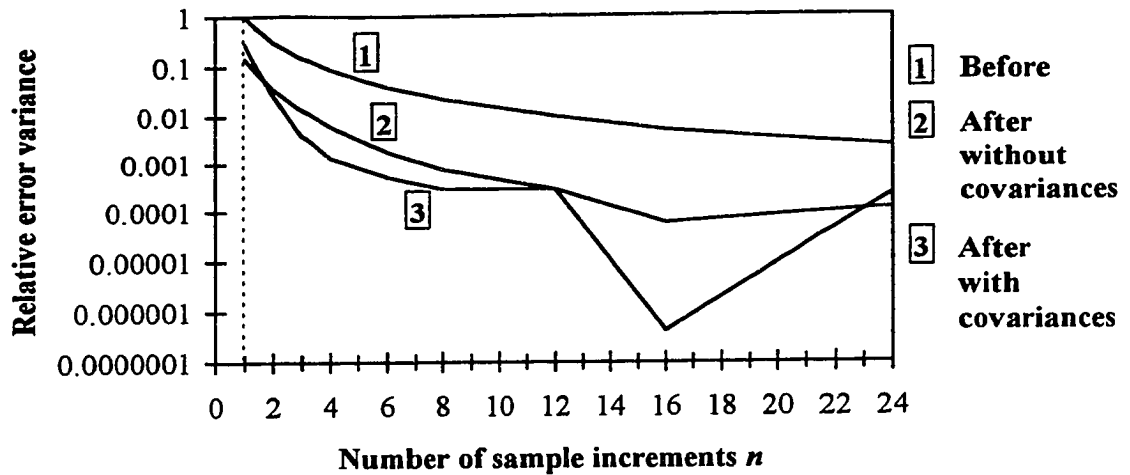


Figure 5.15. Relative error variance before and after material balancing for the first component of the stream 1.

Figure 5.16 exhibits the ratio $\sigma_s^2(n, k)/\sigma_s^2(1, 48)$ for the first component of the stream 1.

The graph indicates the recommended numbers of sample increments:

- 2 for the corrected data including the covariance terms.
- 3 for the corrected data without considering the covariance terms.
- 4 for the uncorrected data.

This means that the recommended number of sample increments decreases with material balancing and with including the covariance terms existing between the different components in a stream and between the different streams.

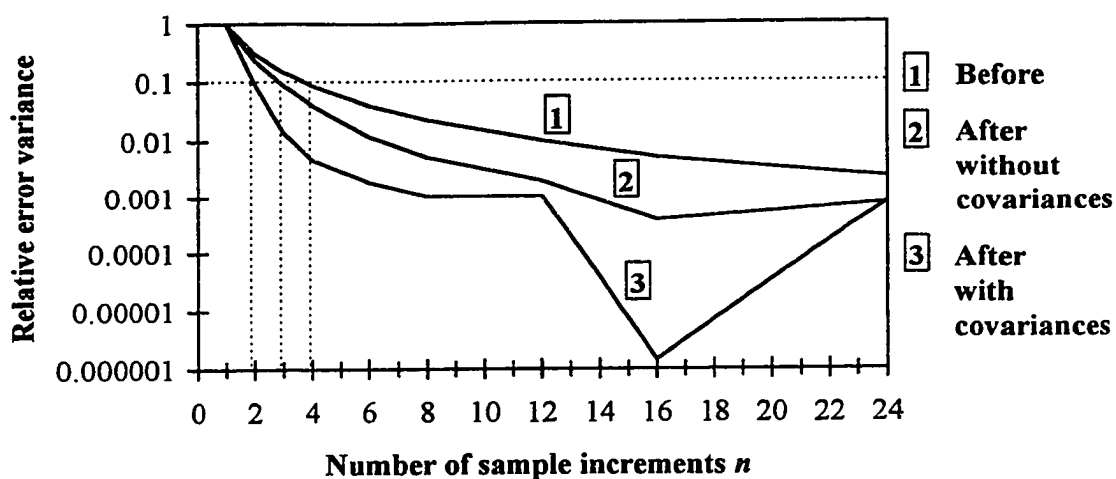


Figure 5.16. Ratio $\sigma_e^2(n,k)/\sigma_e^2(1,48)$ before and after material balancing for the first component of the stream 1.

These simulated data can also be used to illustrate the differences between the sampling error variances after material balancing with and without covariances. Figures 5.17 and 5.18 present the relative error:

$$e_R = \frac{\sigma_{e,2}^2 - \sigma_{e,1}^2}{\sigma_{e,1}^2} \quad (5.78)$$

for all the streams of the flotation circuit where $\sigma_{e,1}^2$ represents the sampling error variance after material balancing including the covariance terms, $\sigma_{e,2}^2$ the sampling error variance after material balancing without considering the covariance terms.

Figures 5.17 and 5.18 show that the relative error is different from zero in most sampling situations. The relative error can be as high as 267, which reveals the importance of including the covariance terms in a material balancing study.

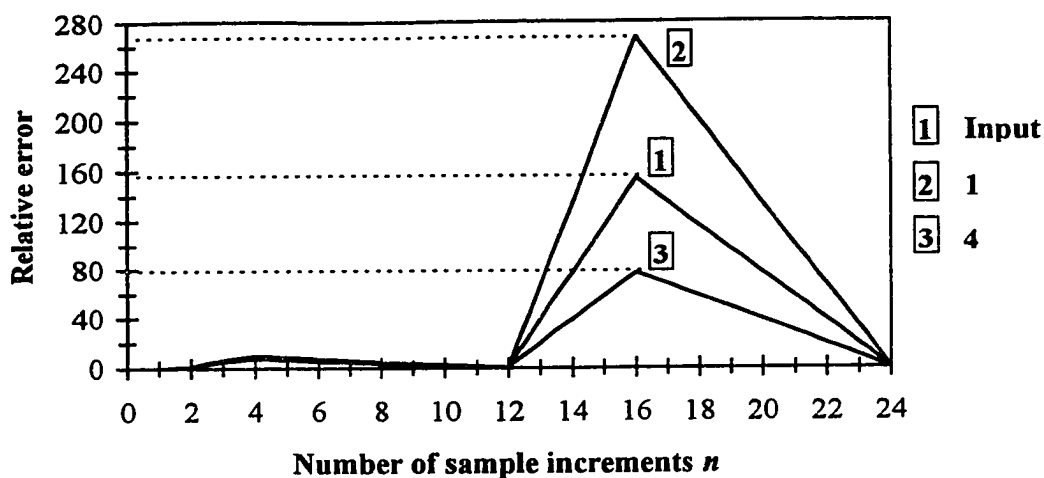


Figure 5.17. Relative error for the first component of the process streams: input, 1, and 4.

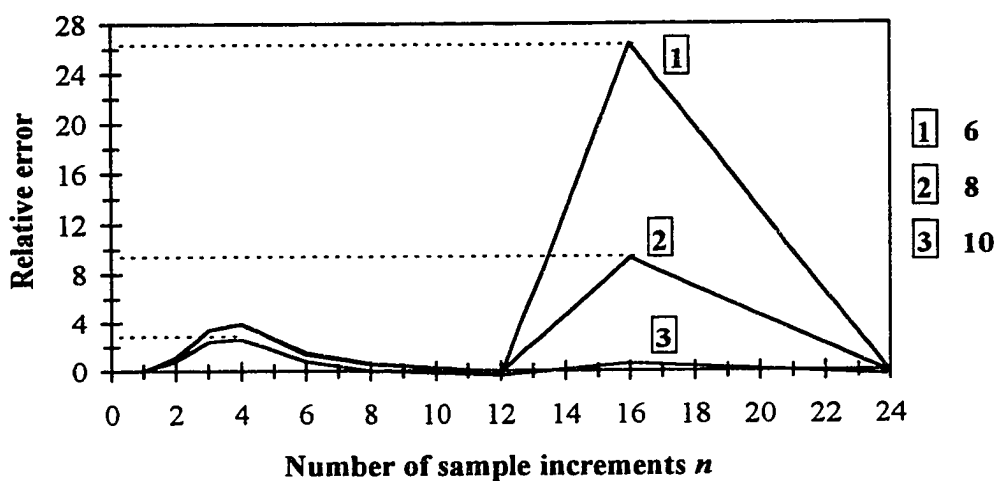


Figure 5.18. Relative error for the first component of the process streams: 6, 8, and 10.

Spollen et al. (1991) discuss also the practicalities of carrying out a material balance using simple nodal sensitivity analysis. They review the calculation of the errors involved and examine the concept of covariance. They conclude that the use of covariances in such

calculations is practically insignificant. However, their observation is based on the assumption that the data is uniformly distributed around a mean. Therefore, it cannot be generalized as indicated by the above results.

In addition, the study of Spollen et al. (1991) is based on variance evaluation for only one component in one node. However, in the case of many nodes and presence of many components, correlation exists between the components of a stream and between the streams of a process network. Indeed, the method proposed in this chapter shows the significance of considering the essential contribution of covariances in a mass balancing study.

CHAPTER 6

FUZZY EVALUATION OF SAMPLE RELIABILITY

6.1. FUZZY LOGIC

The notion of “fuzzy thinking” describes the thought process of a human being. According to Zadeh, the strength of human reasoning lies in the ability to grasp inexact concepts directly rather than approximating some precisely defined process by using a mathematical simulation based on true or false and on or off logic. Zadeh (1965) developed the methodology of fuzzy set theory and approximate reasoning. This process deals with the simulation of the human thought process, by introducing the concepts of vagueness and imprecise measure into the interpretation of information.

A fuzzy set has imprecise boundaries. Mathematically, a fuzzy set \tilde{A} is characterized by a membership function $\mu_{\tilde{A}}(a)$. This function assigns to each member element a a number in the closed interval 0 to 1 that represents the membership grade of a in \tilde{A} . Figure 6.1 illustrates an example of membership function of the fuzzy set “reliable sample”.

Fuzzy theory translates human concepts such as “reliable”, “high”, “low”, and “large” into forms that mathematically represent the vagueness of these concepts. As with ordinary sets, connections between fuzzy sets as unions, intersections and negations can also be defined.

6.2. SAMPLING CORRECTNESS CONDITIONS

The performance of a mineral processing plant is assessed regularly by stream sampling. The analysis of the resulting data is a crucial step in understanding the plant performance.

To control the sampling process efficiently, it is very important to minimize sampling errors and estimate them.

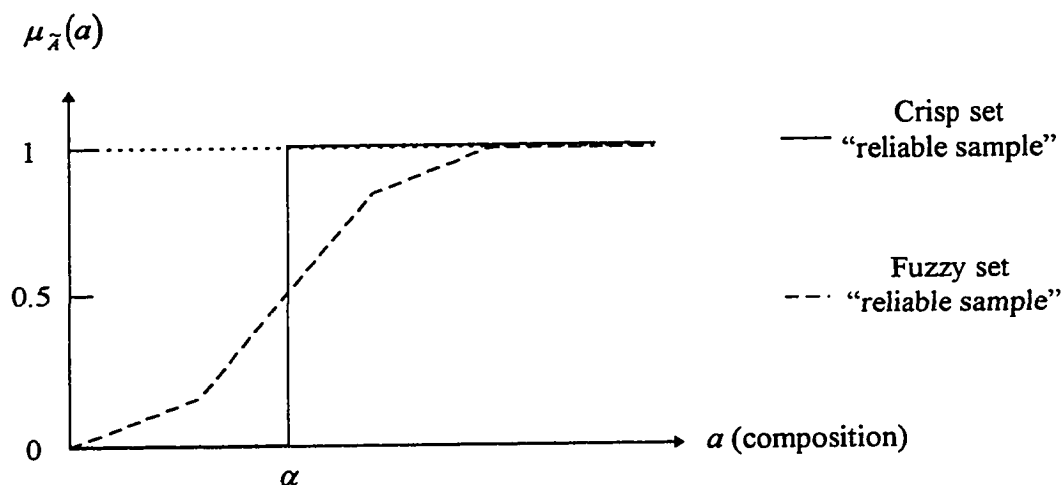


Figure 6.1. Example of crisp and fuzzy sets "reliable sample".

The factors influencing these errors are divided into two categories: material properties and cutter features. The material properties can be particle composition, mineral liberation, particle size distribution, and particle shape. These factors cannot be controlled due to material heterogeneity. However, the cutter features, such as cutter geometry, speed, layout, and path, are controllable. They can be responsible for gross errors if the cutter is not designed and used according to sampling correctness requirements. In this chapter, the stream is assumed to move at constant speed.

Fuzzy logic can be used to assess the sampling performance index that is the sample reliability. The latter is influenced by the cutter features describing the sampling conditions. Fuzzy modeling permits the use of expert's judgment by assigning numeric values that describe uncertainty of their statements. However, the conventional crisp modeling does not take advantage of it. Consequently, fuzzy logic improves the

understanding of the sampling process, which results in efficient control of the plant operation.

Fuzzy logic, with its intuitive nature and closeness to the natural languages, offers significant advantages over traditional approaches in the appraisal of sampling conditions. Fuzzy logic allows sampling situations to be described and processed in linguistic terms such as *very reliable*, *reliable*, *adequate*, *doubtful*, and *very doubtful*. These fuzzy sets lead to the value of the sampling performance index.

A value of sampling performance index is attached to each sample. It can be utilized as a weight in further calculations of average stream compositions for example. Besides, it is the basis for decision-making concerning sampling strategy and data reconciliation.

6.2.1. Cutter geometry

The condition of cutter geometry correctness is that both boundaries delimiting the extended increment should be perfectly superposable by a translation parallel to the extension axis. This means that the distance l between both boundaries should remain strictly constant as shown in Figure 6.2. The cutter geometry is correct if the cutter is rectangular (Pitard, 1993).

Therefore, the increment actually delimited across the stream is an inclined parallelogram (Figure 6.2). Under those conditions, all fragments of the stream are submitted to the selection during the same length of time. This ensures that the sampling probability is uniform for all fragments, irrespective of their position in the stream. However, a correctly designed cutter may slowly degenerate into an incorrect cutter for various reasons (Pitard, 1993).

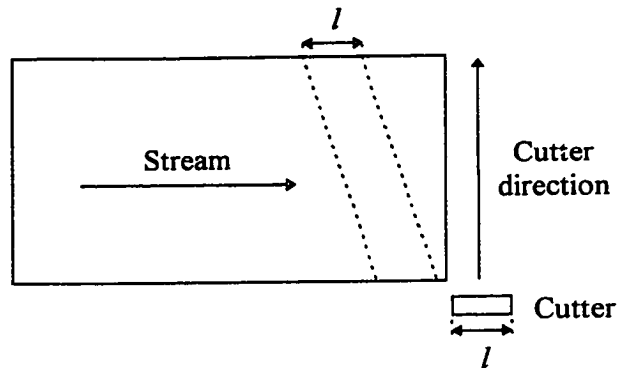


Figure 6.2. Correct cutter geometry.

First, if the cutter is built with fragile material, it can be easily damaged by the violent impact of large fragments. Often, cutters are selected because they are cheap. Consequently, they are used under conditions for which they are not designed. A cutter that can handle 2-in. coal fragments may not be able to handle 2-in. iron ore. The cutter becomes vaulted in the middle and the condition for correct increment delimitation is not respected. The same incorrectness can be the result of no inspection and no maintenance of a correct cutter over weeks, months, and sometimes years.

Secondly, the cutter can become obstructed by sticky materials such as fines. This is observed in flotation plants where samplers are not regularly cleaned and maintained. Sticky materials accumulate around the edges, partially closing the cutter opening. Preventive maintenance and periodic cleaning employing a water hose is an easy solution.

6.2.2. Cutter speed

The conditions of cutter speed correctness are (Pitard, 1993):

- The cutter speed should be constant during the time necessary to cross the stream.

- The cutter speed should remain uniform during the collection of all increments.

The cutters follow, most of the time, the same sequence:

- Starting point: A timer sets the cutter into motion from an idle position that is supposed to be outside the stream.
- Crossing the stream: The cutter proceeds with the extraction of the actual increment.
- Stopping point: A timer turns off the power, then the cutter stops on its new idle position that is supposed to be outside the stream.
- Repetition: This sequence is repeated as needed following a systematic sampling mode.

Pitard (1993) points out that the correctness of these cutters highly depends on the nature of their driving systems. The following driving systems are used:

- Electric.
- Hydraulic.
- Pneumatic.
- Magnetic.
- Manual.

He adds that cutters equipped with an electric motor generally coupled to a reduction gear are very common. They are recommended for all sampling phases. They are cheap, easy to maintain, simple, and very reliable. According to Pitard (1993), the electric drive is the only drive likely to achieve constant velocity of the cutter across a flowing stream of particulate material.

As indicated by Pitard (1993), for this condition to be fulfilled, four other conditions must be put into effect. First, it is necessary to make certain that the system is powerful enough to prevent any slowing down of the cutter as it crosses the stream.

Secondly, the idle positions of the cutter should be far enough from the stream for the cutter to have enough time to attain its cruising speed prior to entering the flowing stream. This distance may vary from 0.2 m for small pulp cutter to 3 m for cutters handling streams with very high flowrates, for example 5,000 to 10,000 tons per hour. The only way this distance could be decreased is by using a stronger electric motor coupled to a reduction gear giving more mechanical advantage.

Thirdly, the speed reduction gears should be designed and constructed to prevent jolting and bouncing sporadic motions. The cutter traveling at constant speed should also travel as smoothly as possible.

Finally, the constant speed of the cutter across the stream can be badly affected by poor or nonexistent maintenance. Regular and well-planned maintenance of the cutters is an absolute must.

6.2.3. Cutter layout

Idle positions of the cutter should be far away from the stream to permit the motor to accelerate to its cruising speed before the cutter attains the stream. The stream should fall exactly in the center of the cutter. The cutter should be able to cross the entire stream. Figure 6.3 illustrates the vertical distance u between the liberation point of the stream and the plane described by the cutter edges during sampling. This distance should be minimized but still at least three times the diameter d of the largest fragments (Pitard, 1993).

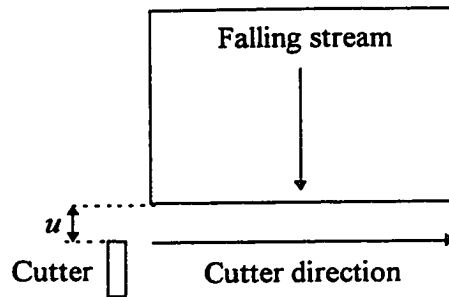


Figure 6.3. Correct cutter layout.

Respecting the minimum distance u prevents accidental obstruction between the cutter and the point of discharge of the stream. In the case of fine particles, this distance should also be minimized but not less than one centimeter. Hence, the following rule of thumb is recommended (Pitard, 1993):

$$u = 3d + 1 \quad (6.1)$$

where u is the minimum distance in centimeters.

A very common faulty layout results when the stream thickness is larger than the cutter length. This occurs when the cutter is undersized for budget reasons. Under such conditions it is probable to see part of the stream fall outside the cutter trajectory. One faulty design is shown in Figure 6.4. The shaded area illustrates the material belonging to the increment that is not collected. For the same reasons, the stream may partially enter the cutter all the time because one idle position is too close to the stream as shown in Figure 6.5. The shaded area illustrates the material that is collected but does not belong to the increment.

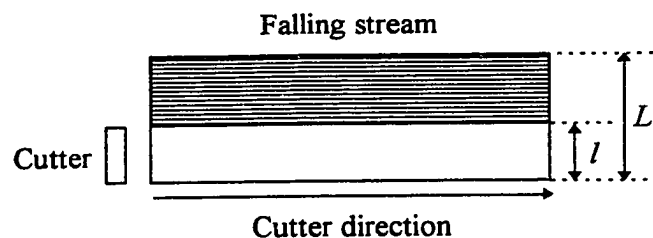


Figure 6.4. Incorrect cutter layout with stream thickness larger than cutter length.

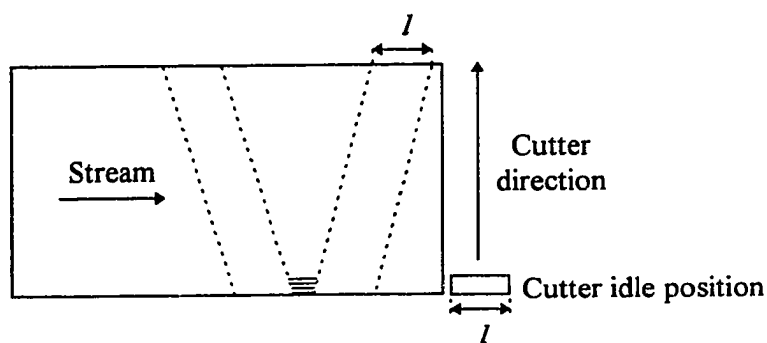


Figure 6.5. Incorrect cutter layout with cutter idle position close to the stream.

Another incorrect layout is noticed when the cutter reverses its motion before it attains an idle position completely outside the flowing stream. In such a case the sampling ratio is smaller on the side where the cutter reverses its trajectory as illustrated in Figure 6.6. The shaded area presents the part of the stream that is not correctly sampled.

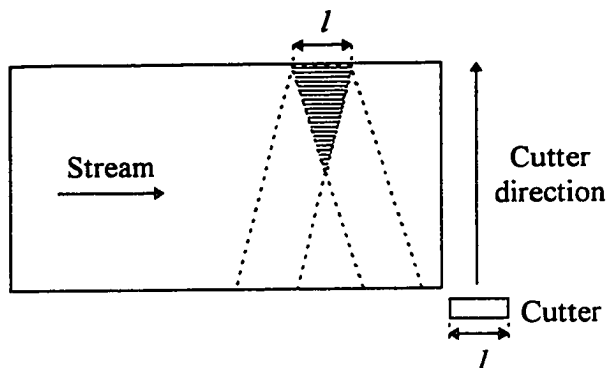


Figure 6.6. Incorrect cutter layout with cutter reversing its motion too soon.

6.2.4. Cutter path

The condition of cutter path correctness is that the cutter path is perpendicular to the direction of the stream. For instance, if the cutter path is parallel to the stream direction, the sampling operation is considered to be performed by the scheme: taking part of the stream all of the time. This is always incorrect. This situation is illustrated in Figure 6.7. The shaded area represents the material collected by the cutter.

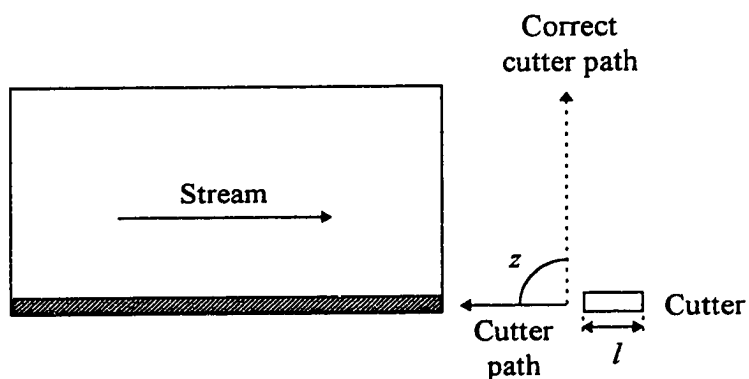


Figure 6.7. Incorrect cutter path with taking part of the stream all of the time.

6.3. FUZZY SETS

6.3.1. Cutter geometry

Here, the cutter is assumed to be resistant to the violent impact of large fragments. In this way, the cutter geometry becomes incorrect only if it is obstructed by sticky materials such as fines. Figure 6.8 illustrates a possible membership function of the fuzzy set “very reliable cutter geometry”:

$$\tilde{A} = \text{“very reliable cutter geometry”} = \{(v, \mu_{\tilde{A}}(v)) | v \in [0,1]\} \quad (6.2)$$

In this expression, v denotes the area fraction of the cutter opening given by:

$$v = A_E / A \quad (6.3)$$

where A_E represents the effective cutter area, and A the total cutter area as displayed in Figure 6.9. The shaded area shows the sticky material.

The corresponding membership function equals:

$$\mu_{\tilde{A}}(v) = v, \quad 0 \leq v \leq 1 \quad (6.4)$$

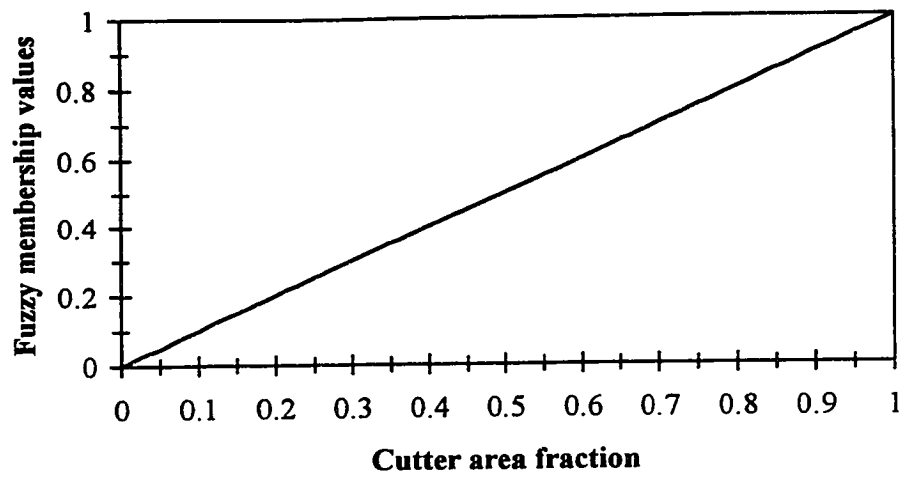


Figure 6.8. Definition of fuzzy set “very reliable cutter geometry”.

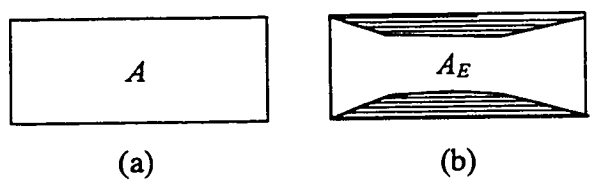


Figure 6.9. Cutter opening before (a) and after (b) obstruction by sticky materials.

6.3.2. Cutter speed

The faulty cutter speed results when the cutter speed is not constant during sampling. Hence, Figure 6.10 exhibits a possible membership function of the fuzzy set “very reliable cutter speed”:

$$\tilde{B} = \text{“very reliable cutter speed”} = \{(x, \mu_{\tilde{B}}(x)) | x \in [0,1]\} \tag{6.5}$$

where x is the constant-speed length fraction of the cutter given by:

$$x = L_A / L_T \quad (6.6)$$

where L_A represents the actual cutter path length that is perpendicular to the stream direction and equal to the length of constant speed, and L_T the total cutter path length that is perpendicular to the stream direction. The corresponding membership function is:

$$\mu_{\bar{B}}(x) = x, \quad 0 \leq x \leq 1 \quad (6.7)$$

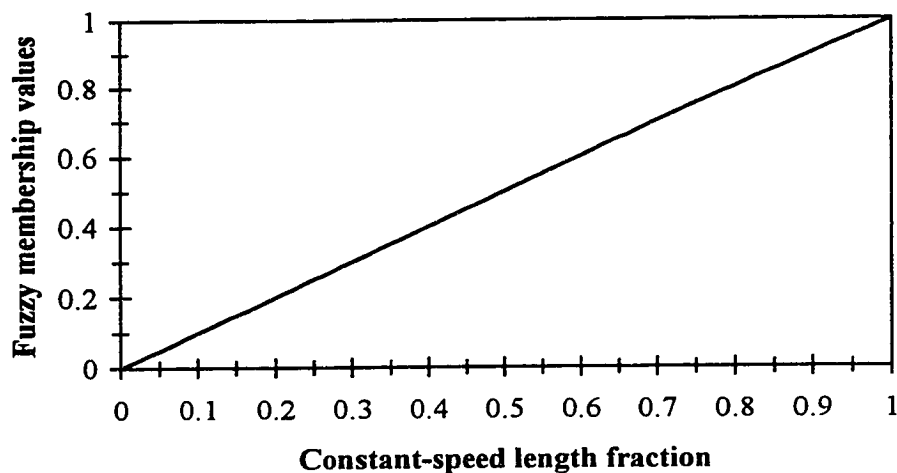


Figure 6.10. Definition of fuzzy set “very reliable cutter speed”.

6.3.3. Cutter layout

Here, the minimum distance u that prevents accidental obstruction between the cutter and the point of stream discharge is assumed to be respected. In addition, the cutter is presumed to have its idle positions completely outside the flowing stream. Thus, the faulty

layout results when the stream thickness is larger than the cutter length. Indeed, Figure 6.11 displays a possible membership function of the fuzzy set “very reliable cutter layout”:

$$\tilde{C} = \text{“very reliable cutter layout”} = \{(y, \mu_{\tilde{C}}(y)) \mid y \in [0, +\infty[\} \quad (6.8)$$

where y is the cutter length fraction given by:

$$y = L/l \quad (6.9)$$

where L and l represent the stream thickness and the cutter length respectively (see Figure 6.4). The related membership function is written as follows:

$$\mu_{\tilde{C}}(y) = \begin{cases} 1, & 0 \leq y \leq 1 \\ 1/y, & y \geq 1 \end{cases} \quad (6.10)$$

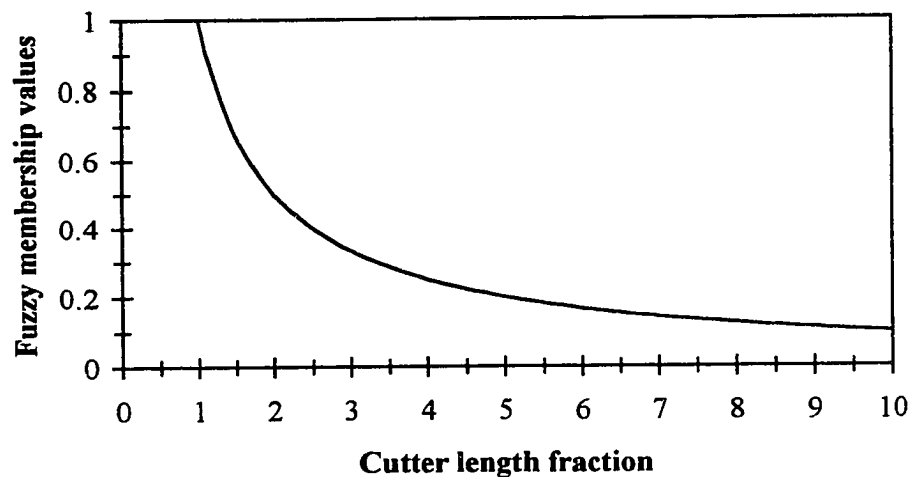


Figure 6.11. Definition of fuzzy set “very reliable cutter layout”.

6.3.4. Cutter path

The incorrect cutter path results when the cutter deviates from the correct route that is perpendicular to the stream direction. Hence, Figure 6.12 shows a possible membership function of the fuzzy set “very reliable cutter direction”:

$$\tilde{D} = \text{“very reliable cutter direction”} = \{(z, \mu_{\tilde{D}}(z)) | z \in [0^\circ, 90^\circ]\} \quad (6.11)$$

where z is the cutter deviation angle, in degrees, from the correct cutter path that is perpendicular to the stream (see Figure 6.7). The related membership function is described by:

$$\mu_{\tilde{D}}(z) = \cos(z), \quad 0 \leq z \leq 90^\circ \quad (6.12)$$

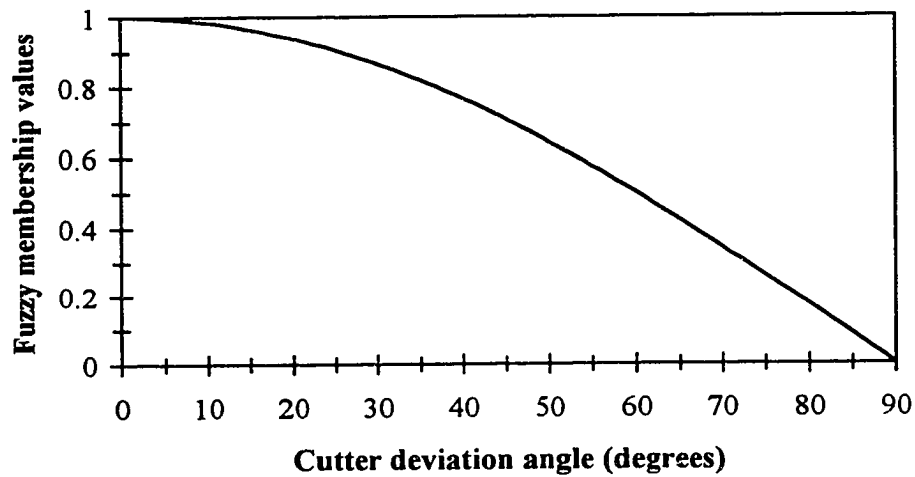


Figure 6.12. Definition of fuzzy set “very reliable sampling direction”.

6.4. FUZZY EVALUATION OF SAMPLE RELIABILITY

The sampling condition of mineral processing streams is influenced by the cutter factors:

- Geometry.
- Speed.
- Layout.
- Path.

In sampling condition assessment a problem commonly encountered is the need to evaluate the compound influence of the cutter factors. This situation is essentially a multi-criterion decision making problem.

The cutter factors are put into the set C :

$$C = [\textit{cutter geometry}, \textit{cutter speed}, \textit{cutter layout}, \textit{cutter path}] \quad (6.13)$$

The sample reliability is described by five indexes, namely *very reliable*, *reliable*, *adequate*, *doubtful*, and *very doubtful*. These indexes are represented by index set R :

$$R = [\textit{very reliable}, \textit{reliable}, \textit{adequate}, \textit{doubtful}, \textit{very doubtful}] = [r_1, r_2, r_3, r_4, r_5] \quad (6.14)$$

Thus, the fuzzy relation \tilde{R} is established:

$$\tilde{R} = "c_i \textit{ is } r_j" \quad (6.15)$$

where $i = 1, \dots, 4$ and $j = 1, \dots, 5$.

Hence, \tilde{R} is defined by the following matrix:

$$\tilde{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} \end{bmatrix} \quad (6.16)$$

where $r_{ij} \in [0,1]$ is the membership grade of the pair (c_i, r_j) given by:

$$r_{ij} = \mu_{\tilde{R}}(c_i, r_j) \quad (6.17)$$

Now, the degree of cutter factor importance is represented by the fuzzy set \tilde{I} :

$$\tilde{I} = [i_1, i_2, i_3, i_4] \quad (6.18)$$

where i_i is the membership grade of factor c_i . The quantifying scales of the perceived importance of each cutter factor can be determined as shown in Figure 6.13.

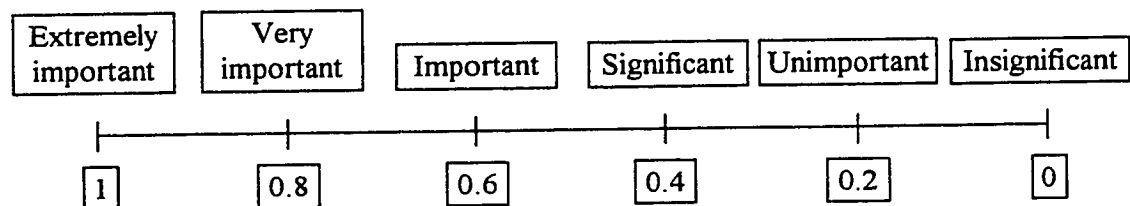


Figure 6.13. Factor importance ruler.

Consequently, the overall assessment result is obtained by:

$$\tilde{W} = \tilde{I} \circ \tilde{R} = [i_1, i_2, i_3, i_4] \circ \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} \end{bmatrix} = [w_1, w_2, w_3, w_4, w_5] \quad (6.19)$$

where \circ stands for maximum-minimum composition (Zimmermann, 1991):

$$w_i = \max[\min(i_1, r_{1i}), \min(i_2, r_{2i}), \min(i_3, r_{3i}), \min(i_4, r_{4i})], \quad i = 1, \dots, 5 \quad (6.20)$$

where max and min symbolize maximum and minimum respectively.

w_i is the membership grade of sample reliability index i . The sample reliability is described by the index i with the highest value w_i . As a result, the degree of sample reliability is evaluated by applying fuzzy logic.

From the previous development, the fuzzy logic appears to be a strong tool for estimating the sample reliability. This estimation is the basis for decision-making concerning sampling strategy and data reconciliation. The judgment of plant operators and sampling experts is used to define all the fuzzy sets.

6.5. FUZZY CORRECTION OF ERRONEOUS COMPOSITION VALUE

Figure 6.14 exhibits a possible membership function of the fuzzy set “very reliable sample”:

$$\tilde{S} = \text{“very reliable sample”} = \{(a, \mu_{\tilde{S}}(a)) | a \in [0,1]\} \quad (6.21)$$

where a is the composition of a stream component and:

$$\mu_{\tilde{S}}(a) = \begin{cases} 0, & 0 \leq a \leq \alpha \\ \frac{a - \alpha}{\beta - \alpha}, & \alpha \leq a \leq \beta \\ 0, & a \geq \beta \end{cases} \quad (6.22)$$

$\mu_{\tilde{S}}(a)$ is also equal to w_1 :

$$\mu_{\tilde{S}}(a) = w_1 \quad (6.23)$$

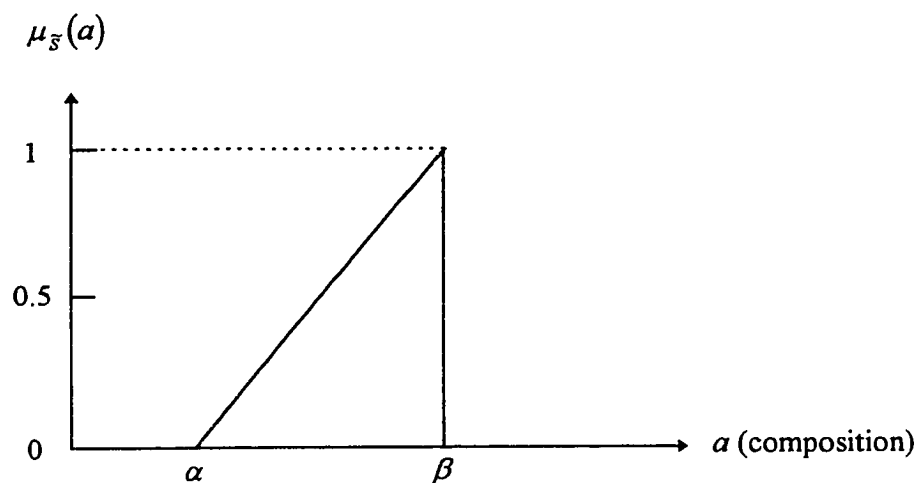


Figure 6.14. Definition of fuzzy set “very reliable sample”.

\tilde{S} is assumed to be a triangular fuzzy number:

$$\tilde{S} = (\alpha, \beta, \gamma) \quad (6.24)$$

where:

$$\beta = \gamma \quad (6.25)$$

Hence, it is written as follows:

$$\tilde{S} = (\alpha, \beta) \quad (6.26)$$

Let α be the measured value of the sample composition. This value is erroneous. The correct sample composition β is expressed as follows:

$$\beta = \alpha + \frac{\alpha - \alpha}{w_1} \quad (6.27)$$

6.6. NUMERICAL EXAMPLE

To illustrate the above formulae, a stream sampling situation is considered. Its membership grade matrix \tilde{R} is given by:

$$\tilde{R} = \begin{bmatrix} 0.3 & 0.5 & 0.2 & 0 & 0 \\ 0.2 & 0.4 & 0.3 & 0.1 & 0 \\ 0.2 & 0.3 & 0.2 & 0.2 & 0.1 \\ 0 & 0.3 & 0.4 & 0.3 & 0 \end{bmatrix} \quad (6.28)$$

In this case, all the cutter factors are extremely important. Therefore, the factor importance set \tilde{I} is obtained:

$$\tilde{I} = [1,1,1,1] \quad (6.29)$$

Then the following judgment result is acquired:

$$\tilde{W} = \tilde{I} \circ \tilde{R} = [0.3,0.5,0.4,0.3,0.1] \quad (6.30)$$

From set \tilde{W} , it appears that the membership grade for index 2, *reliable*, is the maximum. This means that the sampling condition is reliable. Now, the correct composition value can be estimated based on the previous evaluation of sample reliability. If the values of the composition α and the erroneous composition a equal 0.2 and 0.23 respectively, following Equation 6.27, the correct sample composition is:

$$\beta = 0.2 + \frac{0.23 - 0.2}{0.3} = 0.3 \quad (6.31)$$

The resulting triangular fuzzy number “very reliable sample” is:

$$\tilde{S} = (0.2, 0.3) \quad (6.32)$$

and shown in Figure 6.15.

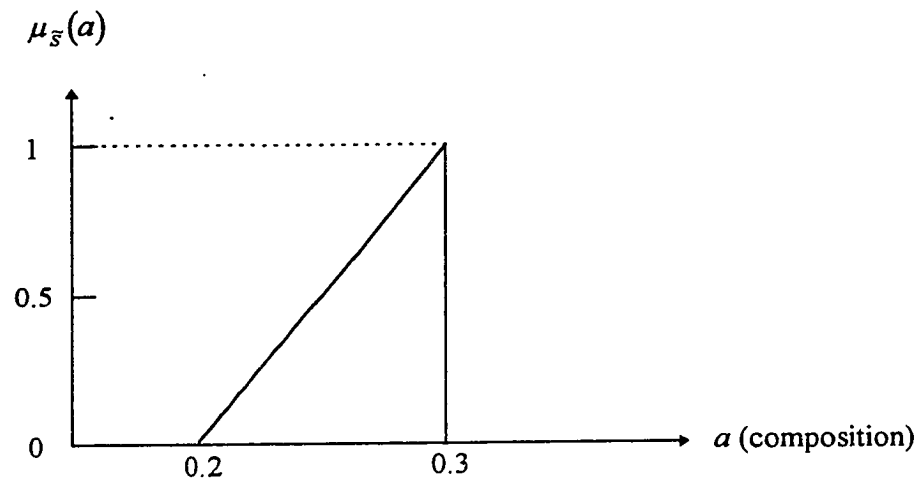


Figure 6.15. Definition of fuzzy set “very reliable sample”.

CHAPTER 7
EXPERT SYSTEMS SAMPLING CORRECTNESS INSPECTOR
AND SAMPLING ERROR EVALUATOR

7.1. EXPERT SYSTEMS

In the study of artificial intelligence, efforts are made to emulate some human activities such as movement, vision, or reasoning. Much research has been done in the field of artificial intelligence in the last decades. One of the most important fields of research in artificial intelligence has been that of knowledge engineering. Knowledge-based systems have been the outcome of this human endeavor to make machines intelligent and to make the computers participate in the decision making process rather than being used as a mere tool for numerical computation (Hayes-Roth et al., 1983).

Knowledge-based systems are computer systems that emulate a defined part of the reasoning capability of an expert (Laguitton and Leung, 1989; Bearman and Milne, 1992). For this reason, knowledge-based systems are also known as expert systems. Besides, they are identified as knowledge systems and rule-based systems. These systems may completely fulfill functions that normally require human expertise, or may play the role of assistants to human decision makers.

Jackson (1990) identifies the features of an expert system as follows:

- 1) An expert system deals with subject matter of reasonable complexity that normally requires a considerable amount of human expertise.
- 2) An expert system must exhibit high performance in terms of speed and reliability in order to be a useful tool.

- 3) An expert system must be capable of explaining and justifying solutions and recommendations in order to convince the user that its reasoning is correct.

Mishkoff (1988) distinguishes between a database system and an expert system. He indicates that a database program retrieves facts that are stored while an expert system uses reasoning to draw conclusions from stored facts. Furthermore, he points out that an expert system contains knowledge about a particular field to assist human experts or provide information to people who do not have access to an expert in the particular field.

Expert systems assemble declarative and procedural knowledge in one system. Such systems make conclusions whereas database systems or simulation programs generate calculated answers. The basic elements of the expert system are knowledge elements as distinct from data elements (Mishkoff, 1988).

An expert system comprises three main parts (see Figure 7.1):

- 1) A knowledge base containing the expertise of one or more human experts in a particular domain.
- 2) An inference engine for manipulating the knowledge in order to solve problems.
- 3) A user interface for acquiring and distributing information.

The knowledge base has the characteristics of a database and a program (Arbour, 1992). Usually a knowledge base is made up of facts and rules relating to particular tasks or applications, that are, in this chapter, sampling correctness inspection and sampling error evaluation. A fact is a knowledge base entry most generally of the form:

$$\textit{expression} = \textit{value}. \tag{7.1}$$

which provides a value for expression. The fact may be qualified by a certainty factor that indicates the degree of certainty with which the fact is believed.

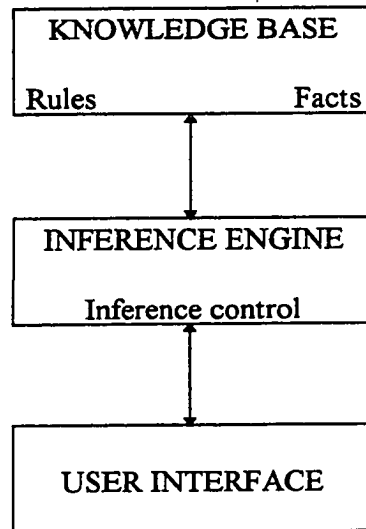


Figure 7.1. Components of an expert system.

A rule is a knowledge base entry of the form:

if *premise* then *conclusion*. (7.2)

where *premise* is generally a proposition of the form exhibited by Equation 7.1 or a combination of such propositions joined by the Boolean connectives *and*, *or*, or *not*.

The inference engine is the component of an expert system that acts upon the rules and other data. The inference engine controls the way in which the rules are accessed, either in forward or backward chaining. An inference engine primarily uses a backward-chaining reasoning process to reach conclusions. Thus, it works with input data and rules in the

knowledge base to deduce facts or conclusions. When conclusions matching a specified pattern are made, a special set of high-priority goals can be activated thereby enabling a forward-chaining capability.

At the heart of an expert system is the mechanism for finding, or seeking, values for expressions by methodically considering previously stored conclusions, relevant knowledge base entries, and user-supplied information. The inference engine is activated when a consultation is begun. It then seeks the values for expressions. When done, the inference engine displays the values for those expressions, along with a justification describing how those conclusions are reached.

The user interface controls the way in which the user interacts with the expert system. Some are interactive, others are not, and some use graphical interfaces and others are character-based. The degree to which the system is interactive or uses graphical interfaces depends to some extent on the problem area with which the system deals. The user interface provides varying opportunities for the user to follow the inference process and to provide explanation features (Arbour, 1992).

The comprehension of expert systems in an industry depends on the nature of that industry. Enormous interest in expert systems has been shown in medicine though ethical and legal barriers to implementation have appeared. Mineral processing has shown an interest similar to that shown by the manufacturing industries: a prudent investigation followed by spot implementations of the technology in disparate places (Jämsä-Jounela, 1990; Reuter and Van Deventer, 1990; Reuter and Van Deventer, 1992; Poirier and Meech, 1993; Farzanegan et al., 1995; Sastry, 1997; Golsan et al., 1998).

Up to 1983, an expert system consisted in interviewing an expert for his expertise and encoding it directly. Each rule corresponded to the problem solution of an expert in a

given situation. Knowledge and know-how were mixed. This led to more or less anarchic system development and especially weak systems difficult to handle (Gueniffey et al., 1995).

Since 1983, second generation expert systems, where knowledge is encoded in a declarative way, have been designed (Gueniffey et al., 1995). Thus, specific area models have been used for classification, design, diagnosis, and control. Three knowledge levels have to be distinguished (Newell, 1991):

- 1) Tasks.
- 2) Models.
- 3) Methods.

Two knowledge sources are consulted:

- 1) A specialist whose knowledge is public and proved by publications.
- 2) An expert whose knowledge is private from a personal experience, transferable with difficulty.

The difficulty of knowledge acquisition consists in the interpretation and modeling problem. The collaboration of domain specialists simplifies this stage. Combining this progress with the object formalism use permits to design new systems that are reusing developed modules for other applications. New architectures allow the generation of multi-expert systems using heterogeneous expertise. The graphical tools simplify the accomplishment of convivial user interfaces to make the communication with the user easier.

7.2. EXPERT SYSTEM EVALUATION

Expert systems are entering a critical stage as interest spreads from university research to practical applications (Hayes-Roth et al., 1983; Harmon and King, 1986; Waterman, 1986). It is often difficult, in analyzing potential expert systems applications, to assess the match between the tasks involved in an application and the current state of the art in expert systems development paradigms and techniques (Laufmann et al., 1990).

The artificial intelligence literature contains a number of important references related to the selection of appropriate domains or problems for the application of expert systems technology (Laufmann et al., 1990). These typically comprise a series of questions designed to elucidate problem or task characteristics that can provide important clues about the applicability of expert systems techniques to specific problems.

Slagle and Wick (1988) provide a technique whereby knowledge engineers can evaluate applications along essential and desirable feature lines. Prerau (1985) groups salient issues into topical lists, with a brief discussion of each question, such that an organization selecting among several potential expert system applications may choose the one that offers the greatest potential for success. Kline and Dolins (1989) present a series of probing questions to assist expert systems developers in relating problem characteristics to specific expert systems techniques, representations, and methodologies.

Laufmann et al. (1990) builds on that previous work and extends it by identifying common principles related to selecting appropriate expert systems applications. They have exploited and organized these principles into a methodology for qualitatively analyzing the applicability of expert systems technologies to specific problems within specific organizational environments.

This methodology is divided into two levels of analysis. The first level is oriented toward those who are less familiar with the mechanical environment within which the application will be developed and fielded. The second level probes more deeply into technical issues, and is oriented toward knowledge engineers and technical project managers. The methodology described by Laufmann et al. (1990) assumes a primarily end-product orientation to the expert systems application analysis. Any ultimate success or failure is therefore defined in terms that can be reduced to production end use, such as productivity, speed, efficiency, accuracy, consistency, and related measures.

The goal of the expert systems development effort is assumed to be a viable system for production use. The goal of both analysis levels is to examine the following factors:

- 1) Definition of success or goals.
- 2) Appropriateness of expert systems technology.
- 3) Availability of resources (for example knowledge, funding for hardware and software tools and related support).
- 4) Nontechnical considerations (for example problem importance, and economic, ethical, and legal considerations of the project).

Developers should try to define the task before beginning the implementation, but should always allow for changes in specifications (Smith, 1984).

In this chapter two expert systems are proposed. The first one is intended to inspect the correctness of sampling operations in a mineral processing plant. The second one is destined for the evaluation of sampling errors in a mineral processing plant. The existence and necessity of human knowledge suggest an expert systems solution in both cases. The appropriateness of an expert systems solution is amply justified by the need to use human expertise in sampling of mineral processing streams (Gy, 1965, 1979, 1988, 1992; Merks,

1985; Holmes, 1991; Ketata, 1991; Pitard, 1992, 1993; Hodouin and Ketata, 1994; Ketata and Rockwell, 1998).

The software M.4 (Teknowledge, 1993, 1995, 1996) is chosen to develop and use these expert systems. M.4 is a powerful and versatile PC-based knowledge system software tool. It is a development tool that makes it relatively easy to develop an interactive expert system. M.4 has been used since 1984 by hundreds of corporations and at many universities to develop knowledge system applications. For instance, the University of Ottawa reviewed expert system shells available on the market and selected M.4 as the best overall package (Hansen, 1996).

Implemented in the C programming language, M.4 can be used quickly and effectively by non-specialist computer programmers with no prior experience in knowledge system technology. Besides, M.4 is powerful enough to support the development of complex, commercial quality knowledge systems of significant value.

Expert systems built with M.4 are designed around a knowledge base and an inference engine. The knowledge base is composed of facts and rules relating to a particular task or application. The inference engine performs the reasoning process to solve specific problems in that application area.

Using a text editor that can output ASCII text, a programmer creates a text file of facts and rules that can be applied to those facts in certain situations. When loaded into M.4, this text file becomes the knowledge base that M.4 uses to solve problems with its inference engine. The accuracy and consistency of the knowledge base can be verified by observing the dynamic association of facts and rules.

7.3. EXPERT SYSTEM DEVELOPMENT

Several approaches for developing expert systems have been proposed (Buchanan et al., 1983; Waterman, 1986; Jones and Barrett, 1989; Arbour, 1992). Waterman (1986) has provided the most widely accepted approach:

- 1) Identification.
- 2) Conceptualization.
- 3) Formalization.
- 4) Implementation.
- 5) Testing.

Figure 7.2 illustrates these essential stages for expert system development.

7.3.1. Identification

Identification is the requirements analysis step carried out in traditional software development. It involves a formal task analysis to determine the external requirements, form of the input and output, setting where the program will be used and determines the user, which is very important. The participants, the problems, the objectives, the resources, the costs, and the time frame need to be clearly identified at this stage.

The participants are the group sponsoring the effort, the domain experts, and the knowledge engineer. Choosing an appropriate domain expert is essential to the success of the project. The domain expert should be a legitimate authority in the subject matter area, as software must possess high quality knowledge, and this person must have time and interest to commit to the project.

Although use of human domain experts is the typical method of development, it should be noted that several successful programs have been produced using reference materials only, or with minimal involvement of a human domain expert.

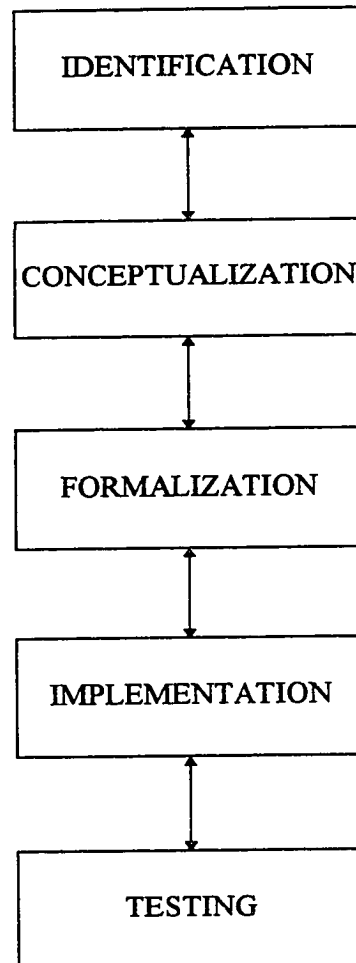


Figure 7.2. Expert system development stages.

To justify the time and cost of development, the problem must be important to the participants and be clearly defined. Although a developer cannot ignore interactions

between the problem and the rest of the subject matter domain, efforts should be made to limit the problem domain so that the recommendations of the program will be specific and valuable instead of generally educational. Choosing depth over breadth not only makes the program more powerful and useful, but also more efficient. Consequently, the amount of information, that must be obtained from the user before a recommendation can be made, is minimized.

For example, it would be more efficient for a user who has a problem with sampling correctness to run a program dealing with correctness in sampling, rather than a program dealing with sampling in general, being forced through a lengthy series of questions, or menus, before finally arriving at a subset of the program that deals with correctness.

Specific goals or purposes of the expert system must be accepted by all the participants. Objectives need to be more than problem solving. It is essential to carefully consider the background and needs of the end-user.

As important and as obvious as a properly designed user interface may seem to be, it is often neglected. Often the struggle to complete the knowledge base is so difficult and time consuming, developers have little energy left for the user interface.

Funding and time are major resources to be considered. Additional resources to be identified include the knowledge sources, computer hardware and development software. As with all programming projects, these estimates are difficult, but they must be realistic. Budgeted costs should include the cost of lost productivity by the expert and the programmer who will be devoting time to the effort and the ongoing cost of maintaining the knowledge base. By the same token, the expected benefits must include an estimate of the savings of valuable time in future years.

Some estimate of the useful life of the program should be made. Additional questions include how frequently the expertise will be needed, the cost and availability of alternate methods of solving the problem and the likely acceptability in the workplace. A realistic appraisal of the costs and benefits can help establish the level of program detail that can be justified.

The hardware available for delivery can greatly affect the choice of computer used for development, since the developer must determine:

- 1) The extent of help messages and graphics.
- 2) The form of question asked.
- 3) The extent and format of output.
- 4) The need to interact with other programs and databases.

Many troubleshooting and classification problems require input based on results of sensory examination of an environment.

High resolution color graphics should be especially useful in mineral processing troubleshooting or classification applications. High quality, inexpensive PC graphics as well as high resolution color scanners and video capture devices should be used where advantageous to reduce potential confusion on the user's part in answering questions posed by the program or in interpreting program output. The less experienced the end-user is with computer hardware and software, the more effort must be taken in the design of the user to machine interface. Expert systems have the added advantage of being more transparent than conventional programs, an ability that should be exploited if the user is likely to be skeptical of "black box" computer output.

7.3.2. Conceptualization

The second stage of expert systems development, conceptualization, involves designing the proposed program to ensure that specific interactions and relationships in the problem domain are understood and defined. The key concepts, relationships between objects and processes and control mechanisms are determined. This is the initial stage of knowledge acquisition. It involves the specific characterization of the situation and determines the expertise needed for the problem solution.

The following questions may be used by the knowledge engineer to help understand what the expert does:

- 1) Exactly what decisions does the expert make?
- 2) What are the decision outcomes?
- 3) Which outcomes require greater reflection, exploration or interaction?
- 4) What resources or inputs are required to reach a decision?
- 5) What conditions are present when a particular outcome is decided?
- 6) How consistently do these conditions predict a given outcome?
- 7) At what point after exposure to influential inputs is a decision made?
- 8) Given the particulars of a specific case, will the outcome predictions of the knowledge engineering team be consistent with those of the expert?

One of several or combinations of several knowledge acquisition methods are used. A typical approach would be to characterize the questions the end-user might pose to the domain expert and the range of possible solutions. One method of getting started is to begin with a range of final recommendations, and then build pathways to these. For example, in expert systems development to troubleshoot sampling incorrectness problems

in a mineral processing plant, the top level of programming might involve the following typical symptom and recommendations:

sampling error variance too high \Rightarrow *clean cutter and/or
increase cutter speed and/or change the cutter driving system and/or
increase the cutter width and/or increase the cutter depth.* (7.3)

The development process beyond this point is mainly one of refinement and addition of details once this top level is in place. For instance, additional information would be added to help determine whether the hypothesis “sampling error variance too high” is true. Indeed, the knowledge base evolves during this refining process to provide a recommendation as accurate as that made by the human expert.

The job of the knowledge engineer is to identify the knowledge sources required by the domain expert when making a specific recommendation, i.e., determine the reference books to be consulted, calculations to be made or other computer programs executed, and what rules-of-thumb or heuristics come into play.

Information the user will likely not know should be determined and represented by additional rules or other knowledge structures. Supplementary information needed to apply these rules can then be obtained from the user or additional rules created. This structure is typically created through frequent and intensive interview sessions with the domain experts or knowledge collection from experts publications.

Opportunities to group, rank, and order knowledge should be sought. The information that is collected and analyzed forms the basis of the scenarios to be presented in the next session with the expert. Correct and complete description of the expert’s logic to solve a problem is difficult because true experts usually do not know exactly how they reach a

decision. Therefore, they are often unable effectively to verbalize their own problem solving process. The careful study of detailed cases often reveals consistent patterns in the solution process that are still obscure. Needed refinements to the concepts and relationships become apparent during in-depth analyses.

7.3.3. Formalization

Formalization involves organizing the key concepts, subproblems, and information flow into formal representations. The program logic is designed at this stage. It is often useful to group or modularize the knowledge collected, perhaps even attempting to display the problem solving steps graphically.

It is the job of the knowledge engineer to build a set of interrelated tree structures for representing the knowledge base. They must decide the attributes to be determined to solve the problem and then which of these attributes should be asked of the user or represented by an internal set of decision trees. While decision trees are appealing in their simplicity and are a good way to begin formalizing knowledge into a knowledge representation scheme that can be visualized, things are rarely this simple in practice and rigid adherence to a tree structure is seldom satisfactory.

The representation of knowledge is important for credibility and acceptance by the user. The questions asked and the rules examined should be in the same sequence as used by the human expert. The questions and their order are determined by presenting the expert with several detailed scenarios. The granularity and structure of the concepts, including how the concepts relate into a logical flow and how uncertainties are involved, are coordinated in making recommendations.

The problem domain is analyzed to uncover obscure behavioral and mathematical models that may exist within the decision making process. The characteristics of the information needed are recognized. It follows that as the uncertainties are defined and explained, the relationships involved become better understood and ultimately may be explained using conventional programming techniques in a more expedient manner.

It is difficult to separate the conceptualization phase from the formalization phase and, in reality, knowledge-base design proceeds almost in parallel with knowledge acquisition. The two items that are the most important in the formalization stage are:

- 1) Refinement of the knowledge pieces into their specific relationships and hierarchy.
- 2) More accurate determination of the expected user interaction with the system

7.3.4. Implementation

During the next stage, implementation, the formalized knowledge is mapped or coded into the framework of the development tool to build a working prototype. The contents of knowledge structures, inference rules, and control strategies established in the previous stages are organized into suitable format. Often, knowledge engineers will have been using the program development tool to build a working prototype to document and organize information collected during the formalization stage, so that implementation is completed at this point. If not, the notes from the earlier phases are coded at this time.

Consideration must be given to long-term maintenance. Modifications to the knowledge base over time must be anticipated. The knowledge base should be extensively documented as it is coded. The potential for later misunderstanding and confusion should be minimized wherever possible. In addition, extensive justifications and explanations should be included to assist the end-users in fully understanding questions posed to them

by the program, so that the users can effectively use the program output, and to show the users, on demand, how the recommendation was logically derived.

The amount of help to be incorporated will depend on the ability of the anticipated user. While a consultant may be interested in quickly obtaining an answer to a question, an expert system intended to be used by those who must accomplish the recommendation is different. Typically, to believe the recommendation the end-user needs access to the assumptions underlying the recommendation and desires a credible justification for program recommendations.

This is also the point where the developer must decide how the program will interact with other computer programs and databases. The first generation of expert systems was standalone programs. Many had no facilities to communicate with the operating system or to read from, or write to databases.

7.3.5. Testing

The last stage, testing or validation, involves considerably more than finding and fixing syntax errors. It covers the verification of individual relationships, validation of program performance and evaluation of the utility of the software package. Testing guides reformulation of concepts, redesign of representations, and other refinements. Verification and validation must occur during the entire development process. Verification proves that the models within the program are true relationships. It ensures that the knowledge is accurately mimicked by having the domain expert operate the program for all possible contingencies.

The most difficult aspect of testing is accurately handling the uncertainty that is incorporated in most expert systems. Certainty factors are one of the most common

methods for handling uncertainty. Verification of the certainty factors assigned to the knowledge base is largely a process of trial and error, refining the initial estimates by the domain expert until the program consistently provides recommendations at a level of certainty that satisfies the expert. To ensure program accuracy, all possible solution paths must be carefully evaluated.

An effective validation procedure is critical to the success and acceptance of the program. During validation the following areas are of concern:

- 1) Correctness, consistency, and completeness of the rules.
- 2) Ability of the control strategy to consider information in the order that corresponds to the problem solving process.
- 3) Appropriateness of information about how conclusions are reached and why certain information is required; and most critical.
- 4) Agreement of the computer program output with the domain expert's corresponding solutions.

How the sequence of questions and output are presented to the end-user may have as much to do with acceptance and use as does the accuracy of the recommendations. The lessons learned from human engineering cannot be ignored if the program is to be successful.

Validation is an ongoing process requiring the output recommendations be accurate for a specific user's case. Validation is enhanced by allowing others to review critically and recommend improvements. A formal project evaluation is helpful to establish whether the system meets the intended original goal. The evaluation process focuses on uncovering problems with the credibility, acceptability, and utility. This can be determined from the program accuracy that is determined from comparisons with the real-world environment.

Included are the understanding and flexibility of the program, ease of use, adaptability of the design and the correctness of solutions.

7.4. EXPERT SYSTEM SAMPLING CORRECTNESS INSPECTOR

7.4.1. Solution process

The expert system *Sampling Correctness Inspector (SCI)* acquires its knowledge base (see Appendix 7.1) through collection of facts and rules from the experts publications (Gy, 1965, 1979, 1988, 1992; Merks, 1985; Holmes, 1991; Pitard, 1992, 1993) in addition to the author's expertise in sampling of mineral processing streams (Ketata, 1991; Hodouin and Ketata, 1994; Ketata and Rockwell, 1998). The main conditions of sampling correctness are exposed in Sections 2.5 and 6.2. This knowledge is then processed by the inference mechanism of the expert system to give an evaluation of the sampling process correctness.

The expert system *SCI* tests the correctness of the following factors:

- 1) Sampling scheme.
- 2) Cutter geometry.
- 3) Cutter strength.
- 4) Cutter clearance.
- 5) Cutter driving system.
- 6) Cutter idle positions.
- 7) Vertical distance between the stream and the cutter.
- 8) Cutter length.
- 9) Cutter path.
- 10) Cutter width.

- 11) Cutter depth.
- 12) Cutter speed.

It is assumed that all these factors are independent. Consequently, a conclusion concerning the sampling process correctness is made. If at least one factor is not correct, then the whole sampling process is not correct. Each of these factors has been further investigated to make a treelike structure as shown in Figure 7.3. It illustrates a detailed division of the correctness inspection procedure. In graph theory this structure is called an And Or graph.

The *SCI* obtains the information about a sampling situation through an interactive session with the user as shown in the next subsection. The user is asked to respond to a set of questions dealing with the previous factors. The information obtained this way is then fed into the rule base for processing purposes. The knowledge base of *SCI* is made up of 152 entries that are facts or rules.

The expert system *SCI* inspects the correctness of the sampling factors indicated formerly and then give the corresponding conclusion and advice to correct the faulty situation if necessary. Finally, a general conclusion is displayed indicating whether the sampling process is correct or not.

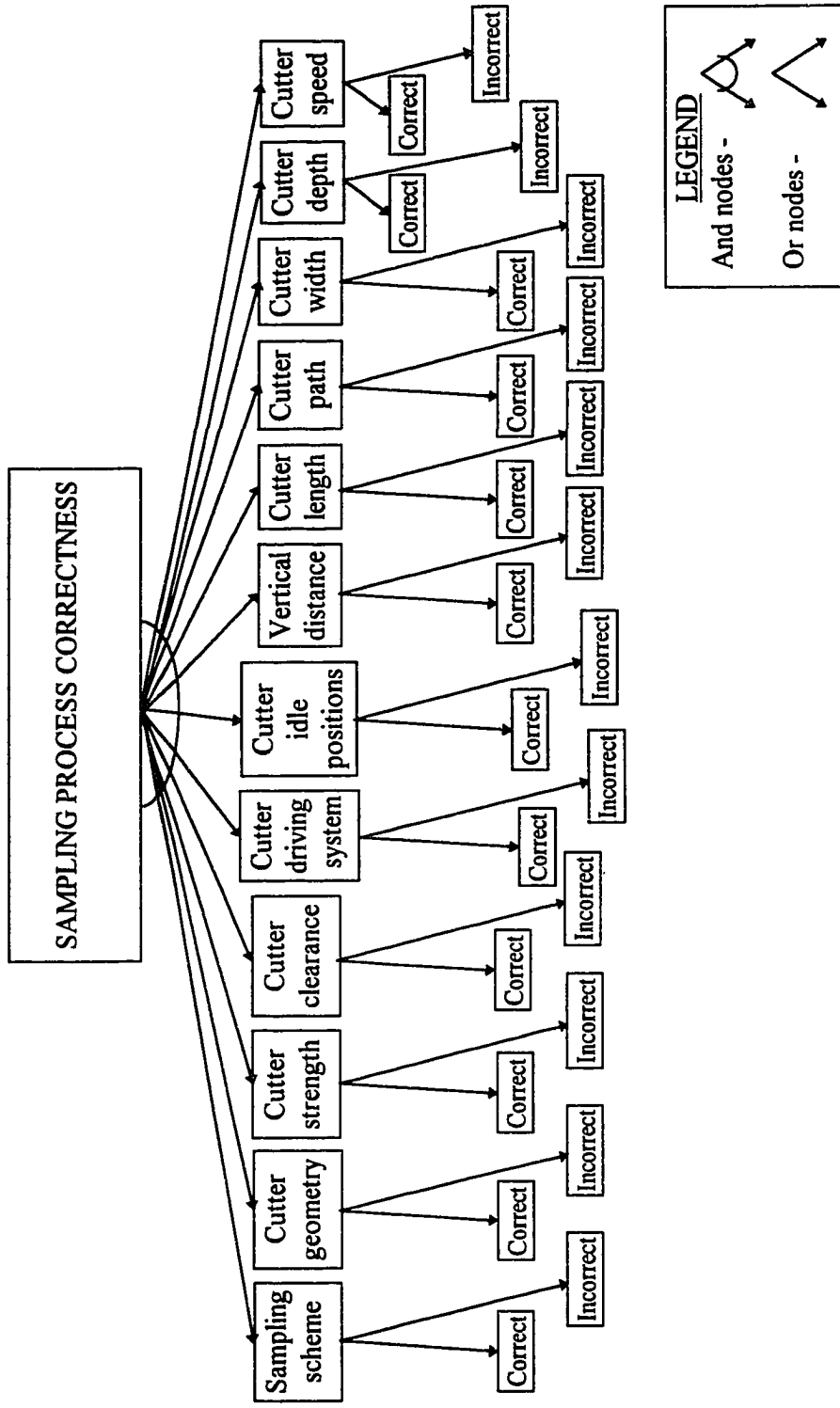


Figure 7.3. Network diagram of the Sampling Correctness Inspector.

7.4.2. Example run of the *Sampling Correctness Inspector*

The behavior of the expert system *SCI*, during a consultation, is as follows:

```

"----- SAMPLING PROCESS INSPECTOR -----
"
"
"
"What is your sampling scheme?
"
"  taking-the-whole-stream-part-of-the-time
"
"  taking-part-of-the-stream-all-of-the-time
"
"  taking-part-of-the-stream-part-of-the-time
"
">"taking-the-whole-stream-part-of-the-time"
"
"SAMPLING SCHEME:
"
"The sampling scheme is correct.
"
"
"
"What is your cutter geometry?
"
"  rectangular
"
"  square
"
"  circular
"
"  triangular
"
"  trapezoidal
"
"  other
"
">"rectangular"
"
"CUTTER GEOMETRY:

```

"
"The cutter geometry is correct.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"yes"
"
"CUTTER OBSTRUCTION:
"
"The presence of sticky material in the cutter is not correct.
"
"ADVICE: Clean the cutter with a water hose to clear it of sticky material.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"

" magnetic

"

" manual

"

">"electric"

"

"CUTTER DRIVING SYSTEM:

"

"The cutter driving system is correct.

"

"

"

"Are all the cutter idle positions located outside the stream?

"

" yes

"

" no

"

">"no"

"

"CUTTER IDLE POSITIONS:

"

"The cutter idle positions are not correct.

"

"ADVICE: The cutter idle positions should be outside the stream.

"

"

"

"What is the vertical distance, in cm, between the stream and the cutter?

"

">10

"

"What is the largest particle diameter in cm?

"

">0.02

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is correct.

"

"

"

"What is the cutter length in cm?

```
"
">10
"
"What is the stream thickness in cm?
"
">6
"
"CUTTER LENGTH:
"
"The cutter length is correct.
"
"
"
"Is the cutter path normal to the stream?
"
"  yes
"
"  no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
```

"The cutter depth is correct.

"

"

"

"What is the cutter speed in m/s?

"

">1

"

"Is the cutter speed constant?

"

" yes

"

" no

"

">"no"

"

"CUTTER SPEED:

"

"The cutter speed is not correct.

"

"ADVICE: The cutter speed should be constant, higher than 0.6 m/s, and lower than 0.8660380 m/s.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"-----END-----"

7.4.3. Validation

The expert system *SCI* is used to test the sampling process correctness. An example run is illustrated in the previous subsection. To validate this expert system, thirteen case studies (see Appendices 7.2 through 7.14) are considered. In the first sampling situation, all the factors are correct. Therefore, the sampling process is correct. However, in the remaining sampling situations, one factor is faulty. As a result, these sampling processes are not correct. Table 7.1 reveals the faulty factor in the incorrect sampling situations. The results of these case studies are compared to the results of the experts. Indeed, all the case studies lead to correct conclusions and advices (see Appendices 7.2 through 7.14).

Table 7.1. Twelve faulty sampling situations.

Case study #	Incorrect sampling factor
2	Sampling scheme
3	Cutter geometry
4	Cutter strength
5	Cutter clearance
6	Cutter driving system
7	Cutter idle positions
8	Vertical distance between the stream and the cutter
9	Cutter length
10	Cutter path
11	Cutter width
12	Cutter depth
13	Cutter speed

7.5. EXPERT SYSTEM SAMPLING ERROR EVALUATOR

7.5.1. Solution process

The suggested expert system *Sampling Error Evaluator (SEE)* acquires its knowledge base (see Appendix 7.15) through collection of facts and rules from the experts publications (Gy, 1965, 1979, 1988, 1992; Merks, 1985; Holmes, 1991; Pitard, 1992, 1993) in addition to the author's expertise in sampling of mineral processing streams (Ketata, 1991; Hodouin and Ketata, 1994; Ketata and Rockwell, 1998). This knowledge is subsequently processed by the inference mechanism of the expert system to give an evaluation of the sampling error.

If the sampling process is correct as assessed by the expert system *SCI*, then the sampling error is assumed to be random and equal to the continuous selection error (see Section 2.4). Therefore the sampling error is characterized by a given average, nil, and a given variance, never nil. The latter is computed by *SEE*. The knowledge base of *SEE* comprises two distinct parts:

- 1) Fundamental error or composition error.
- 2) Grouping and segregation error or distribution error.

Each of these two parts has been further explored to make a treelike structure as shown in Figure 7.4. A detailed division of each part is illustrated. In graph theory this structure is called an And Or graph. It takes into account the following factors:

- 1) Particle shape.
- 2) Minimum particle size.
- 3) Maximum particle size.

- 4) Critical component composition.
- 5) Critical component density.
- 6) Noncritical component density.
- 7) Material constitution heterogeneity.
- 8) Liberation size.
- 9) Lot mass.
- 10) Solid flowrate of the stream
- 11) Cutter width.
- 12) Cutter speed.
- 13) Material distribution heterogeneity.

The *SEE* obtains the information about a sampling situation through an interactive session with the user. The user is asked to respond to a set of questions dealing with the previous factors. An example run of *SEE* is displayed in the next subsection. The information obtained this way is then fed into the rule base for processing purposes. The knowledge base of *SEE* is made up of 112 entries that are facts or rules.

In addition to other data, the number of sample increments extracted for the sample analysis is entered by the user. As a result, the expert system *SEE* evaluates the sampling error corresponding to this number of sample increments. Besides, it proposes the minimum number of increments based on the variance ratio r :

$$r = \sigma_e^2(n) / \sigma_e^2(1) \quad (7.4)$$

where $\sigma_e^2(n)$ and $\sigma_e^2(1)$ symbolize the sampling error variances for n increments and one increment respectively. Here, n represents the minimum number of increments to respect the ratio requirement. The value of the ratio r is supplied by the user.

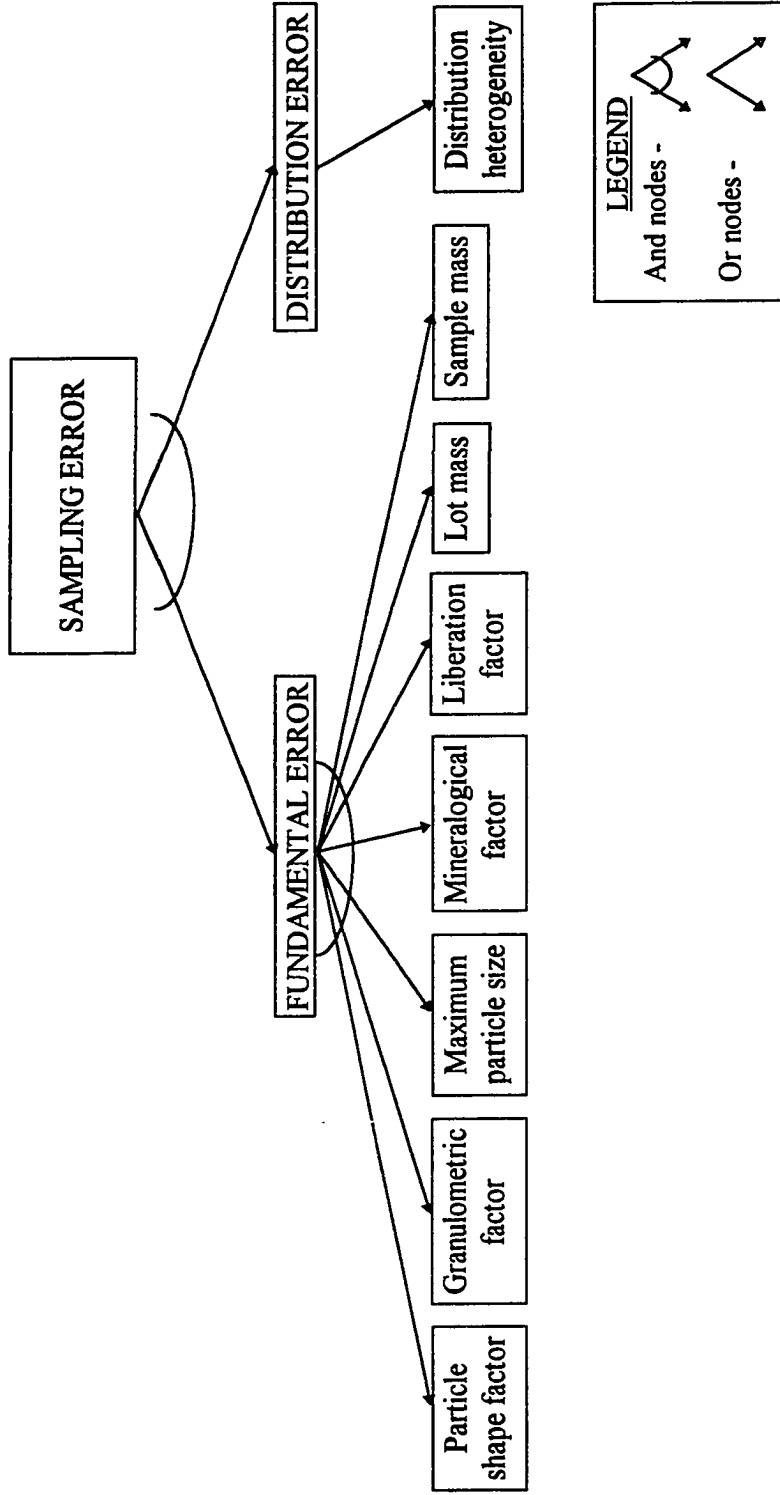


Figure 7.4. Network diagram of the Sampling Error Evaluator.

7.5.2. Example run of the *Sampling Error Evaluator*

The behavior of the expert system *SEE*, during a consultation, is as follows:

```
"----- SAMPLING ERROR EVALUATOR -----"  
"  
"  
"  
"What is the solid flowrate of the stream in t/h?  
"  
">10  
"  
"What is the cutter width in cm?  
"  
">2  
"  
"What is the cutter speed in m/s?  
"  
">0.6  
"  
"INCREMENT MASS:  
"  
"The mass of an increment equals 92.59260 g.  
"  
"  
"  
"What is the shape of the particles?  
"  
"  cube  
"  
"  sphere  
"  
"  flake  
"  
"  nugget  
"  
"  needle  
"  
">"sphere"  
"  
"What is the minimum particle size in cm?
```

"
">0.002
"
"What is the maximum particle size in cm?
"
">0.02
"
"What is the composition of the critical stream component in %?
"
">20
"
"What is the density of the critical stream component in g/cm³?
"
">4.3
"
"What is the density of the noncritical stream components in g/cm³?
"
">2.8
"
"How is the constitution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"average"
"
"What is the liberation size in cm?
"
">0.02
"
"What is the lot mass in t?
"

">100
"
"How is the distribution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"average"
"
"INCREMENT SAMPLING ERROR:
"
"The sampling error variance equals 2.71123e-005 % for one sample increment.
"
"
"
"What is the number of increments extracted for the sample analysis?
"
">10
"
"SAMPLING PERIOD:
"
"The sampling period equals 60.0 min for 10 sample increments.
"
"
"
"SAMPLING ERROR:
"
"The sampling error variance equals 2.71121e-006 % for 10 sample increments.
"
"
"

"What is the ratio, in %, of the sampling error variance for the minimum number of increments to the sampling error variance for one increment?"

"

">10

"

"MINIMUM INCREMENT NUMBER:

"

"The minimum increment number equals 11 for a sampling error variance ratio of 10 %.

"

"

"

"MAXIMUM SAMPLING PERIOD:

"

"The sampling period equals 54.54550 min for 11 sample increments.

"

"

"

"MAXIMUM SAMPLING ERROR:

"

"The sampling error variance equals 2.46473e-006 % for 11 sample increments.

"

"

"

"----- END -----"

7.5.3. Validation

The expert system *SEE* is used to evaluate the sampling error given that the sampling process is correct. An example run of *SEE* is shown in the previous subsection. To validate this expert system, three case studies (see Appendices 7.16 through 7.18) are considered. The user responds to a set of questions through an interactive session. Three different sampling situations are investigated. The results of these case studies are compared to the results of the experts. Indeed, all the case studies lead to correct evaluations (see Appendices 7.16 through 7.18).

CHAPTER 8

CONCLUSION

8.1. GENERAL CONCLUSIONS

This thesis includes six main chapters:

- Stream sampling.
- Stochastic modeling.
- Stochastic evaluation of sampling error.
- Sampling error filtering by material balance.
- Fuzzy evaluation of sample reliability.
- Expert systems *Sampling Correctness Inspector* and *Sampling Error Evaluator*.

8.1.1. Stream sampling

Stream sampling is introduced. The objective of stream sampling is to collect a sample that represents the average properties of the stream during a given time interval. This sample is obtained by gathering and blending the extracted increments following the systematic sampling strategy. In addition, the sampling errors due to material heterogeneity and cutter design are presented.

8.1.2. Stochastic modeling

Stochastic modeling is investigated. It deals with stochastic processes. A stochastic process is a family of time random variables. The theory of probability and statistics plays such an important role in sampling of mineral processing streams that some knowledge and understanding of the most important and useful concepts are required. Indeed, normal

probability distribution is defined. It is the most common model in use in conventional statistics.

Besides, random and systematic errors are described. Further, probabilistic selecting process is explained. Moreover, stationary time series and state space models and their properties are specified. A time series is an ordered sequence of observations through time. The state space representation of a system is a fundamental concept in modern control theory (Wei, 1990). The state of a mineral process is defined to be a minimum set of information from the present and past. Thus, the future behavior of the system can be completely described by the knowledge of the present state and the present input.

8.1.3. Stochastic evaluation of sampling error

The influence of the data variation on sampling errors is studied for a stream component. Besides, the proposed evaluation method is compared with that of Gy since these two procedures are different. A stream is sampled for the determination of its average composition. The latter is obtained by an averaging process of component mass flowrates. If the stream mass flowrate is constant through time, then the average stream composition is acquired by an averaging process of incremental compositions. The generated algorithms are developed using the programming language APL of Manugistics, Inc. (1992).

A stream is made up of large number of increments. It is sampled for the estimation of its average composition. A small number of increments is collected to compose the sample. Sampling is assumed to be correct. This means that all particles in the stream have a uniform probability of being sampled. Then, the sampling error is random. Therefore, its mean is zero and its variance is different from zero. The sampling error variance is

calculated based on the autocorrelation function used to describe the variation of stream incremental component flowrates.

This chapter presents a new expression for the variance of the error that is made when estimating an average stream composition by systematic incremental sampling. The stream component flowrates are assumed to randomly vary with a known autocovariance. The variance of the sample composition is a function of the number of sample increments and the time interval between the increments.

The samples collected from streams with higher autocorrelations are more representative. Finally, the proposed method for sampling error evaluation is compared to Gy's stochastic approach. For some sampling period values in the correlation length of the stream incremental composition process, the difference between the proposed formula and Gy's expression for sampling error variance is very high. This difference increases with the number of sample increments. Hence, Gy's expression is inaccurate.

The variance of the sampling error can be reduced drastically by data reconciliation techniques such as material balance. However, these techniques require that streams interconnected by process nodes are simultaneously sampled and that the variances and eventually the covariances between the errors on the various streams are known. Therefore, it is useful to estimate the sampling error variances.

8.1.4. Sampling error filtering by material balance

The influence of the data variation on sampling errors throughout a two-stage flotation circuit is analyzed. The material balance technique is used to upgrade the measured variables. The weighted least-squares method is applied to minimize the estimation errors. The sampling errors are evaluated before and after material balancing for comparison. In

one case, the covariance terms between the different components in a stream and between the different streams in the flotation circuit are included. In the other case, these terms are not included. The produced algorithms are elaborated using the programming language APL of Manugistics, Inc. (1992).

This chapter proves that, by including the covariance terms in the calculation, the variances of the sampling errors are reduced and therefore the reliability of the material balancing is improved.

8.1.5. Fuzzy evaluation of sample reliability

Fuzzy logic can be used to assess the sampling performance index that is the sample reliability. The latter is influenced by the cutter features describing the sampling conditions. Fuzzy logic with its intuitive nature and closeness to the natural languages offer significant advantages over traditional approaches in the appraisal of sampling conditions. Fuzzy logic allows sampling situations to be described and processed in linguistic terms such as *very reliable*, *reliable*, *adequate*, *doubtful*, and *very doubtful*. These fuzzy sets lead to the value of the sampling performance index.

A stream is made up of large number of increments. It is sampled for the estimation of its average composition. A small number of increments is collected to compose the sample. Therefore, the sampling process correctness should be inspected and evaluated. This means that all particles in the stream should have a uniform probability of being sampled.

The sampling process correctness is influenced by the cutter geometry, speed, layout, and path. Then, the degree of sample reliability is evaluated by applying fuzzy logic. The degree of sample reliability is also called sampling performance index. From this chapter development, the fuzzy logic appears to be a strong tool for estimating the sample

reliability. This estimation is the basis for decision-making concerning sampling strategy and data reconciliation. The judgment of plant operators and sampling experts is used to define all the fuzzy sets.

As a result, a value of sampling performance index is attached to each sample increment. It can be utilized as a weight in further calculations of average stream compositions or correct composition values for example. In this way, the degree of sample reliability is estimated by applying fuzzy logic.

8.1.6. Expert systems *Sampling Correctness Inspector* and *Sampling Error Evaluator*

The expert systems, also known as knowledge-based systems, emulate part of the human reasoning capabilities based on the knowledge of an expert or a specialist to solve problems. The task of developing expert systems is not easy but there are large benefits to be realized if the appropriate approach is taken. As research into the development of these systems continues, the practical benefits are being understood and used in industry.

This chapter deals with two expert systems for sampling correctness inspection, called *Sampling Correctness Inspector (SCI)*, and sampling error evaluation, named *Sampling Error Evaluator (SEE)*. They are developed through the utilization of an expert system shell M.4 of Teknowledge Corporation (1993, 1995, 1996). The knowledge is collected from experts publications in sampling of mineral processing streams (Gy, 1965, 1979, 1988, 1992; Merks, 1985; Holmes, 1991; Pitard, 1992, 1993) in addition to the author's expertise in the considered domain (Ketata, 1991; Hodouin and Ketata, 1994; Ketata and Rockwell, 1998). These expert systems take into account the stream properties, the sampler features, and the sampling manner.

The expert system *SCI* tests the correctness of the following factors:

- 1) Sampling scheme.
- 2) Cutter geometry.
- 3) Cutter strength.
- 4) Cutter clearance.
- 5) Cutter driving system.
- 6) Cutter idle positions.
- 7) Vertical distance between the stream and the cutter.
- 8) Cutter length.
- 9) Cutter path.
- 10) Cutter width.
- 11) Cutter depth.
- 12) Cutter speed.

It is assumed that all these factors are independent. Consequently, a conclusion concerning the sampling process correctness is made. If at least one factor is not correct, then the whole sampling process is not correct. Each of these factors has been further investigated to make a treelike structure.

The *SCI* obtains the information about a sampling situation through an interactive session with the user. The user is asked to respond to a set of questions dealing with the previous factors. The information obtained this way is then fed into the rule base for processing purposes. The knowledge base of *SCI* is made up of 152 entries that are facts or rules.

The expert system *SCI* inspects the correctness of the sampling factors indicated formerly and then give the corresponding conclusion and advice to correct the faulty situation if necessary. Finally, a general conclusion is displayed indicating whether the sampling process is correct or not.

To validate this expert system, thirteen case studies are considered. In the first sampling situation, all the factors are correct. Therefore, the sampling process is correct. However, in the remaining sampling situations, one factor is faulty. As a result, these sampling processes are not correct. The results of these case studies are compared to the results of the experts. Indeed, all the case studies lead to correct conclusions and advices.

If the sampling process is correct as assessed by the expert system *SCI*, then the sampling error is assumed to be random and equal to the continuous selection error. Therefore the sampling error is characterized by a given average, nil, and a given variance, never nil. The latter is computed by *SEE*. The knowledge base of *SEE* comprises two distinct parts:

- 1) Fundamental error or composition error.
- 2) Grouping and segregation error or distribution error.

Each of these two parts has been further explored to make a treelike structure. It takes into account the following factors:

- 1) Particle shape.
- 2) Minimum particle size.
- 3) Maximum particle size.
- 4) Critical component composition.
- 5) Critical component density.
- 6) Noncritical component density.
- 7) Material constitution heterogeneity.
- 8) Liberation size.
- 9) Lot mass.
- 10) Solid flowrate of the stream
- 11) Cutter width.

- 12) Cutter speed.
- 13) Material distribution heterogeneity.

The *SEE* obtains the information about a sampling situation through an interactive session with the user. The user is asked to respond to a set of questions dealing with the previous factors. The information obtained this way is then fed into the rule base for processing purposes. The knowledge base of *SEE* is made up of 112 entries that are facts or rules.

In addition to other data, the number of sample increments extracted for the sample analysis is entered by the user. As a result, the expert system *SEE* evaluates the sampling error corresponding to this number of sample increments. Besides, it proposes the minimum number of increments based on the variance ratio r defined by Equation 7.4. The value of the ratio r is supplied by the user.

To validate this expert system, three case studies are considered. The user responds to a set of questions through an interactive session. Three different sampling situations are investigated. The results of these case studies are compared to the results of the experts. Indeed, all the case studies lead to correct evaluations.

8.1.7. Comments

Given the previous description of the thesis structure, the ultimate thesis objective is the development of an intelligent sampling controller. Consequently, the actions taken by the operator to correct sampling conditions are assisted by the intelligent sampling hybrid system composed of:

- 1) *Sampling Error Filter (SEF)*. This is the product of Chapters 4 and 5.
- 2) *Sampling Performance Indexer (SPI)*. This is the product of Chapter 6.

- 3) *Sampling Correctness Inspector (SCI)*. This is the product of a part of Chapter 7
- 4) *Sampling Error Evaluator (SEE)*. This is the product of a part of Chapter 7.

The development of the expert systems *SCI* and *SEE* results in contributions to stream sampling in the mineral processing plants and to the application of expert systems. By collecting the sampling knowledge, an important knowledge domain is documented in an innovative way. The implementation of the formalized knowledge into the framework of the development tool to build the expert systems makes it available for a wide range of users. This represents a significant step. The generated expert systems provide a comprehensive approach to sampling of mineral processing streams.

8.2. RECOMMENDATIONS

This research work has introduced the concept of *Intelligent Sampling Control System* to control the sampling operations using plant data, operators experience, and experts knowledge. The findings from this highlight a number of important issues that need to be addressed in future research work.

8.2.1. Uncertain knowledge

The elaborated expert systems *SCI* and *SEE* represent and use crisp knowledge. Nevertheless, the sampling conditions encountered in a mineral processing plant are characterized by uncertain data and fuzzy variables such as material heterogeneity and cutter design. Hence, this uncertainty should be implemented to permit these expert systems to deal with uncertain situations.

If uncertain knowledge is encoded through the association of certainty factors with facts and rules, then these certainty factors, or measures of likelihood, are automatically used

during system execution. The aim is to control the inference process and qualify conclusions, thus allowing the representation and use of fuzzy data and uncertain knowledge. In a consultation, this feature lets a knowledge system to reach realistic conclusions, even with incomplete and conflicting user input.

8.2.2. Cross-stream cutters

This research work is restricted to straight path cutters. However, other categories of cross-stream cutters, such as circular path cutters, are used. Therefore, research to study other cutter types is needed to establish the correctness conditions and whether there is a different approach taken within these cases to ensure correct sampling and evaluate sampling errors.

8.2.3. Sampling instrumentation

The developed expert systems *SCI* and *SEE* cannot identify the correct sampling equipment supplied and used by the industry. Consequently, it is paramount to include a list of the industry sampling products including their features to assist the users in selecting the best quality samplers and sampling systems available in the market.

8.2.4. Material handling

This research is confined to straight path cutters for stream sampling in mineral processing plants. Nevertheless, the material handling industries are various. Specific cutters are utilized in each industry application. Thus, it is important to add more sampling aspects by incorporating sampling information and knowledge corresponding to each material handling industry.

REFERENCES

Abraham B. and Ledolter J., 1983. Statistical Methods for Forecasting. John Wiley & Sons, New York, New York, 445 pp.

Arbour J. H., 1992. Development and Validation of an Expert System for Soil Erosion Control Planning in Prince Edward Island. Ph.D. Thesis, Technical University of Nova Scotia, Halifax, Nova Scotia, 564 pp.

Bearman R.A. and Milne R., 1992. Expert Systems:- Opportunities in the Minerals Industries. *Minerals Engineering*, Vol. 5, Nos 10-12, pp. 1307-1323.

Bilonick R.A., 1986. How Far Does Gy's Theory Go toward Fulfilling the Role of a General Particulate Material Sampling Theory? Presentation at the SME Fall Meet., St. Louis, September 7.

Box G.E.P. and Jenkins G.M., 1976. Time Series Analysis: Forecasting and Control. Holden Day, San Francisco, California, 575 pp.

Buchanan B.G., Barstow D., Bechtel R., Bennett J., Clancey W., Kulikowski C., Mitchel T. and Waterman D.A., 1983. Constructing an Expert System. In: *Building Expert Systems*, F. Hayes-Roth, D.A. Waterman and D.A. Lenat eds., Addison-Wesley Publishing Company Inc., Reading, Massachusetts.

Cooper H.R., 1985. Theory and Practice of Incremental Sampling. *SME Mineral Processing Handbook*, N.L. Weiss (Editor), Volume 2, SME of the AIME, 30-4—30-20.

Crowe C.M., 1992. The Maximum-Power Test for Gross Errors in the Original Constraints in Data Reconciliation. *Can. J. Chem. Eng.*, pp. 1030-1036.

David M., 1977. Geostatistical Ore Reserve Estimation. Elsevier, Amsterdam, 364 pp.

Farzanegan A., Laplante A.R. and Lowther D.A., 1995. Application of Knowledge-Based Systems Technology in Optimization of Mineral Grinding Circuits. CAMI '95, Montreal, October 22-25, pp. 725-732.

Golsan J.E., Huffman J.E., Mazerolle D.J. and Vignando F.P., 1998. Intelligent Systems Technology for Semi-Autogenous Grinding Circuit. CIM/CMMI/MIGA '98 on CD ROM, Montreal, May 4-7.

Gray A.R. and Kasabov N.K., 1997. Round-Trip System Engineering in Neuro-Fuzzy Hybrid Systems. *Journal of Intelligent and Fuzzy Systems*, to appear.

Gueniffey Y., Kouniali S., Baroudi H. and Piguet J. P., 1995. Les Systèmes Experts dans l'Industrie Minière: Bilan de l'Expérience SOUT et Orientations Futures (The Expert Systems in the Mining Industry: Assessment of the SOUT Experience and Future Orientations). In French, CAMI'95, Montreal, pp. 560-566.

Gy P., 1965. The Notion of Heterogeneity. In French, *Bull. Bur. Rech. Geol. Minières*, p. 91.

Gy P., 1979. Sampling of Particulate Materials, Theory and Practice. Elsevier, Amsterdam, 431 pp.

- Gy P., 1988. Hétérogénéité, Échantillonnage, Homogénéisation (Heterogeneity, Sampling, Homogenization). In French, Masson, Paris, 608 pp.
- Gy P., 1992. Sampling of Heterogeneous and Dynamic Material Systems. Elsevier, Amsterdam, 654 pp.
- Hansen D., 1996. Memorandum Regarding Proposal for Department of Civil Engineering to Purchase Teknowledge M.4 and RoboHELP software Packages. Technical University of Nova Scotia, Halifax, July 30.
- Harmon P. and King D., 1986. Expert Systems: Artificial Intelligence in Business. Wiley, New York, New York.
- Hayes-Roth F., Waterman D.A. and Lenat, D.B., eds., 1983. Building Expert Systems. Addison-Wesley, Reading, Massachusetts.
- Hodouin D., Alliot N. and Flament F., 1991. Redundancy Analysis of Complex Sets of Mineral Processing Data for Mass Balance Computation. *Int. J. Miner. Process.*, 32, pp. 213-231.
- Hodouin D. and Everell M.D., 1980. A Hierarchical Procedure for Adjustment and Material Balancing of Mineral Processes Data. *Int. J. Miner. Process.*, 7(2), pp. 91-116.
- Hodouin D. and Flament F., 1985. BILMAT Computer Program. In: D. Laguitton (Editor), *The SPOC Manual*, Ch. 3.1, CANMET, Energy, Mines and Resources Canada, Ottawa, 141 pp.

- Hodouin D. and Flament F., 1989. New Developments in Material Balance Calculations for the Mineral Processing Industry. Proceedings SME/AIME Annual Meeting, Las Vegas, Paper 89-147.
- Hodouin D. and Flament F., 1991. Influence of Data Collection and Conditioning Strategies on the Significance of Performance Indices in Mineral Processing Plants. In: Proceedings of the Int. Symp. Evaluation and Optimization of Metallurgical Performance, SME/AIME, pp. 195-208.
- Hodouin D., Flament F. and Bazin C., 1989. Reliability of Material Balance Calculations — A Sensitivity Approach. Miner. Eng., 2(2), pp. 157-170.
- Hodouin D. and Ketata C., 1994. Variance of Average Stream Compositions Obtained by Automatic Incremental Sampling. Int. J. Miner. Process., 40, pp. 199-223.
- Holmes R.J., 1991. Sampling Methods: Problems and Solutions. In: Proceedings of the Int. Symp. Evaluation and Optimization of Metallurgical Performance, SME/AIME, pp. 157-167.
- Jackson P., 1990. Introduction to Expert Systems. Second Edition, Addison-Wesley Publishing Company, Reading, Massachusetts, 526 pp.
- Jämsä-Jounela S.-L., 1990. Expert System for Phosphate Flotation Control. Conference Paper, Proceedings. APCOM XXII, Volume I, Berlin, pp. 281-294.
- Jones D.D. and Barrett J.R., 1989. Building Expert Systems. In J.R. Barrett and D.D. Jones (eds), Knowledge Engineering in Agriculture, ASAE Monograph No. 8, ASAE, St. Joseph, MI.

Journel A. and Huijbregts C., 1978. Mining Geostatistics. Academic Press, New York, New York, 600 pp.

Ketata C., 1991. Influence de la Structure de Variation des Signaux dans les Usines Minéralurgiques sur les Erreurs d'Échantillonnage et la Précision des Bilans de Matière (Influence of the Signal Variation Structure in the Mineral Processing Plants on the Sampling Errors and the Material Balance Precision). M.Sc. Thesis, in French, Laval University, Ste-Foy, Quebec, 153 pp.

Ketata C. and Rockwell M.C., 1998. Stochastic Evaluation of Sampling Errors in Mineral Processing Streams. CIM/CMMI/MIGA '98 on CD ROM, Montreal, May 4-7.

Kline P.J. and Dolins S.B., 1989. Designing Expert Systems: A Guide to Selecting Implementation Techniques. John Wiley & Sons, New York.

Laguitton D., 1985. MATBAL Computer Program. In: D. Laguitton (Editor), The SPOC Manual, Ch. 3.2, CANMET, Energy, Mines and Resources Canada, Ottawa, 58 pp.

Laguitton D. and Hodouin D., 1985. Material Balance. In: D. Laguitton (Editor), The SPOC Manual, Ch. 3, CANMET, Energy, Mines and Resources Canada, Ottawa, 22 pp.

Laguitton D. and Leung J., 1989. Advances in Expert System Applications in Mineral Processing. Processing of Complex Ores, ed. G.S. Dobby & S.R. Rao, Pergamon Press, New York, 565.

Laufmann S.C., DeVaney D.M. and Whiting M.A., 1990. A Methodology for Evaluating Potential KBS Applications. IEEE Expert, December, pp. 43-61.

Lyman G.I., 1986. Application of Gy's Sampling Theory to Coal: A Simplified Explanation and Illustration of Some Basic Concepts. *Int. J. Miner. Process.*, 17, pp. 1-22.

Manugistics, Inc., 1992. APL*PLUS Special Edition. Maryland.

Matheron G., 1965. Les Variables Régionalisées et leur Estimation (The Regionalized Variables and their Estimation). In French, Masson, Paris.

Merks J.W., 1985. Sampling and Weighing of Bulk Solids. Trans. Tech. Publications, Germany, 410 pp.

Mishkoff H.C., 1988. Understanding Artificial Intelligence. Howard W Sams and Company, Indianapolis, pp. 65-67.

Newell A., 1991. Unified Theories of Cognition. Harvard University Press. Cambridge, Massachusetts.

Pitard F., 1992. Pierre Gy's Sampling Theory and Sampling Practice. Vols. I and II, CRC Press, Boca Raton, Florida, 496 pp.

Pitard F., 1993. Pierre Gy's Sampling Theory and Sampling Practice: Heterogeneity, Sampling Correctness, and Statistical Process Control. 2nd Ed., CRC Press, Boca Raton, Florida, 488 pp.

Poirier P.J. and Meech J.A., 1993. Using Fuzzy Logic for On-Line Trend Analysis. Second IEEE Conference on Control Applications, September 13-16, Vancouver, pp. 83-86.

Prerau D.S., 1985. Selection of an Appropriate Domain for an Expert System. *AI Magazine*, Vol. 6, No. 2, pp. 26-30.

Reuter M.A. and Van Deventer J.S.J., 1990. A Knowledge Based System for the Simulation of Batch and Continuous Carbon-in-Leach Systems. Conference Paper, Proceedings, APCOM XXII, Volume I, Berlin, pp. 343-356.

Reuter M.A. and Van Deventer J.S.J., 1992. The Simulation and Identification of Flotation Processes by Use of a Knowledge Based Model. *International Journal of Mineral Processing*, 35, Elsevier Science Publishers B.V., Amsterdam, pp. 13-49.

Rockwell M.C., 1989. Sample Handling and Preparation. *Canadian Offshore Non-Fuel Mineral Resources: Assessment and Planning*, Ed. Hale, P., EMR.

Rose C.D., 1983. Variances in Sampling Streams of Coal. *Journal of Testing and Evaluation*, Vol. 11, No. 5, September, pp. 320-326.

Sastry K.V.S., 1997. Databases for Mineral Processing. *Mining Magazine*, November, pp. 306-309.

Saunders I.W., Robinson G.K., Lwin T. and Holmes R.J., 1989. A Simplified Variogram Method for Determining the Estimation Error Variance in Sampling from a Continuous Stream. *Int. J. Miner. Process.*, 25, pp. 175-198.

Slagle J. and Wick M., 1988. A Method for Evaluating Candidate Expert System Applications. *AI Magazine*, Vol. 9, No. 4, pp. 44-53.

Smith H.W., 1985. Sampling Methodology. In: D. Laguitton (Editor), The SPOC Manual, Ch. 2, CANMET, Energy, Mines and Resources Canada, Ottawa, 26 pp.

Smith R.G., 1984. On the Development of Commercial Expert Systems. AI Magazine, Vol. 5, No. 3, pp. 61-73.

Spollen C.F., Powell N. and Wills B.A., 1991. The Importance of Covariance in Mass Balancing Using Sensitivity Analysis Techniques. In: Proceedings of the Int. Symp. Evaluation and Optimization of Metallurgical Performance, SME/AIME, pp. 231-237.

Stephanopoulos G., 1984. Chemical Process Control: An Introduction to Theory and Practice. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 696 pp.

Teknowledge Corporation, 1993. M.4 Version 3.0 Release Notes: A Guide to Developing and Delivering Knowledge Systems. Palo Alto, California, 92 pp.

Teknowledge Corporation, 1995. M.4 User's Guide. Palo Alto, California.

Teknowledge Corporation, 1996. M.4 Version 3.01 Release Notes: A Guide to Developing and Delivering Knowledge Systems. Palo Alto, California, 29 pp.

Visman J., 1972. A General Theory of Sampling — Discussion 3. J. Mater., 7, pp. 345-350.

Waterman D.A., 1986. A Guide to Expert Systems. Addison-Wesley Publishing Co., Inc., Reading, Massachusetts.

Wei W.W.S., 1990. Time Series Analysis: Univariate and Multivariate Methods. Addison Wesley, Reading, Massachusetts, 478 pp.

Zadeh L.A., 1965. Fuzzy Sets. *Information and Control*, Vol. 8, pp. 338-353.

Zimmermann, H.-J., 1991. Fuzzy Set Theory and its Applications. 2nd Edition, Kluwer Academic Publishers, Boston, Massachusetts, 399 pp.

APPENDIX A
APL WORKSPACES OF THE SAMPLING ERROR FILTER

The *Sampling Error Filter* is composed of two APL workspaces (see Appendix H):

- 1) CHAPTER4.AWS (see Table A1).
- 2) CHAPTER5.AWS (see Table A2).

Table A1. Functions of the APL workspace CHAPTER4.AWS.

Function	Property
AR	computes the sampling errors of a simulated <i>AR</i> process as detailed in Chapter 4.
AR48	computes the sampling errors of a simulated <i>AR</i> process for a shift of eight hours as detailed in Chapter 4.
MA	computes the sampling errors of a simulated <i>MA</i> process as detailed in Chapter 4.
MA48	computes the sampling errors of a simulated <i>MA</i> process for a shift of eight hours as detailed in Chapter 4.
ARMA	computes the sampling errors of a simulated <i>ARMA</i> process as detailed in Chapter 4.
ARMA48	computes the sampling errors of a simulated <i>ARMA</i> process for a shift of eight hours as detailed in Chapter 4.
MAA	computes the sampling errors of a simulated <i>MA</i> process as detailed in Appendix C.

Table A2. Main function of the APL workspace CHAPTER5.AWS.

Function	Property
MB	computes the sampling errors of a simulated mineral process as detailed in Chapter 5.

APPENDIX B
INFLUENCE OF COMPOSITION VARIATIONS
ON SAMPLING ERRORS IN MINERAL PROCESSING STREAMS

B1. INTRODUCTION

During the operation of a mineral processing plant, the unit processes are investigated by analyzing the data extracted by means of stream sampling. The data structure is identical to the time series form. Data analysis is a crucial step in understanding the plant performance. To control the plant performance efficiently, it is very important to avoid sampling errors, which is impossible. However, it is possible to minimize these errors.

The factors influencing these errors are divided into two categories: material properties illustrating its heterogeneity, and cutter features. The material properties can be the sample mass, particle composition, mineral liberation, particle size distribution, or particle shape. The cutter features can be cutter geometry or speed, or sampling geometry. Some of these variables are controllable like the cutter features. However, the material properties cannot be controlled which leads to unavoidable sampling errors. The cutter features can be responsible for gross errors if the sampler is not designed and used according to the sampling correctness.

The purpose of this appendix is to study the influence of the composition variation on sampling errors, and to compare this evaluation method with that of Gy (1988, 1992) since these two procedures are different. Therefore, this appendix is divided into four main sections. First, stream sampling is introduced. A stream is made up of large number of increments. It is sampled for the estimation of its average composition. A small number of increments is collected to compose the sample. Sampling is assumed to be correct. This means that all particles in the stream have a uniform probability of being sampled.

Secondly, the sampling error is defined. Since stream sampling is considered to be correct, the sampling error is random. Then its mean is zero and its variance is different from zero. Thirdly, the sampling error variance is calculated based on the autocorrelation function used to describe the variation of stream incremental compositions. The evaluation of sampling error variance is useful for the selection of the number of increments that is required to obtain a specified accuracy of the average stream composition. Further, it helps for the evaluation of the weighting factors in data reconciliation techniques (Hodouin et al., 1989) used for process performance evaluation.

Finally, the proposed method for sampling error evaluation is compared to Gy's approach since his sampling theory of particulate material is a reference in mineral processing (Smith, 1985; Bilonick, 1986; Rockwell, 1989; Saunders et al., 1989; Ketata, 1991; Pitard, 1993; Hodouin and Ketata, 1994).

B2. STREAM SAMPLING

A stream is sampled for the estimation of its average composition. It is a slurry containing mineral particles. The stream composition is determined by mass fractions of the stream components. The considered sampling strategy is systematic. In this strategy the increments are extracted from the stream discharge following a sampling period given by:

$$k = N / n \quad (\text{B1})$$

where N and n are the numbers of increments in the stream and the sample respectively. The increments are extracted by a cross stream cutter (see Figure B1).

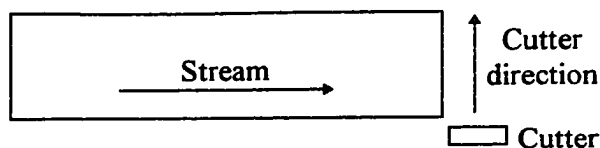


Figure B1. Stream sampling by a cross stream cutter.

B3. SAMPLING ERROR

For one stream component, the average stream composition is given by:

$$a_L = \frac{\sum_{i=1}^N t(i)m(i)a(i)}{\sum_{i=1}^N t(i)m(i)} \quad (\text{B2})$$

where i is the index of the increment, $m(i)$ the mass flowrate of the slurry at the time of the extraction of the i th increment, $a(i)$ the component composition at the time of extraction of the i th increment, and $t(i)$ the time during which each increment of the stream is collected. In this case, the average sample composition is obtained by:

$$a_S = \frac{\sum_{i=1}^n t(i)m(i)a(i)}{\sum_{i=1}^n t(i)m(i)} \quad (\text{B3})$$

To assure that the sampling process is correct, the cutter velocity should be constant. Therefore $t(i)$ is constant. Furthermore $m(i)$ is assumed to be constant. Consequently, the stream and sample compositions become respectively:

$$a_L = \frac{1}{N} \sum_{i=1}^N a(i) \quad (\text{B4})$$

$$a_S = \frac{1}{n} \sum_{i=1}^n a(i) \quad (\text{B5})$$

In the following sections, it is supposed that the variable $a(i)$ obey a normal distribution.

Let l be the time index of the first increment in a sample. Then the average sample composition equals:

$$a_S = \frac{1}{n} \sum_{i=0}^{n-1} a(l+ik) \quad (\text{B6})$$

Although the true average stream composition a is unknown, it is assumed to be equal to:

$$a_L = \frac{1}{nk} \sum_{i=1}^{nk} a(i) \quad (\text{B7})$$

Now, the sampling error e is defined by:

$$e = a_S - a_L \quad (\text{B8})$$

It can be expressed as follows:

$$e = \frac{1}{n} \sum_{i=1}^n [a(i) - a_L] \quad (\text{B9})$$

B4. SAMPLING ERROR VARIANCE

The sampling process is supposed to be correct. Therefore the sampling error is random. Consequently, it is characterized by a zero mean and a variance different from zero:

$$\sigma_e^2 = E(e^2) \quad (\text{B10})$$

where $E(e^2)$ is the mathematical expectation of the random variable e^2 . The variance of e is then:

$$\begin{aligned} \sigma_e^2 = & \frac{1}{n^2} E \left[\sum_{i=0}^{n-1} (a(l+ik) - a) \right]^2 + \frac{1}{(nk)^2} E \left[\sum_{i=1}^{nk} (a(i) - a) \right]^2 \\ & - \frac{2}{n^2 k} E \left[\left(\sum_{j=0}^{n-1} (a(l+jk) - a) \right) \left(\sum_{i=1}^{nk} (a(i) - a) \right) \right] \end{aligned} \quad (\text{B11})$$

which equals:

$$\begin{aligned} \sigma_e^2(n, k) = & \left(1 + \frac{1}{k} \right) \frac{\sigma^2}{n} \\ & + \frac{2\sigma^2}{n^2} \left[\sum_{i=1}^{n-1} (n-i) \rho(ik) + \frac{1}{k^2} \sum_{i=1}^{nk-1} (nk-i) \rho(i) - \frac{1}{k} \sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho(l+jk-i) \right] \end{aligned} \quad (\text{B12})$$

where σ^2 and $\rho(i)$ are respectively the variance and the autocorrelation function of the random variable $a(i)$. If the sample is composed of one increment, the sampling error variance becomes:

$$\sigma_e^2(1, k) = \sigma^2 \left[1 + \frac{1}{k} - \frac{2}{k} \sum_{i=1}^k \rho(l-i) + \frac{2}{k^2} \sum_{i=1}^{k-1} (k-i) \rho(i) \right] \quad (\text{B13})$$

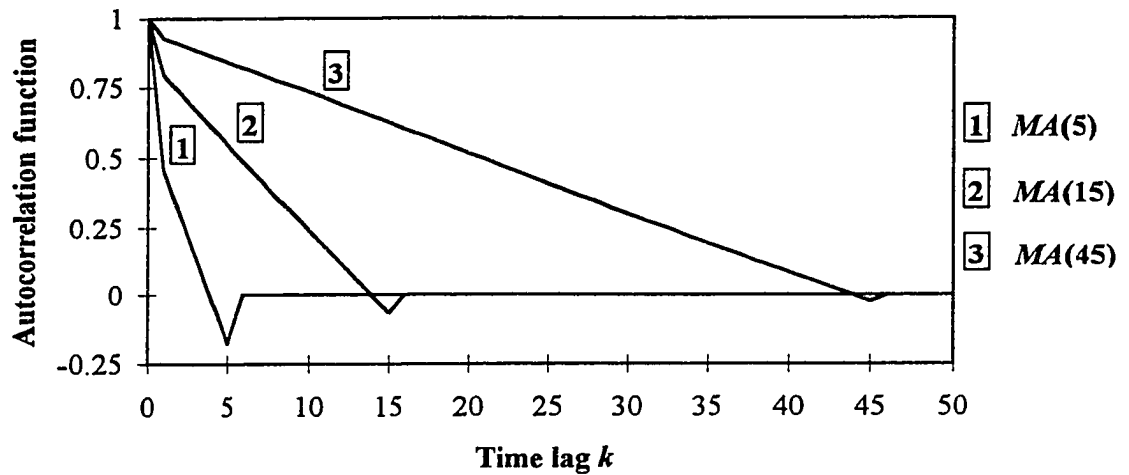


Figure B2. Autocorrelation functions of the three *MA* series.

Figure B2 illustrates the autocorrelation functions of three moving average (*MA*) processes describing three stream incremental compositions. These stochastic processes are generated following the moving average models (Wei, 1990). In this case, three moving average processes are obtained through *MA*(5), *MA*(15), and *MA*(45) models. The corresponding autocorrelation functions show significant correlation up to a lag of 5, 15, and 45 sampling periods respectively. These lags define the correlation lengths.

Figures B3 through B5 illustrate the relative sampling error variance:

$$\sigma_R^2 = \sigma_s^2(n, k) / \sigma^2 \quad (\text{B14})$$

when the sample consists of one, five, and ten increments respectively.

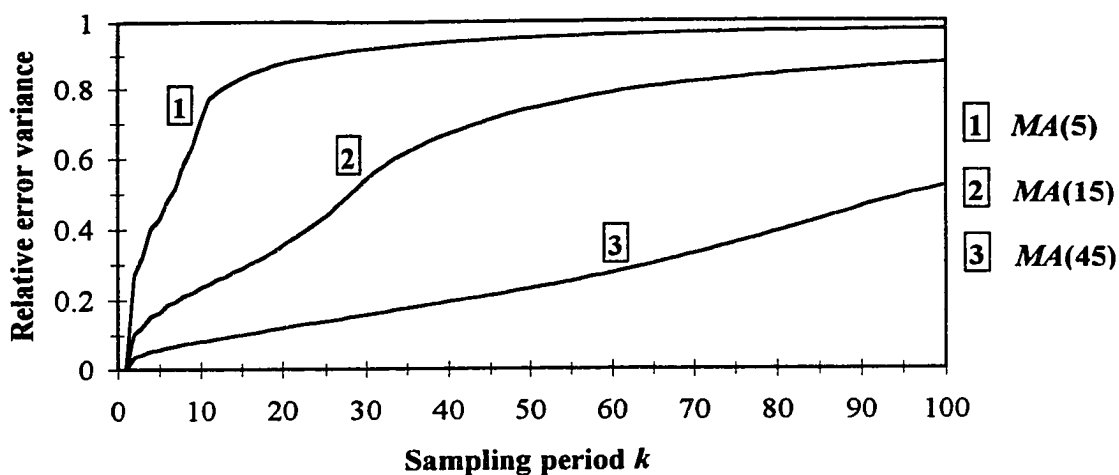


Figure B3. Relative sampling error variance σ_R^2 for a one-increment sample.

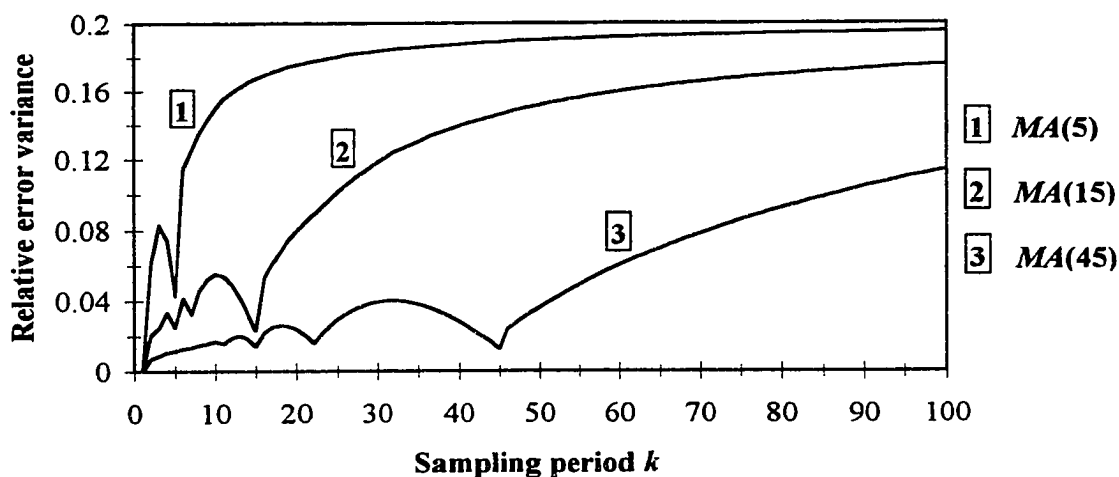


Figure B4. Relative sampling error variance σ_R^2 for a five-increment sample.

These figures demonstrate that the longer the correlation length the lower the sampling error variance. Thus, the samples collected from streams showing higher autocorrelations are more representative. Furthermore, it is revealed that the sampling error variance tends

towards the ratio of the stream composition variance to the number of sample increments. In addition, Figures B4 and B5 show that for more than one sample increment, oscillatory behavior is generated. This is explained by the periodicity in the systematic sampling strategy.

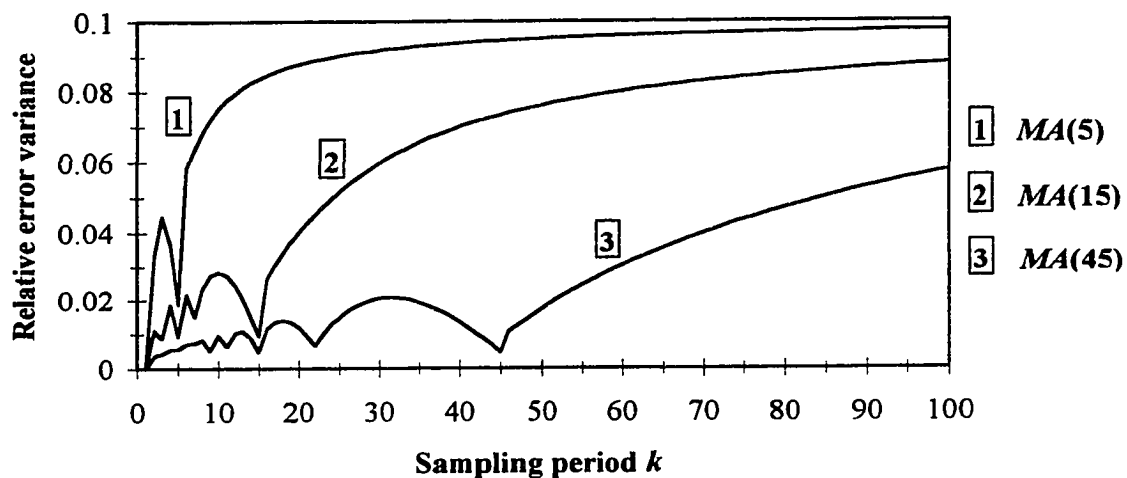


Figure B5. Relative sampling error variance σ_R^2 for a ten-increment sample.

B5. COMPARISON WITH GY'S METHOD

To evaluate the sampling error, Gy uses the heterogeneity contribution $h(i)$ as a stream descriptor instead of the component composition $a(i)$:

$$h(i) = \frac{a(i) - a_L}{a_L} \quad (\text{B15})$$

The sampling error, given by Equation B9, can be expressed as a function of $h(i)$:

$$e = \frac{a_L}{n} \sum_{i=1}^n h(i) \quad (\text{B16})$$

Moreover, to describe the data variation, he uses the variogram $v(k)$ instead of the autocorrelation function $\rho(k)$. The relationship between these tools is:

$$v(k) = \frac{\sigma^2}{a_L^2} (1 - \rho(k)) \quad (\text{B17})$$

The sampling error variance evaluated by Gy is:

$$\sigma_e^2(n, k) = \frac{1}{na_L} \sigma_e^2(1, k) \quad (\text{B18})$$

This indicates that Gy neglects the correlation between the sample increments. If $a(i)$ is considered instead of $h(i)$, Equation B18 becomes:

$$\sigma_e^2(n, k) = \frac{1}{n} \sigma_e^2(1, k) \quad (\text{B19})$$

Therefore the term neglected by Gy equals:

$$\begin{aligned} \sigma_e^2(n, k) - \frac{1}{n} \sigma_e^2(1, k) = & \frac{2\sigma^2}{n^2} \left[\sum_{i=1}^{n-1} (n-i)\rho(ik) + \frac{1}{k^2} \sum_{i=1}^{nk-1} (nk-i)\rho(i) \right. \\ & \left. - \frac{1}{k} \sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho(p+jk-i) + \frac{n}{k} \sum_{i=1}^k \rho(p-i) - \frac{n}{k^2} \sum_{i=1}^{k-1} (k-i)\rho(i) \right] \end{aligned} \quad (\text{B20})$$

Gy defines an error generator as a function of the sampling period k only:

$$W(k) = \sigma_e^2(1, k) / \sigma^2 \quad (\text{B21})$$

However the method proposed in this appendix defines the error generator as a function of both the number of sample increments n and the sampling period k :

$$W(n, k) = n \sigma_e^2(n, k) / \sigma^2 \quad (\text{B22})$$

Table B1 summarizes the differences between the proposed method and Gy's procedure for the evaluation of the sampling error.

Table B1. Comparison of the proposed method with Gy's procedure.

	Proposed method	Gy's method
Stream descriptor	composition $a(i)$	heterogeneity $h(i)$
Data variation descriptor	autocorrelation function $\rho(k)$	variogram $v(k)$
Sampling error	e	e/a_L
Sampling error variance	$\sigma_e^2(n, k)$	$\sigma_e^2(1, k) / na_L^2$
Error generator	$W(n, k) = n \sigma_e^2(n, k) / \sigma^2$	$W(k) = \sigma_e^2(1, k) / \sigma^2 a_L^2$

Figures B6 through B8 illustrate the relative error:

$$e_R = \frac{\sigma_e^2(n, k) - \frac{\sigma_e^2(1, k)}{n}}{\sigma_e^2(n, k)} \quad (\text{B23})$$

when the component composition $a(i)$ follows an $MA(5)$, $MA(15)$, and $MA(45)$ respectively.

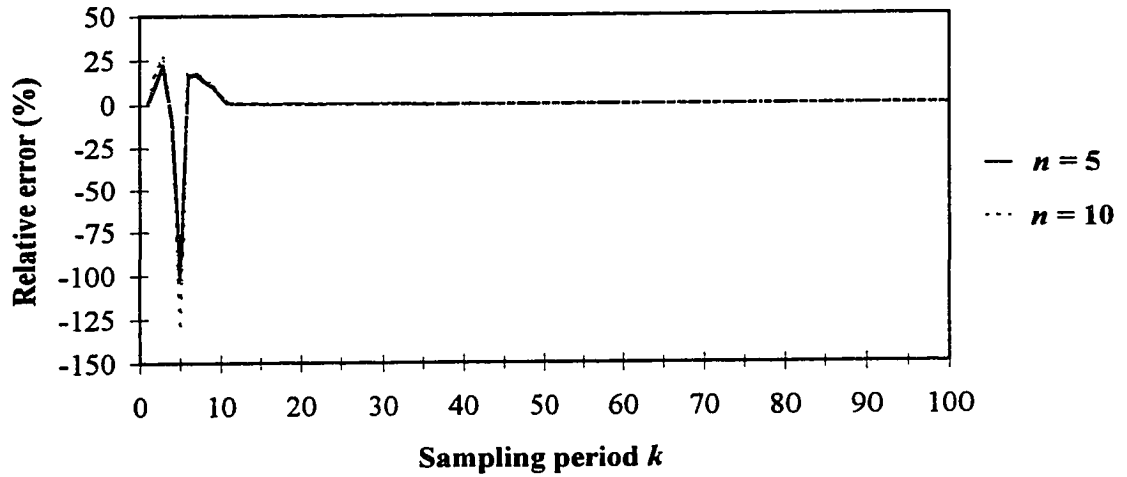


Figure B6. Relative error e_R for the $MA(5)$ series.

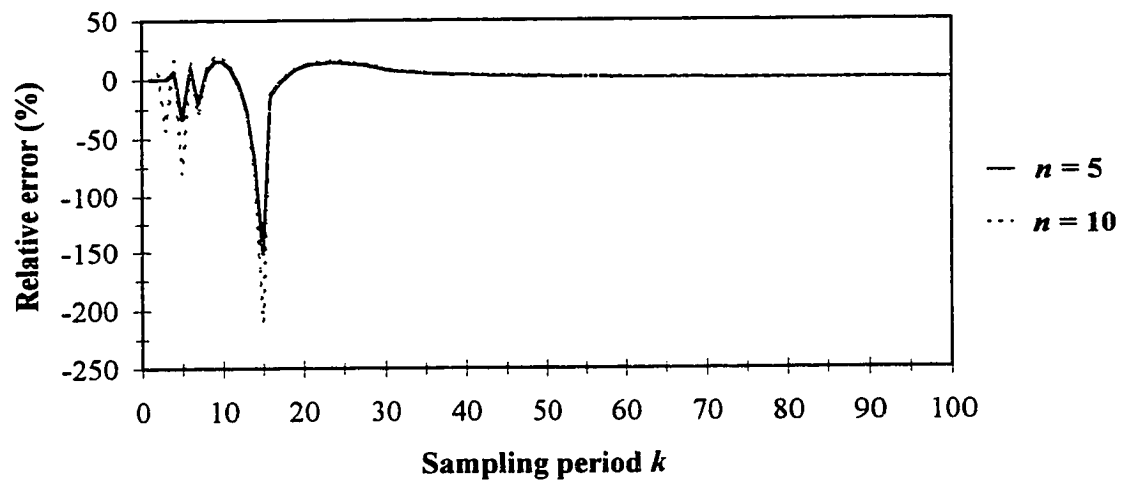


Figure B7. Relative error e_R for the $MA(15)$ series.

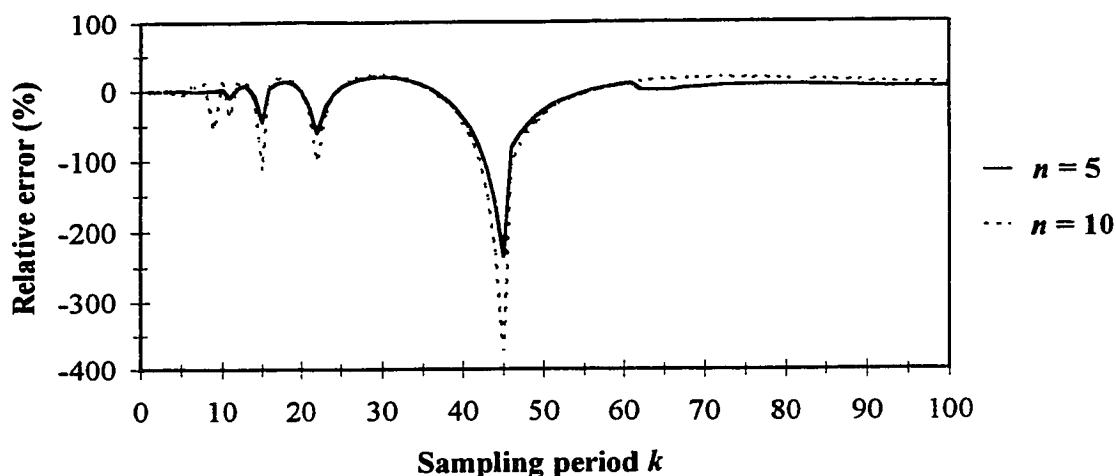


Figure B8. Relative error e_R for the $MA(45)$ series.

These figures show that the relative error is the highest at the sampling period corresponding to the correlation length. In this case the sampling error variance is overestimated by Gy. Besides, for some low values of k compared to the correlation length the approximation by Equation B13 is inaccurate. In addition, it is demonstrated that the relative error increases with the number of sample increments.

B6. CONCLUSION

A stream is made up of large number of increments. It is sampled for the estimation of its average composition. A small number of increments is collected to compose the sample. Sampling is assumed to be correct. This means that all particles in the stream have a uniform probability of being sampled. Then, the sampling error is random. Therefore, its mean is zero and its variance is different from zero.

The sampling error variance is calculated based on the autocorrelation function used to describe the variation of stream incremental compositions. The samples collected from

streams with higher autocorrelations are more representative. Finally, the proposed method for sampling error evaluation is compared to Gy's approach. For some k values in the correlation length of the stream incremental composition process, the difference between the proposed formula and Gy's expression for sampling error variance is very high. This difference increases with the number of sample increments.

APPENDIX C
INFLUENCE OF COMPOSITION AND MASS FLOWRATE VARIATIONS
ON SAMPLING ERRORS IN MINERAL PROCESSING STREAMS

C1. INTRODUCTION

The performance of a mineral processing plant is assessed regularly by stream sampling. The analysis of the resulting data is a crucial step in understanding the plant performance. The data structure is identical to the time series form. To control the sampling process efficiently, it is very important to minimize sampling errors and estimate them.

The factors influencing these errors are divided into two categories: material properties and cutter features. The material properties can be particle composition, mineral liberation, particle size distribution, and particle shape. These factors cannot be controlled due to material heterogeneity. However, the cutter features, such as cutter geometry, speed, layout, and path, are controllable. They can be responsible for gross errors if the cutter is not designed and used according to sampling correctness requirements.

The purpose of this appendix is to study the influence of the composition and mass flowrate variations on sampling errors. Furthermore, the proposed evaluation method is compared with that of Gy (1988, 1992) since these two procedures are different.

This appendix is divided into six main sections. First, stream sampling is introduced. A stream is made up of large number of increments. It is sampled for the estimation of its average composition. A small number of increments is collected to compose the sample. Sampling is assumed to be correct. This means that all particles in the stream have a uniform probability of being sampled.

Secondly, the sampling error is defined. Since stream sampling is considered to be correct, the sampling error is random. Then its mean is zero and its variance is different from zero. The average stream and sample compositions are not linear with respect to the mass flowrates. To apply linear statistics, the average stream and sample compositions must be linearized around their average values. This is described in the third and fourth sections.

Fifthly, the sampling error variance is calculated based on the autocorrelation functions used to describe the variations of stream incremental compositions and mass flowrates. The evaluation of sampling error variance is useful for the selection of the number of increments that is required to obtain a specified accuracy of the average stream composition. Further, it helps for the evaluation of the weighting factors in data reconciliation techniques (Hodouin et al., 1989) used for process performance evaluation.

Finally, the proposed method for sampling error evaluation is compared to Gy's approach since his sampling theory of particulate material is a reference in mineral processing (Smith, 1985; Bilonick, 1986; Rockwell, 1989; Saunders et al., 1989; Ketata, 1991; Pitard, 1993; Hodouin and Ketata, 1994; Ketata and Rockwell, 1998).

C2. STREAM SAMPLING

A stream is sampled for the estimation of its average composition. It is a slurry containing mineral particles. The stream composition is determined by mass fractions of the stream components. The considered sampling strategy is systematic. In this strategy the increments are extracted from the stream discharge following a sampling period given by:

$$k = N / n \quad (C1)$$

where N and n are the numbers of increments in the stream and the sample respectively. The sample increments are extracted by a cross stream cutter (see Figure B1).

C3. SAMPLING ERROR

For one component, the average stream composition is given by:

$$a_L = \frac{\sum_{i=1}^N t(i)m(i)a(i)}{\sum_{i=1}^N t(i)m(i)} \quad (\text{C2})$$

where i is the index of the increment, $m(i)$ the mass flowrate of the slurry at the time of the extraction of the i th increment, $a(i)$ the component composition at the time of extraction of the i th increment, and $t(i)$ the time during which each increment of the stream is collected. In this case, the average sample composition is obtained by:

$$a_S = \frac{\sum_{i=1}^n t(i)m(i)a(i)}{\sum_{i=1}^n t(i)m(i)} \quad (\text{C3})$$

To assure that the sampling process is correct, the cutter velocity should be constant. Therefore $t(i)$ is constant. Consequently, the stream and sample compositions become respectively:

$$a_L = \frac{\sum_{i=1}^N m(i)a(i)}{\sum_{i=1}^N m(i)} \quad (\text{C4})$$

$$a_s = \frac{\sum_{i=1}^n m(i)\alpha(i)}{\sum_{i=1}^n m(i)} \quad (\text{C5})$$

In the following sections, it is supposed that the variables $a(i)$ and $m(i)$ obey normal distributions.

Let l be the time index of the first increment in a sample. Then the average sample composition equals:

$$a_s = \frac{\sum_{i=0}^{n-1} m(l+ik)\alpha(l+ik)}{\sum_{i=1}^n m(l+ik)} \quad (\text{C6})$$

Although the true average stream composition a is unknown, it is assumed to be equal to a_L .

Now, the sampling error e is defined by:

$$e = a_s - a_L \quad (\text{C7})$$

C4. LINEARIZATION OF AVERAGE STREAM COMPOSITION

For one component, the average stream composition is presented by Equation C4. The latter shows that the composition a_L is not linear with respect to the mass flowrates $m(i)$, $i = 1, \dots, nk$. To apply linear statistics, the composition a_L must be linearized around its average value a as follows:

$$a_L = a + \sum_{i=1}^{nk} \left(\frac{\partial a_L}{\partial m(i)} \right) (m(i) - m) \quad (\text{C8})$$

where m symbolizes the mass flowrate average.

$\frac{\partial a_L}{\partial m(i)}$ equals:

$$\frac{\partial a_L}{\partial m(i)} = \frac{a(i)}{\sum_{j=1}^{nk} m(j)} - \frac{\sum_{j=1}^{nk} a(j)m(j)}{\left(\sum_{j=1}^{nk} m(j) \right)^2} \quad (\text{C9})$$

which is written:

$$\frac{\partial a_L}{\partial m(i)} = \frac{a(i) \left(\sum_{j=1}^{nk} m(j) \right) - \sum_{j=1}^{nk} a(j)m(j)}{\left(\sum_{j=1}^{nk} m(j) \right)^2} \quad (\text{C10})$$

As a result, a_L becomes equal to:

$$a_L = a + \frac{1}{mnk} \left(\sum_{i=1}^{nk} a(i)m(i) \right) - \frac{1}{mn^2k^2} \left(\sum_{j=1}^{nk} a(j) \right) \left(\sum_{i=1}^{nk} m(i) \right) \quad (\text{C11})$$

C5. LINEARIZATION OF AVERAGE SAMPLE COMPOSITION

For one component, the average sample composition is presented by Equation C6. The latter indicates that the composition α_S is not linear with respect to the mass flowrates

$m(i)$, $i = 1, \dots, n$. To apply linear statistics, the composition a_s must be linearized around its average value a as follows:

$$a_s = a + \sum_{i=1}^n \left(\frac{\partial a_s}{\partial m(i)} \right)_m (m(i) - m) \quad (\text{C12})$$

$\frac{\partial a_s}{\partial m(i)}$ equals:

$$\frac{\partial a_s}{\partial m(i)} = \frac{a(i)}{\sum_{j=1}^n m(j)} - \frac{\sum_{j=1}^n a(j)m(j)}{\left(\sum_{j=1}^n m(j) \right)^2} \quad (\text{C13})$$

which is written:

$$\frac{\partial a_s}{\partial m(i)} = \frac{a(i) \left(\sum_{j=1}^n m(j) \right) - \sum_{j=1}^n a(j)m(j)}{\left(\sum_{j=1}^n m(j) \right)^2} \quad (\text{C14})$$

As a result, a_s becomes equal to:

$$a_s = a + \frac{1}{mn} \left(\sum_{i=1}^n a(i)m(i) \right) - \frac{1}{mn^2} \left(\sum_{j=1}^n a(j) \right) \left(\sum_{i=1}^n m(i) \right) \quad (\text{C15})$$

C6. SAMPLING ERROR VARIANCE

The sampling process is supposed to be correct. Therefore the sampling error is random. Consequently, it is characterized by a zero mean and a variance different from zero:

$$\sigma_e^2 = E(e^2) \quad (\text{C16})$$

where $E(e^2)$ is the mathematical expectation of the random variable e^2 . The variance of e is thus:

$$\sigma_e^2 = E[(a_s - a)^2] + E[(a_L - a)^2] - 2E[(a_s - a)(a_L - a)] \quad (\text{C17})$$

which equals:

$$\begin{aligned}
\sigma_e^2(n, k) = & \frac{\sigma_a^2 \sigma_m^2}{n^2 m^2} \left[\left(n + 2 \sum_{i=1}^{n-1} (n-i) \rho_m(ik) \rho_a(ik) \right) \right. \\
& + \frac{1}{n^2} \left(n + 2 \sum_{i=1}^{n-1} (n-i) \rho_a(ik) \right) \left(n + 2 \sum_{i=1}^{n-1} (n-i) \rho_m(ik) \right) \\
& + \frac{1}{k^2} \left(nk + 2 \sum_{i=1}^{nk-1} (nk-i) \rho_m(i) \rho_a(i) \right) \\
& + \frac{1}{n^2 k^4} \left(nk + 2 \sum_{i=1}^{nk-1} (nk-i) \rho_a(i) \right) \left(nk + 2 \sum_{i=1}^{nk-1} (nk-i) \rho_m(i) \right) \\
& - \frac{2}{n} \left(\sum_{p=0}^{n-1} \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} \rho_a((p-j)k) \rho_m((p-i)k) \right) \\
& - \frac{2}{nk^3} \left(\sum_{p=1}^{nk} \sum_{j=1}^{nk} \sum_{i=1}^{nk} \rho_a(p-j) \rho_m(p-i) \right) \\
& - \frac{2}{k} \left(\sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho_a(l+jk-i) \rho_m(l+jk-i) \right) \\
& + \frac{2}{nk^2} \left(\sum_{p=0}^{n-1} \sum_{j=1}^{nk} \sum_{i=1}^{nk} \rho_a(l+pk-j) \rho_m(l+pk-i) \right) \\
& + \frac{2}{nk} \left(\sum_{p=0}^{n-1} \sum_{j=0}^{n-1} \sum_{i=1}^{nk} \rho_a(l+pk-i) \rho_m(l+jk-i) \right) \\
& \left. - \frac{2}{n^2 k^2} \left(\sum_{q=0}^{n-1} \sum_{p=0}^{n-1} \sum_{j=1}^{nk} \sum_{i=1}^{nk} \rho_a(l+qk-j) \rho_m(l+pk-i) \right) \right] \tag{C18}
\end{aligned}$$

where σ_a^2 and $\rho_a(i)$ are respectively the variance and the autocorrelation function of the random variable $\alpha(i)$, and σ_m^2 and $\rho_m(i)$ the variance and the autocorrelation function of the random variable $m(i)$.

If the sample is composed of one increment, the sampling error variance becomes:

$$\sigma_e^2(1, k) = \sigma_a^2(1 + \alpha), \quad k \geq 2 \tag{C19}$$

where α is given by:

$$\begin{aligned}
\alpha = & \frac{\sigma_m^2}{m^2} \left[\frac{1}{k^2} \left(k + 2 \sum_{i=1}^{k-1} (k-i) \rho_m(i) \rho_a(i) \right) \right. \\
& + \frac{1}{k^4} \left(k + 2 \sum_{i=1}^{k-1} (k-i) \rho_a(i) \right) \left(k + 2 \sum_{i=1}^{k-1} (k-i) \rho_m(i) \right) \\
& \left. - \frac{2}{k^3} \left(\sum_{p=1}^k \sum_{j=1}^k \sum_{i=1}^k \rho_a(p-j) \rho_m(p-i) \right) \right] \quad (C20)
\end{aligned}$$

If k equals 1, the following is obtained:

$$\sigma_e^2(1,1) = 0 \quad (C21)$$

In practice, α is negligible compared to 1 (see Figure C1). Therefore Equation C19 becomes:

$$\sigma_e^2(1,k) = \sigma_a^2, \quad k \geq 2 \quad (C22)$$

which indicates that the sampling error variance is described by the composition variance.

In the following development, the composition $a(i)$ is assumed to be $MA(5)$, whereas the mass flowrate $m(i)$ is considered to be $MA(5)$, $MA(15)$, or $MA(45)$.

Figure B2 illustrates the autocorrelation functions of three moving average (MA) processes describing three stream incremental mass flowrates. These stochastic processes are generated following the moving average models (Wei, 1990). In this case, three moving average processes are obtained through $MA(5)$, $MA(15)$, and $MA(45)$ models. The corresponding autocorrelation functions show significant correlation up to a lag of 5, 15, and 45 sampling periods respectively. These lags define the correlation lengths.

Figure C1 illustrates the relative sampling error variance:

$$\sigma_\alpha^2 = \alpha m^2 / \sigma_m^2 \tag{C23}$$

when the sample comprises one increment. This diagram points out the insignificant contribution of α to the sampling error variance.

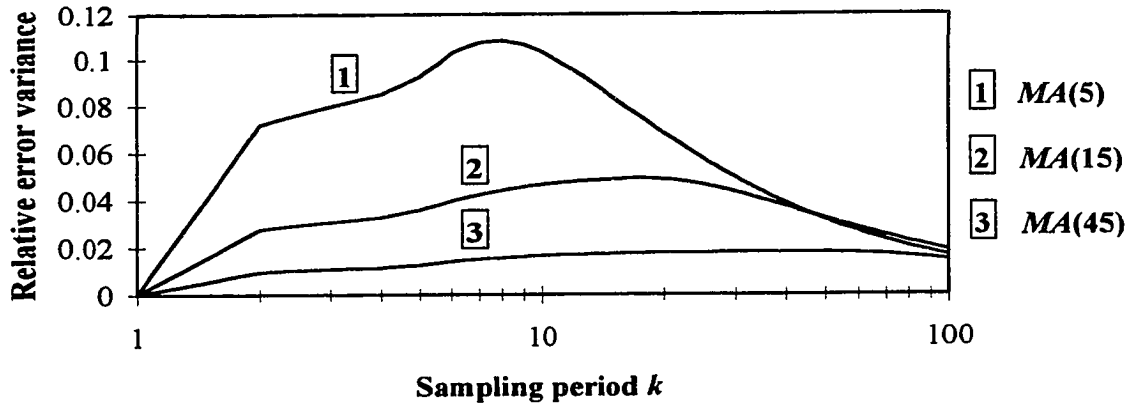


Figure C1. Relative sampling error variance σ_α^2 for a one-increment sample.

Figure C2 exhibits the relative sampling error variance:

$$\sigma_1^2 = \sigma_e^2(1, k) / \sigma_\alpha^2 \tag{C24}$$

Figures C3 and C4 display the relative sampling error variance:

$$\sigma_R^2 = m^2 \sigma_e^2(n, k) / \sigma_\alpha^2 \sigma_m^2 \tag{C25}$$

when the sample consists of five and ten increments respectively.

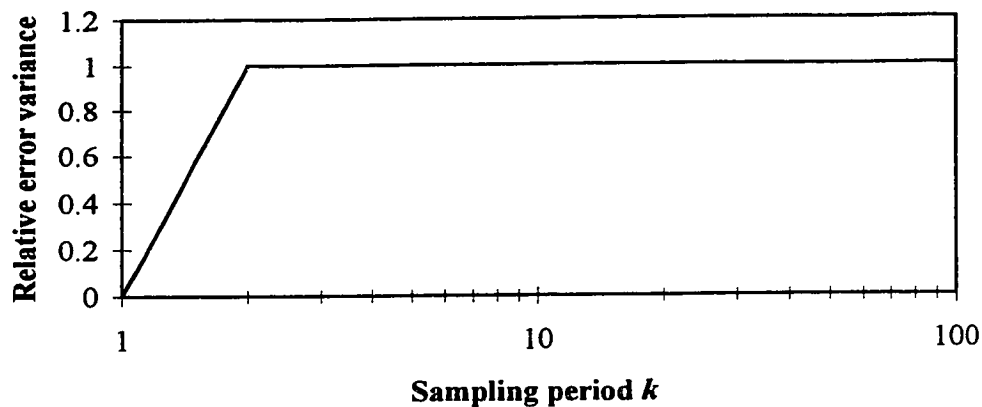


Figure C2. Relative sampling error variance σ_1^2 for a one-increment sample.

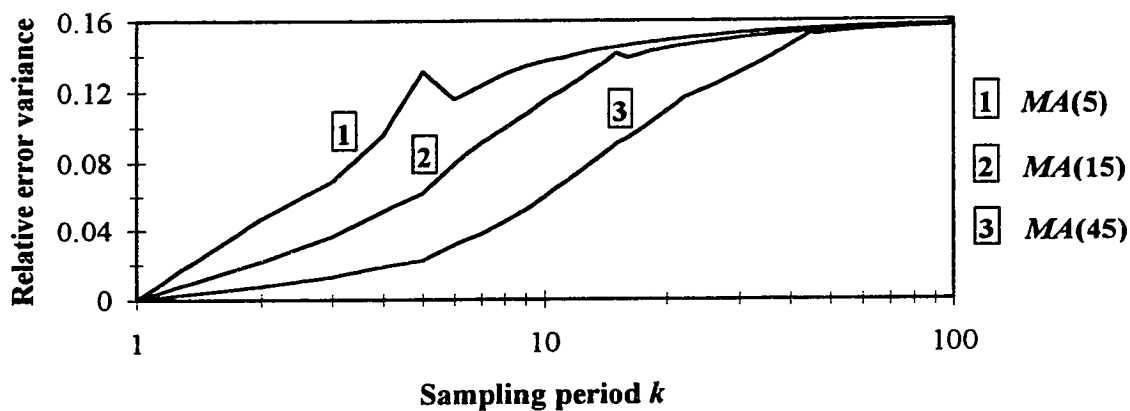


Figure C3. Relative sampling error variance σ_R^2 for a five-increment sample.

These figures demonstrate that the longer the correlation length the lower the sampling error variance. Thus, the samples collected from streams showing higher autocorrelations are more representative. Furthermore, it is revealed that the sampling error variance tends towards the ratio $\sigma_a^2 \sigma_m^2 / m^2 n$. In addition, Figures C3 and C4 show that for more than

one sample increment, oscillatory behavior is generated. This is explained by the periodicity in the systematic sampling strategy.

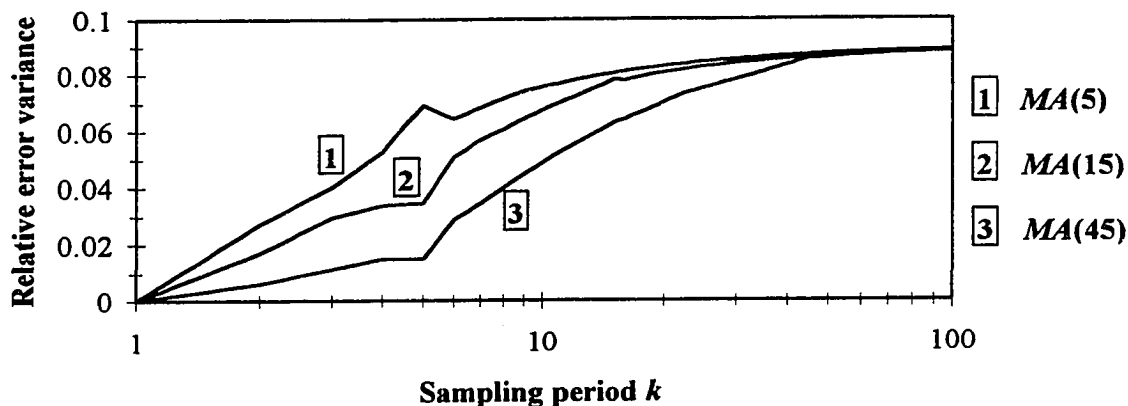


Figure C4. Relative sampling error variance σ_R^2 for a ten-increment sample.

If the mass flowrate $m(i)$ is considered to be $MA(5)$ and the composition $a(i)$ is assumed to be $MA(5)$, $MA(15)$, or $MA(45)$, the same results are obtained respectively.

C7. COMPARISON WITH GY'S METHOD

To evaluate the sampling error, Gy uses the heterogeneity contribution $h(i)$ as a stream descriptor instead of the component composition $a(i)$:

$$h(i) = \frac{a(i) - a_L}{a_L} \frac{m(i)}{m} \quad (\text{C26})$$

The sampling error, given by Equation C7, can be expressed as a function of $h(i)$:

$$e = \frac{a_L m}{\sum_{i=1}^n m(i)} \sum_{i=1}^n h(i) \quad (\text{C27})$$

Moreover, to describe the data variation, he uses the variogram $\nu(k)$ instead of the autocorrelation functions $\rho_a(k)$ and $\rho_m(k)$. The relationship between these tools is:

$$\nu(k) = \frac{\sigma_m^2}{m^2} (1 - \rho_m(k)) + \frac{\sigma_a^2 \sigma_m^2}{a^2 m^2} (1 - \rho_a(k) \rho_m(k)) \quad (\text{C28})$$

The sampling error variance evaluated by Gy is:

$$\sigma_e^2(n, k) = \frac{\sigma_m^2}{na^2 m^2} \sigma_e^2(1, k) \quad (\text{C29})$$

This indicates that Gy neglects the correlation between the increments making up the sample. If $a(i)$ is considered instead of $h(i)$, Equation C29 becomes:

$$\sigma_e^2(n, k) = \frac{1}{n} \sigma_e^2(1, k) \quad (\text{C30})$$

Hence, the term neglected by Gy equals:

$$\beta = \sigma_e^2(n, k) - \frac{1}{n} \sigma_e^2(1, k) \quad (\text{C31})$$

Gy defines an error generator as a function of the sampling period k only:

$$W(k) = \sigma_e^2(1, k) / \sigma_a^2 \sigma_m^2 \quad (\text{C32})$$

However the method proposed in this appendix defines the error generator as a function of both the number of sample increments n and the sampling period k :

$$W(n, k) = n \sigma_e^2(n, k) / \sigma_a^2 \sigma_m^2 \quad (\text{C33})$$

Table C1 summarizes the differences between the proposed method and Gy's procedure for the evaluation of the sampling error.

Figure C5 illustrates the relative error:

$$e_R = \frac{\frac{\sigma_e^2(1, k)}{n} - \sigma_e^2(n, k)}{\frac{\sigma_e^2(1, k)}{n}} \quad (\text{C34})$$

This figure shows that the sampling error variance is overestimated by Gy. In this way, the approximation by Equation C22 is inaccurate.

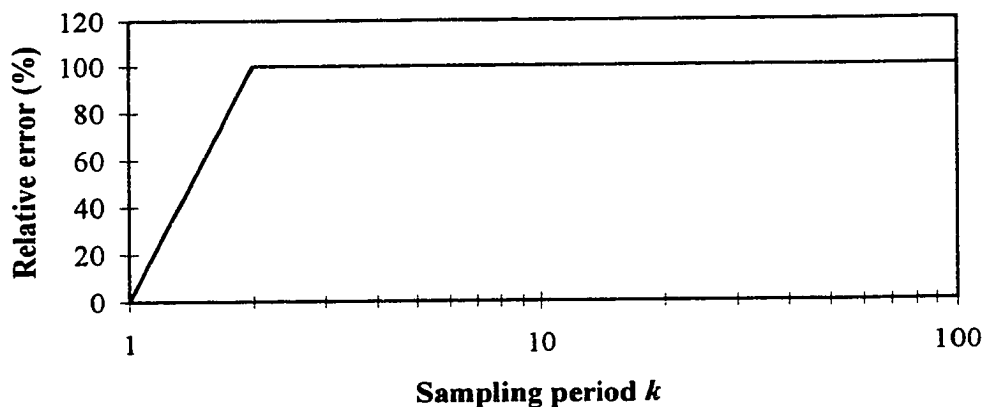


Figure C5. Relative error e_R .

Table C1. Comparison of the proposed method with Gy's procedure.

	Proposed method	Gy's method
Stream descriptor	composition $a(i)$ and mass flowrate $m(i)$	heterogeneity $h(i)$
Data variation descriptor(s)	autocorrelation functions $\rho_a(k)$ and $\rho_m(k)$	variogram $v(k)$
Sampling error	e	$e \left(\sum_{i=1}^n m(i) \right) / nam$
Sampling error variance	$\sigma_e^2(n, k)$	$\sigma_e^2(1, k) \sigma_m^2 / na^2 m^2$
Error generator	$W(n, k)$	$W(k)$

C8. CONCLUSION

A stream is made up of large number of increments. It is sampled for the estimation of its average composition. A small number of increments is collected to compose the sample. Sampling is assumed to be correct. This means that all particles in the stream have a uniform probability of being sampled. Then, the sampling error is random. Therefore, its mean is zero and its variance is different from zero.

The sampling error variance is calculated based on the autocorrelation function used to describe the variation of stream incremental compositions and mass flowrates. The samples collected from streams with higher autocorrelations are more representative. Finally, the proposed method for sampling error evaluation is compared to Gy's approach.

The difference between the proposed formula and Gy's expression for sampling error variance is very high. Obviously, the sampling error variance is overestimated by Gy.

APPENDIX D
KNOWLEDGE BASE OF
THE EXPERT SYSTEM SAMPLING CORRECTNESS INSPECTOR

/*----- Sampling Correctness Inspector Knowledge Base
Copyright (C) 1998 Chefi Ketata -----*/

initialdata = [welcome,end-inspection].

/*----- Sampling Scheme -----*/

question(sampling-scheme) = 'What is your sampling scheme?'.
legalvals(sampling-scheme) = [taking-the-whole-stream-part-of-the-time,taking-part-of-
the-stream-all-of-the-time,taking-part-of-the-stream-part-of-the-time].

automaticmenu(sampling-scheme).

/*----- Cutter Geometry -----*/

question(cutter-geometry) = 'What is your cutter geometry?'.
legalvals(cutter-geometry) = [rectangular,square,circular,triangular,trapezoidal,other].

automaticmenu(cutter-geometry).

/*----- Cutter Strength -----*/

question(cutter-resistant) = 'Is the cutter resistant to the violent impact of large particles?'.
legalvals(cutter-resistant) = [yes,no].

automaticmenu(cutter-resistant).

/*----- Cutter Obstruction -----*/

question(cutter-obstructed) = 'Is the cutter obstructed by sticky materials such as fines?'.
legalvals(cutter-obstructed) = [yes,no].

automaticmenu(cutter-obstructed).

/*----- Cutter Driving System -----*/

question(cutter-driving-system) = 'What is the cutter driving system?'.
legalvals(cutter-driving-system) = [yes,no].

legalvals(cutter-driving-system) = [electric,hydraulic,pneumatic,magnetic>manual].

automaticmenu(cutter-driving-system).

/*----- Cutter Idle Positions -----*/

question(cutter-idle-positions-outside) = 'Are all the cutter idle positions located outside the stream?'

legalvals(cutter-idle-positions-outside) = [yes,no].

automaticmenu(cutter-idle-positions-outside).

/*----- Largest particle diameter -----*/

question(largest-particle-diameter) = 'What is the largest particle diameter in cm?'

legalvals(largest-particle-diameter) = number.

/*----- Minimum Vertical Distance -----*/

if (largest-particle-diameter = D and
 (D <= 0.3) and
 (((3*D)+1) = U0))
 then minimum-vertical-distance = U0.

if (largest-particle-diameter = D and
 (D > 0.3) and
 ((3*D) = U0))
 then minimum-vertical-distance = U0.

/*----- Actual Vertical Distance -----*/

question(stream-cutter-distance) = 'What is the vertical distance, in cm, between the stream and the cutter?'

legalvals(stream-cutter-distance) = number.

/*----- Stream Thickness -----*/

question(stream-thickness) = 'What is the stream thickness in cm?'

legalvals(stream-thickness) = number.

/*----- Minimum Cutter Length -----*/

```

if stream-thickness = T and
  T = L0
then minimum-cutter-length = L0.

```

```

/*----- Actual Cutter Length -----*/

```

```

question(cutter-length) = 'What is the cutter length in cm?'.
legalvals(cutter-length) = number.

```

```

/*----- Cutter Path Normality to the Stream -----*/

```

```

question(cutter-path-normal-to-stream) = 'Is the cutter path normal to the stream?'.
legalvals(cutter-path-normal-to-stream) = [yes,no].

```

```

automaticmenu(cutter-path-normal-to-stream).

```

```

/*----- Minimum Cutter Width -----*/

```

```

if (largest-particle-diameter = D and
  (D <= 0.3) and
  (((3*D)+1) = W0))
then minimum-cutter-width = W0.

```

```

if (largest-particle-diameter = D and
  (D > 0.3) and ((3*D) = W0))
then minimum-cutter-width = W0.

```

```

/*----- Actual Cutter Width -----*/

```

```

question(cutter-width) = 'What is the cutter width in cm?'.
legalvals(cutter-width) = number.

```

```

/*----- Minimum Cutter Depth -----*/

```

```

if (largest-particle-diameter = D and
  (D <= 0.3) and (((3*D)+1) = CD0))
then minimum-cutter-depth = CD0.

```

```

if (largest-particle-diameter = D and
  (D > 0.3) and ((3*D) = CD0))
then minimum-cutter-depth = CD0.

```

```
/*----- Actual Cutter Depth -----*/
```

```
question(cutter-depth) = 'What is the cutter depth in cm?'.
legalvals(cutter-depth) = number.
```

```
/*----- Minimum Cutter Speed -----*/
```

```
minimum-cutter-speed = 0.6.
```

```
/*----- Maximum Cutter Speed -----*/
```

```
if cutter-width = W and
   minimum-cutter-width = W0 and
   ((1+(W/W0))*0.3) = S1 and
   minimum-cutter-speed = S0 and
   S1 > S0
then maximum-cutter-speed = S1.
```

```
if cutter-width = W and
   minimum-cutter-width = W0 and
   ((1+(W/W0))*0.3) = S1 and
   minimum-cutter-speed = S0 and
   S1 <= S0
then maximum-cutter-speed = S0.
```

```
/*----- Actual Cutter Speed -----*/
```

```
question(cutter-speed) = 'What is the cutter speed in m/s?'.
legalvals(cutter-speed) = number.
```

```
/*----- Cutter Speed Constancy -----*/
```

```
question(cutter-speed-constant) = 'Is the cutter speed constant?'.
legalvals(cutter-speed-constant) = [yes,no].
```

```
automaticmenu(cutter-speed-constant).
```

```
/*----- Sampling Scheme Correctness -----*/
```

```
if sampling-scheme = taking-the-whole-stream-part-of-the-time
then correct-sampling-scheme.
```

if sampling-scheme = taking-part-of-the-stream-all-of-the-time or
 sampling-scheme = taking-part-of-the-stream-part-of-the-time
 then correct-sampling-scheme = no.

/*----- Cutter Geometry Correctness -----*/

if cutter-geometry = rectangular or
 cutter-geometry = square
 then correct-cutter-geometry.

if cutter-geometry = circular or
 cutter-geometry = triangular or
 cutter-geometry = trapezoidal or
 cutter-geometry = other
 then correct-cutter-geometry = no.

/*----- Cutter Strength Correctness -----*/

if cutter-resistant
 then cutter-material = strong.

if cutter-material = strong
 then correct-cutter-strength.

if cutter-resistant = no
 then cutter-material = fragile.

if cutter-material = fragile
 then correct-cutter-strength = no.

/*----- Cutter Clearance Correctness -----*/

if cutter-obstructed = no
 then correct-cutter-clearance.

if cutter-obstructed
 then correct-cutter-clearance = no.

/*----- Cutter Driving System Correctness -----*/

if cutter-driving-system = electric
 then correct-cutter-driving-system.

```

if cutter-driving-system = hydraulic or
  cutter-driving-system = pneumatic or
  cutter-driving-system = magnetic or
  cutter-driving-system = manual
then correct-cutter-driving-system = no.

```

```

/*----- Cutter Idle Positions Correctness -----*/

```

```

if cutter-idle-positions-outside
then correct-cutter-idle-positions.

```

```

if cutter-idle-positions-outside = no
then correct-cutter-idle-positions = no.

```

```

/*----- Vertical Distance Correctness -----*/

```

```

if stream-cutter-distance = U and
  minimum-vertical-distance = U0 and
  U >= U0
then correct-stream-cutter-distance.

```

```

if stream-cutter-distance = U and
  minimum-vertical-distance = U0 and
  U < U0
then correct-stream-cutter-distance = no.

```

```

/*----- Cutter Length Correctness -----*/

```

```

if cutter-length = L and
  minimum-cutter-length = L0 and
  L >= L0
then correct-cutter-length.

```

```

if cutter-length = L and
  minimum-cutter-length = L0 and
  L < L0
then correct-cutter-length = no.

```

```

/*----- Cutter Path Correctness -----*/

```

if cutter-path-normal-to-stream
then correct-cutter-path.

if cutter-path-normal-to-stream = no
then correct-cutter-path = no.

/*----- Cutter Width Correctness -----*/

if cutter-width = W and
 minimum-cutter-width = W0 and
 W >= W0
then correct-cutter-width.

if cutter-width = W and
 minimum-cutter-width = W0 and
 W < W0
then correct-cutter-width = no.

/*----- Cutter Depth Correctness -----*/

if cutter-depth = CD and
 minimum-cutter-depth = CD0 and
 CD >= CD0
then correct-cutter-depth.

if cutter-depth = CD and
 minimum-cutter-depth = CD0 and
 CD < CD0
then correct-cutter-depth = no.

/*----- Cutter Speed Correctness -----*/

if cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 > S0 and
 S >= S0 and
 S <= S1 and
 cutter-speed-constant
then correct-cutter-speed.

if cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 > S0 and
 S >= S0 and
 S <= S1 and
 cutter-speed-constant = no
 then correct-cutter-speed = no.

if cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 = S0 and
 S = S0 and
 cutter-speed-constant
 then correct-cutter-speed.

if cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 = S0 and
 S = S0 and
 cutter-speed-constant = no
 then correct-cutter-speed = no.

if cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 > S0 and
 (S < S0 or
 S > S1) and
 cutter-speed-constant
 then correct-cutter-speed = no.

if cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 = S0 and
 (S > S0 or
 S < S0) and
 cutter-speed-constant
 then correct-cutter-speed = no.

```

if cutter-speed = S and
  minimum-cutter-speed = S0 and
  maximum-cutter-speed = S1 and
  S1 > S0 and
  (S < S0 or
  S > S1) and
  cutter-speed-constant = no
then correct-cutter-speed = no.

```

```

if cutter-speed = S and
  minimum-cutter-speed = S0 and
  maximum-cutter-speed = S1 and
  S1 = S0 and
  (S > S0 or
  S < S0) and
  cutter-speed-constant = no
then correct-cutter-speed = no.

```

```

/*----- Conclusion -----*/

```

```

if correct-sampling-scheme = no
then sampling-scheme-conclusion = ['SAMPLING SCHEME:',nl,'The sampling scheme is
not correct.',nl].

```

```

if correct-cutter-geometry = no
then cutter-geometry-conclusion = ['CUTTER GEOMETRY:',nl,'The cutter geometry is
not correct.',nl].

```

```

if correct-cutter-strength = no
then cutter-strength-conclusion = ['CUTTER STRENGTH:',nl,'The cutter strength is not
correct.',nl].

```

```

if correct-cutter-clearance = no
then cutter-clearance-conclusion = ['CUTTER OBSTRUCTION:',nl,'The presence of
sticky material in the cutter is not correct.',nl].

```

```

if correct-cutter-driving-system = no
then cutter-driving-system-conclusion = ['CUTTER DRIVING SYSTEM:',nl,'The cutter
driving system is not correct.',nl].

```

```

if correct-cutter-idle-positions = no
then cutter-idle-positions-conclusion = ['CUTTER IDLE POSITIONS:',nl,'The cutter idle
positions are not correct.',nl].

```

if correct-stream-cutter-distance = no
then vertical-distance-conclusion = ['VERTICAL DISTANCE:',nl,'The vertical distance between the stream and the cutter is not correct.',nl].

if correct-cutter-length = no
then cutter-length-conclusion = ['CUTTER LENGTH:',nl,'The cutter length is not correct.',nl].

if correct-cutter-path = no
then cutter-path-conclusion = ['CUTTER PATH:',nl,'The cutter path is not correct.',nl].

if correct-cutter-width = no
then cutter-width-conclusion = ['CUTTER WIDTH:',nl,'The cutter width is not correct.',nl].

if correct-cutter-depth = no
then cutter-depth-conclusion = ['CUTTER DEPTH:',nl,'The cutter depth is not correct.',nl].

if correct-cutter-speed = no
then cutter-speed-conclusion = ['CUTTER SPEED:',nl,'The cutter speed is not correct.',nl].

if correct-sampling-scheme
then sampling-scheme-conclusion = ['SAMPLING SCHEME:',nl,'The sampling scheme is correct.',nl,nl].

if correct-cutter-geometry
then cutter-geometry-conclusion = ['CUTTER GEOMETRY:',nl,'The cutter geometry is correct.',nl,nl].

if correct-cutter-strength
then cutter-strength-conclusion = ['CUTTER STRENGTH:',nl,'The cutter strength is correct.',nl,nl].

if correct-cutter-clearance
then cutter-clearance-conclusion = ['CUTTER CLEARANCE:',nl,'The absence of sticky material in the cutter is correct.',nl,nl].

if correct-cutter-driving-system
then cutter-driving-system-conclusion = ['CUTTER DRIVING SYSTEM:',nl,'The cutter driving system is correct.',nl,nl].

if correct-cutter-idle-positions
 then cutter-idle-positions-conclusion = ['CUTTER IDLE POSITIONS:',nl,'The cutter idle positions are correct.',nl,nl].

if correct-stream-cutter-distance
 then vertical-distance-conclusion = ['VERTICAL DISTANCE:',nl,'The vertical distance between the stream and the cutter is correct.',nl,nl].

if correct-cutter-length
 then cutter-length-conclusion = ['CUTTER LENGTH:',nl,'The cutter length is correct.',nl,nl].

if correct-cutter-path
 then cutter-path-conclusion = ['CUTTER PATH:',nl,'The cutter path is correct.',nl,nl].

if correct-cutter-width
 then cutter-width-conclusion = ['CUTTER WIDTH:',nl,'The cutter width is correct.',nl,nl].

if correct-cutter-depth
 then cutter-depth-conclusion = ['CUTTER DEPTH:',nl,'The cutter depth is correct.',nl,nl].

if correct-cutter-speed
 then cutter-speed-conclusion = ['CUTTER SPEED:',nl,'The cutter speed is correct.',nl,nl].

/*----- Advice -----*/

if correct-sampling-scheme = no
 then sampling-scheme-advice = ['ADVICE: The sampling scheme should be taking the whole stream part of the time.',nl,nl].

if correct-cutter-geometry = no
 then cutter-geometry-advice = ['ADVICE: The cutter geometry should be rectangular.',nl,nl].

if correct-cutter-strength = no
 then cutter-strength-advice = ['ADVICE: The cutter material should be strong enough to resist to the violent impact of large particles.',nl,nl].

if correct-cutter-clearance = no
 then cutter-clearance-advice = ['ADVICE: Clean the cutter with a water hose to clear it of sticky material.',nl,nl].

if correct-cutter-driving-system = no
 then cutter-driving-system-advice = ['ADVISE: The cutter driving system should be electric.',nl,nl].

if correct-cutter-idle-positions = no
 then cutter-idle-positions-advice = ['ADVISE: The cutter idle positions should be outside the stream.',nl,nl].

if minimum-vertical-distance = U0 and
 correct-stream-cutter-distance = no
 then vertical-distance-advice = ['ADVISE: The vertical distance between the stream and the cutter should be higher than ',U0,' cm.',nl,nl].

if minimum-cutter-length = L0 and
 correct-cutter-length = no
 then cutter-length-advice = ['ADVISE: The cutter length should be higher than ',L0,' cm.',nl,nl].

if correct-cutter-path = no
 then cutter-path-advice = ['ADVISE: The cutter path should be normal to the stream.',nl,nl].

if correct-cutter-width = no and
 minimum-cutter-width = W0
 then cutter-width-advice = ['ADVISE: The cutter width should be higher than ',W0,' cm.',nl,nl].

if correct-cutter-depth = no and
 minimum-cutter-depth = CD0
 then cutter-depth-advice = ['ADVISE: The cutter depth should be higher than ',CD0,' cm.',nl,nl].

if correct-cutter-speed = no and
 cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 > S0 and
 (S > S1 or
 S < S0) and
 cutter-speed-constant

then cutter-speed-advice = ['ADVICE: The cutter speed should be higher than 0.6 m/s and lower than ',S1,' m/s.',nl,nl].

if correct-cutter-speed = no and
 cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 = S0 and
 (S > S0 or
 S < S0) and
 cutter-speed-constant

then cutter-speed-advice = ['ADVICE: The cutter speed should be equal to 0.6 m/s.',nl,nl].

if correct-cutter-speed = no and
 cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 > S0 and
 (S > S1 or
 S < S0) and
 cutter-speed-constant = no

then cutter-speed-advice = ['ADVICE: The cutter speed should be constant, higher than 0.6 m/s, and lower than ',S1,' m/s.',nl,nl].

if correct-cutter-speed = no and
 cutter-speed = S and
 minimum-cutter-speed = S0 and
 maximum-cutter-speed = S1 and
 S1 = S0 and
 S = not(S0) and
 cutter-speed-constant = no

then cutter-speed-advice = ['ADVICE: The cutter speed should be constant and equal to 0.6 m/s.',nl,nl].

/*----- Sampling Correctness -----*/

if correct-sampling-scheme and
 correct-cutter-geometry and
 correct-cutter-strength and
 correct-cutter-clearance and
 correct-cutter-driving-system and
 correct-cutter-idle-positions and
 correct-stream-cutter-distance and

```

correct-cutter-length and
correct-cutter-path and
correct-cutter-width and
correct-cutter-depth and
correct-cutter-speed
then correct-sampling.

```

```

if correct-sampling-scheme = no or
correct-cutter-geometry = no or
correct-cutter-strength = no or
correct-cutter-clearance = no or
correct-cutter-driving-system = no or
correct-cutter-idle-positions = no or
correct-stream-cutter-distance = no or
correct-cutter-length = no or
correct-cutter-path = no or
correct-cutter-width = no or
correct-cutter-depth = no or
correct-cutter-speed = no
then correct-sampling = no.

```

```

/*----- General Conclusion -----*/

```

```

if correct-sampling
then general-conclusion = ['GENERAL CONCLUSION:',nl,'The sampling process is
correct.',nl,nl].

```

```

if correct-sampling = no
then general-conclusion = ['GENERAL CONCLUSION:',nl,'The sampling process is not
correct.',nl,
'Thus the sampling process is biased.',nl,
'This means that the sampling error average is different from zero.',nl,
'To ensure a correct sampling you should act according to the advices given
previously.',nl,nl].

```

```

/*----- Conclusion displayed -----*/

```

```

if sampling-scheme-conclusion = CONCLUSION
then display(CONCLUSION).

```

```

if cutter-geometry-conclusion = CONCLUSION
then display(CONCLUSION).

```

```
if cutter-strength-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-clearance-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-driving-system-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-idle-positions-conclusion = CONCLUSION
then display(CONCLUSION).

if vertical-distance-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-length-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-path-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-width-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-depth-conclusion = CONCLUSION
then display(CONCLUSION).

if cutter-speed-conclusion = CONCLUSION
then display(CONCLUSION).

/*----- Advice displayed -----*/

if correct-sampling-scheme = no and
sampling-scheme-conclusion = CONCLUSION and
display(CONCLUSION) and
sampling-scheme-advice = ADVICE
then display(ADVICE).

if correct-cutter-geometry = no and
cutter-geometry-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-geometry-advice = ADVICE
```


then display(ADVICE).

if correct-cutter-strength = no and
cutter-strength-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-strength-advice = ADVICE
then display(ADVICE).

if correct-cutter-clearance = no and
cutter-clearance-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-clearance-advice = ADVICE
then display(ADVICE).

if correct-cutter-driving-system = no and
cutter-driving-system-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-driving-system-advice = ADVICE
then display(ADVICE).

if correct-cutter-idle-positions = no and
cutter-idle-positions-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-idle-positions-advice = ADVICE
then display(ADVICE).

if correct-stream-cutter-distance = no and
vertical-distance-conclusion = CONCLUSION and
display(CONCLUSION) and
vertical-distance-advice = ADVICE
then display(ADVICE).

if correct-cutter-length = no and
cutter-length-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-length-advice = ADVICE
then display(ADVICE).

if correct-cutter-path = no and
cutter-path-conclusion = CONCLUSION and
display(CONCLUSION) and
cutter-path-advice = ADVICE

then display(ADVICE).

if correct-cutter-width = no and
 cutter-width-conclusion = CONCLUSION and
 display(CONCLUSION) and
 cutter-width-advice = ADVICE
 then display(ADVICE).

if correct-cutter-depth = no and
 cutter-depth-conclusion = CONCLUSION and
 display(CONCLUSION) and
 cutter-depth-advice = ADVICE
 then display(ADVICE).

if correct-cutter-speed = no and
 cutter-speed-conclusion = CONCLUSION and
 display(CONCLUSION) and
 cutter-speed-advice = ADVICE
 then display(ADVICE).

/*----- General Conclusion displayed -----*/

if ((correct-sampling-scheme and
 sampling-scheme-conclusion = CONCLUSION1 and
 display(CONCLUSION1)) or
 (correct-sampling-scheme = no and
 sampling-scheme-conclusion = CONCLUSION1 and
 display(CONCLUSION1) and
 sampling-scheme-advice = ADVICE1 and
 display(ADVICE1))) and
 ((correct-cutter-geometry and
 cutter-geometry-conclusion = CONCLUSION2 and
 display(CONCLUSION2)) or
 (correct-cutter-geometry = no and
 cutter-geometry-conclusion = CONCLUSION2 and
 display(CONCLUSION2) and
 cutter-geometry-advice = ADVICE2 and
 display(ADVICE2))) and
 ((correct-cutter-strength and
 cutter-strength-conclusion = CONCLUSION3 and
 display(CONCLUSION3)) or
 (correct-cutter-strength = no and
 cutter-strength-conclusion = CONCLUSION3 and

display(CONCLUSION3) and
 cutter-strength-advice = ADVICE3 and
 display(ADVICE3))) and
 ((correct-cutter-clearance and
 cutter-clearance-conclusion = CONCLUSION4 and
 display(CONCLUSION4)) or
 (correct-cutter-clearance = no and
 cutter-clearance-conclusion = CONCLUSION4 and
 display(CONCLUSION4) and
 cutter-clearance-advice = ADVICE4 and
 display(ADVICE4))) and
 ((correct-cutter-driving-system and
 cutter-driving-system-conclusion = CONCLUSION5 and
 display(CONCLUSION5)) or
 (correct-cutter-driving-system = no and
 cutter-driving-system-conclusion = CONCLUSION5 and
 display(CONCLUSION5) and
 cutter-driving-system-advice = ADVICE5 and
 display(ADVICE5))) and
 ((correct-cutter-idle-positions and
 cutter-idle-positions-conclusion = CONCLUSION7 and
 display(CONCLUSION7)) or
 (correct-cutter-idle-positions = no and
 cutter-idle-positions-conclusion = CONCLUSION7 and
 display(CONCLUSION7) and
 cutter-idle-positions-advice = ADVICE7 and
 display(ADVICE7))) and
 ((correct-stream-cutter-distance and
 vertical-distance-conclusion = CONCLUSION8 and
 display(CONCLUSION8)) or
 (correct-stream-cutter-distance = no and
 vertical-distance-conclusion = CONCLUSION8 and
 display(CONCLUSION8) and
 vertical-distance-advice = ADVICE8 and
 display(ADVICE8))) and
 ((correct-cutter-length and
 cutter-length-conclusion = CONCLUSION9 and
 display(CONCLUSION9)) or
 (correct-cutter-length = no and
 cutter-length-conclusion = CONCLUSION9 and
 display(CONCLUSION9) and
 cutter-length-advice = ADVICE9 and
 display(ADVICE9))) and

```

((correct-cutter-path and
 cutter-path-conclusion = CONCLUSION10 and
 display(CONCLUSION10)) or
 (correct-cutter-path = no and
 cutter-path-conclusion = CONCLUSION10 and
 display(CONCLUSION10) and
 cutter-path-advice = ADVICE10 and
 display(ADVICE10))) and
 ((correct-cutter-width and
 cutter-width-conclusion = CONCLUSION11 and
 display(CONCLUSION11)) or
 (correct-cutter-width = no and
 cutter-width-conclusion = CONCLUSION11 and
 display(CONCLUSION11) and
 cutter-width-advice = ADVICE11 and
 display(ADVICE11))) and
 ((correct-cutter-depth and
 cutter-depth-conclusion = CONCLUSION12 and
 display(CONCLUSION12)) or
 (correct-cutter-depth = no and
 cutter-depth-conclusion = CONCLUSION12 and
 display(CONCLUSION12) and
 cutter-depth-advice = ADVICE12 and
 display(ADVICE12))) and
 ((correct-cutter-speed and
 cutter-speed-conclusion = CONCLUSION6 and
 display(CONCLUSION6)) or
 (correct-cutter-speed = no and
 cutter-speed-conclusion = CONCLUSION6 and
 display(CONCLUSION6) and
 cutter-speed-advice = ADVICE6 and
 display(ADVICE6))) and
 general-conclusion = GCONCLUSION
then display(GCONCLUSION).

/*----- User Informed -----*/

if ((correct-sampling-scheme and
 sampling-scheme-conclusion = CONCLUSION1 and
 display(CONCLUSION1)) or
 (correct-sampling-scheme = no and
 sampling-scheme-conclusion = CONCLUSION1 and
 display(CONCLUSION1) and

```

sampling-scheme-advice = ADVICE1 and
 display(ADVICE1))) and
 ((correct-cutter-geometry and
 cutter-geometry-conclusion = CONCLUSION2 and
 display(CONCLUSION2)) or
 (correct-cutter-geometry = no and
 cutter-geometry-conclusion = CONCLUSION2 and
 display(CONCLUSION2) and
 cutter-geometry-advice = ADVICE2 and
 display(ADVICE2))) and
 ((correct-cutter-strength and
 cutter-strength-conclusion = CONCLUSION3 and
 display(CONCLUSION3)) or
 (correct-cutter-strength = no and
 cutter-strength-conclusion = CONCLUSION3 and
 display(CONCLUSION3) and
 cutter-strength-advice = ADVICE3 and
 display(ADVICE3))) and
 ((correct-cutter-clearance and
 cutter-clearance-conclusion = CONCLUSION4 and
 display(CONCLUSION4)) or
 (correct-cutter-clearance = no and
 cutter-clearance-conclusion = CONCLUSION4 and
 display(CONCLUSION4) and
 cutter-clearance-advice = ADVICE4 and
 display(ADVICE4))) and
 ((correct-cutter-driving-system and
 cutter-driving-system-conclusion = CONCLUSION5 and
 display(CONCLUSION5)) or
 (correct-cutter-driving-system = no and
 cutter-driving-system-conclusion = CONCLUSION5 and
 display(CONCLUSION5) and
 cutter-driving-system-advice = ADVICE5 and
 display(ADVICE5))) and
 ((correct-cutter-idle-positions and
 cutter-idle-positions-conclusion = CONCLUSION7 and
 display(CONCLUSION7)) or
 (correct-cutter-idle-positions = no and
 cutter-idle-positions-conclusion = CONCLUSION7 and
 display(CONCLUSION7) and
 cutter-idle-positions-advice = ADVICE7 and
 display(ADVICE7))) and
 ((correct-stream-cutter-distance and

vertical-distance-conclusion = CONCLUSION8 and
 display(CONCLUSION8)) or
 (correct-stream-cutter-distance = no and
 vertical-distance-conclusion = CONCLUSION8 and
 display(CONCLUSION8) and
 vertical-distance-advice = ADVICE8 and
 display(ADVICE8))) and
 ((correct-cutter-length and
 cutter-length-conclusion = CONCLUSION9 and
 display(CONCLUSION9)) or
 (correct-cutter-length = no and
 cutter-length-conclusion = CONCLUSION9 and
 display(CONCLUSION9) and
 cutter-length-advice = ADVICE9 and
 display(ADVICE9))) and
 ((correct-cutter-path and
 cutter-path-conclusion = CONCLUSION10 and
 display(CONCLUSION10)) or
 (correct-cutter-path = no and
 cutter-path-conclusion = CONCLUSION10 and
 display(CONCLUSION10) and
 cutter-path-advice = ADVICE10 and
 display(ADVICE10))) and
 ((correct-cutter-width and
 cutter-width-conclusion = CONCLUSION11 and
 display(CONCLUSION11)) or
 (correct-cutter-width = no and
 cutter-width-conclusion = CONCLUSION11 and
 display(CONCLUSION11) and
 cutter-width-advice = ADVICE11 and
 display(ADVICE11))) and
 ((correct-cutter-depth and
 cutter-depth-conclusion = CONCLUSION12 and
 display(CONCLUSION12)) or
 (correct-cutter-depth = no and
 cutter-depth-conclusion = CONCLUSION12 and
 display(CONCLUSION12) and
 cutter-depth-advice = ADVICE12 and
 display(ADVICE12))) and
 ((correct-cutter-speed and
 cutter-speed-conclusion = CONCLUSION6 and
 display(CONCLUSION6)) or
 (correct-cutter-speed = no and

```
cutter-speed-conclusion = CONCLUSION6 and
display(CONCLUSION6) and
cutter-speed-advice = ADVICE6 and
display(ADVICE6))) and
general-conclusion = GCONCLUSION and
display(GCONCLUSION)
then user-informed.
```

```
/*----- End Inspection -----*/
```

```
if user-informed and
  display(['----- END -----
-----',nl,nl])
then end-inspection.
```

```
/*----- Welcome -----*/
```

```
if display(['----- SAMPLING PROCESS INSPECTOR -----
-----',nl,nl])
then welcome.
```

APPENDIX E
VALIDATION OF
THE EXPERT SYSTEM SAMPLING CORRECTNESS INSPECTOR

E1. CASE STUDY NUMBER 1

"----- SAMPLING PROCESS INSPECTOR -----"
"
"
"
"What is your sampling scheme?
"
" taking-the-whole-stream-part-of-the-time
"
" taking-part-of-the-stream-all-of-the-time
"
" taking-part-of-the-stream-part-of-the-time
"
">"taking-the-whole-stream-part-of-the-time"
"
"SAMPLING SCHEME:
"
"The sampling scheme is correct.
"
"
"
"What is your cutter geometry?
"
" rectangular
"
" square
"
" circular
"
" triangular
"
" trapezoidal
"
" other
"

">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic

"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"
"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">10
"
"What is the largest particle diameter in cm?
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct.
"
"
"
"What is the cutter length in cm?
"

">10
"
"What is the stream thickness in cm?
"
">7
"
"CUTTER LENGTH:
"
"The cutter length is correct.
"
"
"
"Is the cutter path normal to the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.

```
"  
"  
"  
"What is the cutter speed in m/s?  
"  
">0.6  
"  
"Is the cutter speed constant?  
"  
" yes  
"  
" no  
"  
">"yes"  
"  
"CUTTER SPEED:  
"  
"The cutter speed is correct.  
"  
"  
"  
"GENERAL CONCLUSION:  
"  
"The sampling process is correct.  
"  
"  
"  
"----- END -----  
---  
"  
"  
"
```

E2. CASE STUDY NUMBER 2

```
"----- SAMPLING PROCESS INSPECTOR -----  
"  
"  
"  
"What is your sampling scheme?  
"
```

" taking-the-whole-stream-part-of-the-time
"
" taking-part-of-the-stream-all-of-the-time
"
" taking-part-of-the-stream-part-of-the-time
"
">"taking-part-of-the-stream-all-of-the-time"
"
"SAMPLING SCHEME:
"
"The sampling scheme is not correct.
"
"ADVICE: The sampling scheme should be taking the whole stream part of the time.
"
"
"
"What is your cutter geometry?
"
" rectangular
"
" square
"
" circular
"
" triangular
"
" trapezoidal
"
" other
"
">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no

"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"

"Are all the cutter idle positions located outside the stream?"

"

" yes

"

" no

"

">"yes"

"

"CUTTER IDLE POSITIONS:

"

"The cutter idle positions are correct.

"

"

"

"What is the vertical distance, in cm, between the stream and the cutter?"

"

">10

"

"What is the largest particle diameter in cm?"

"

">0.02

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is correct.

"

"

"

"What is the cutter length in cm?"

"

">10

"

"What is the stream thickness in cm?"

"

">7

"

"CUTTER LENGTH:

"

"The cutter length is correct.

"

"

"

"Is the cutter path normal to the stream?"

"
" yes
"
" no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.
"
"
"
"What is the cutter speed in m/s?
"
">0.6
"
"Is the cutter speed constant?
"
" yes
"
" no
"

">"yes"

"

"CUTTER SPEED:

"

"The cutter speed is correct.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"----- END -----

"

"

"

E3. CASE STUDY NUMBER 3

"----- SAMPLING PROCESS INSPECTOR -----

"

"

"

"What is your sampling scheme?

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"
"SAMPLING SCHEME:
"
"The sampling scheme is correct.
"
"
"
"What is your cutter geometry?
"
" rectangular
"
" square
"
" circular
"
" triangular
"
" trapezoidal
"
" other
"
">"circular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is not correct.
"
"ADVICE: The cutter geometry should be rectangular.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"

"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"

"

"CUTTER IDLE POSITIONS:

"

"The cutter idle positions are correct.

"

"

"

"What is the vertical distance, in cm, between the stream and the cutter?"

"

">10

"

"What is the largest particle diameter in cm?"

"

">0.02

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is correct.

"

"

"

"What is the cutter length in cm?"

"

">10

"

"What is the stream thickness in cm?"

"

">6

"

"CUTTER LENGTH:

"

"The cutter length is correct.

"

"

"

"Is the cutter path normal to the stream?"

"

" yes

"

" no

"

">"yes"

"

"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.
"
"
"
"What is the cutter speed in m/s?
"
">0.6
"
"Is the cutter speed constant?
"
" yes
"
" no
"
">"yes"
"
"CUTTER SPEED:
"
"The cutter speed is correct.
"
"
"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"----- END -----"

"

"

"

E4. CASE STUDY NUMBER 4

"----- SAMPLING PROCESS INSPECTOR -----"

"

"

"

"What is your sampling scheme?

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"

"SAMPLING SCHEME:

"

"The sampling scheme is correct.

"

"

"

"What is your cutter geometry?"

"

" rectangular

"

" square

"

" circular

"

" triangular

"

" trapezoidal

"

" other

"

">"rectangular"

"

"CUTTER GEOMETRY:

"

"The cutter geometry is correct.

"

"

"

"Is the cutter resistant to the violent impact of large particles?"

"

" yes

"

" no

"

">"no"

"

"CUTTER STRENGTH:

"

"The cutter strength is not correct.

"

"ADVICE: The cutter material should be strong enough to resist to the violent impact of large particles.

"

"

"

"Is the cutter obstructed by sticky materials such as fines?"

"

" yes

"

" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"

"
"What is the vertical distance, in cm, between the stream and the cutter?"
"
">8
"
"What is the largest particle diameter in cm?"
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct."
"
"
"
"What is the cutter length in cm?"
"
">10
"
"What is the stream thickness in cm?"
"
">6
"
"CUTTER LENGTH:
"
"The cutter length is correct."
"
"
"
"Is the cutter path normal to the stream?"
"
" yes
"
" no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct."
"
"
"

"What is the cutter width in cm?

"

">2

"

"CUTTER WIDTH:

"

"The cutter width is correct.

"

"

"

"What is the cutter depth in cm?

"

">5

"

"CUTTER DEPTH:

"

"The cutter depth is correct.

"

"

"

"What is the cutter speed in m/s?

"

">0.6

"

"Is the cutter speed constant?

"

" yes

"

" no

"

">"yes"

"

"CUTTER SPEED:

"

"The cutter speed is correct.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"
 "This means that the sampling error average is different from zero.
 "
 "To ensure a correct sampling you should act according to the advices given previously.
 "
 "
 "
 "----- END -----
 --
 "
 "
 "

E5. CASE STUDY NUMBER 5

"----- SAMPLING PROCESS INSPECTOR -----
 "
 "
 "
 "What is your sampling scheme?
 "
 " taking-the-whole-stream-part-of-the-time
 "
 " taking-part-of-the-stream-all-of-the-time
 "
 " taking-part-of-the-stream-part-of-the-time
 "
 ">"taking-the-whole-stream-part-of-the-time"
 "
 "SAMPLING SCHEME:
 "
 "The sampling scheme is correct.
 "
 "
 "
 "What is your cutter geometry?
 "
 " rectangular
 "
 " square
 "

" circular
"
" triangular
"
" trapezoidal
"
" other
"
">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"yes"
"
"CUTTER OBSTRUCTION:
"
"The presence of sticky material in the cutter is not correct.
"
"ADVICE: Clean the cutter with a water hose to clear it of sticky material.

"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"
"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">7
"
"What is the largest particle diameter in cm?
"

">0.02

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is correct.

"

"

"

"What is the cutter length in cm?

"

">10

"

"What is the stream thickness in cm?

"

">7

"

"CUTTER LENGTH:

"

"The cutter length is correct.

"

"

"

"Is the cutter path normal to the stream?

"

" yes

"

" no

"

">"yes"

"

"CUTTER PATH:

"

"The cutter path is correct.

"

"

"

"What is the cutter width in cm?

"

">2

"

"CUTTER WIDTH:

"

"The cutter width is correct.

"

"

"

"What is the cutter depth in cm?

"

">5

"

"CUTTER DEPTH:

"

"The cutter depth is correct.

"

"

"

"What is the cutter speed in m/s?

"

">0.6

"

"Is the cutter speed constant?

"

" yes

"

" no

"

">"yes"

"

"CUTTER SPEED:

"

"The cutter speed is correct.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"-----END-----"

 "
 "
 "

E6. CASE STUDY NUMBER 6

"----- SAMPLING PROCESS INSPECTOR -----"
 "
 "
 "
 "What is your sampling scheme?
 "
 " taking-the-whole-stream-part-of-the-time
 "
 " taking-part-of-the-stream-all-of-the-time
 "
 " taking-part-of-the-stream-part-of-the-time
 "
 ">"taking-the-whole-stream-part-of-the-time"
 "
 "SAMPLING SCHEME:
 "
 "The sampling scheme is correct.
 "
 "
 "
 "What is your cutter geometry?
 "
 " rectangular
 "
 " square
 "
 " circular
 "
 " triangular
 "
 " trapezoidal
 "
 " other

"
">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"

" pneumatic
"
" magnetic
"
" manual
"
">"hydraulic"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is not correct.
"
"ADVICE: The cutter driving system should be electric.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"
"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">10
"
"What is the largest particle diameter in cm?
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct.
"
"
"

"
"What is the cutter length in cm?
"
">10
"
"What is the stream thickness in cm?
"
">7
"
"CUTTER LENGTH:
"
"The cutter length is correct.
"
"
"
"Is the cutter path normal to the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"

"CUTTER DEPTH:

"

"The cutter depth is correct.

"

"

"

"What is the cutter speed in m/s?

"

">0.6

"

"Is the cutter speed constant?

"

" yes

"

" no

"

">"yes"

"

"CUTTER SPEED:

"

"The cutter speed is correct.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"----- END -----"

"

"

"

"

E7. CASE STUDY NUMBER 7

"----- SAMPLING PROCESS INSPECTOR -----"
"
"
"
"What is your sampling scheme?
"
" taking-the-whole-stream-part-of-the-time
"
" taking-part-of-the-stream-all-of-the-time
"
" taking-part-of-the-stream-part-of-the-time
"
">"taking-the-whole-stream-part-of-the-time"
"
"SAMPLING SCHEME:
"
"The sampling scheme is correct.
"
"
"
"What is your cutter geometry?
"
" rectangular
"
" square
"
" circular
"
" triangular
"
" trapezoidal
"
" other
"
">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"

"
"
"Is the cutter resistant to the violent impact of large particles?"
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?"
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?"
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"

"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"no"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are not correct.
"
"ADVICE: The cutter idle positions should be outside the stream.
"
"
"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">10
"
"What is the largest particle diameter in cm?
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct.
"
"
"
"What is the cutter length in cm?
"
">10
"
"What is the stream thickness in cm?
"

">8
"
"CUTTER LENGTH:
"
"The cutter length is correct.
"
"
"
"Is the cutter path normal to the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.
"
"
"
"What is the cutter speed in m/s?


```

"
">0.6
"
"Is the cutter speed constant?
"
"  yes
"
"  no
"
">"yes"
"
"CUTTER SPEED:
"
"The cutter speed is correct.
"
"
"
"GENERAL CONCLUSION:
"
"The sampling process is not correct.
"
"Thus the sampling process is biased.
"
"This means that the sampling error average is different from zero.
"
"To ensure a correct sampling you should act according to the advices given previously.
"
"
"
"----- END -----
---
"
"
"

```

E8. CASE STUDY NUMBER 8

```

"----- SAMPLING PROCESS INSPECTOR -----
"
"
"

```

"What is your sampling scheme?

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"

"SAMPLING SCHEME:

"

"The sampling scheme is correct.

"

"

"

"What is your cutter geometry?

"

" rectangular

"

" square

"

" circular

"

" triangular

"

" trapezoidal

"

" other

"

">"rectangular"

"

"CUTTER GEOMETRY:

"

"The cutter geometry is correct.

"

"

"

"Is the cutter resistant to the violent impact of large particles?

"

" yes

"

" no

"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"

"Are all the cutter idle positions located outside the stream?"

"

" yes

"

" no

"

">"yes"

"

"CUTTER IDLE POSITIONS:

"

"The cutter idle positions are correct.

"

"

"

"What is the vertical distance, in cm, between the stream and the cutter?"

"

">1

"

"What is the largest particle diameter in cm?"

"

">0.05

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is not correct.

"

"ADVICE: The vertical distance between the stream and the cutter should be higher than 1.150 cm.

"

"

"

"What is the cutter length in cm?"

"

">10

"

"What is the stream thickness in cm?"

"

">7

"

"CUTTER LENGTH:

"

"The cutter length is correct.

"

"
"
"Is the cutter path normal to the stream?"
"
" yes
"
" no
"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct."
"
"
"
"What is the cutter width in cm?"
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct."
"
"
"
"What is the cutter depth in cm?"
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct."
"
"
"
"What is the cutter speed in m/s?"
"
">0.6
"
"Is the cutter speed constant?"
"
" yes

"

" no

"

">"yes"

"

"CUTTER SPEED:

"

"The cutter speed is correct.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"----- END -----"

"

"

"

E9. CASE STUDY NUMBER 9

"----- SAMPLING PROCESS INSPECTOR -----"

"

"

"

"What is your sampling scheme?

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"

"SAMPLING SCHEME:

"

"The sampling scheme is correct.

"

"

"

"What is your cutter geometry?

"

" rectangular

"

" square

"

" circular

"

" triangular

"

" trapezoidal

"

" other

"

">"rectangular"

"

"CUTTER GEOMETRY:

"

"The cutter geometry is correct.

"

"

"

"Is the cutter resistant to the violent impact of large particles?

"

" yes

"

" no

"

">"yes"

"

"CUTTER STRENGTH:

"

"The cutter strength is correct.

"

"

"

"Is the cutter obstructed by sticky materials such as fines?"

"

" yes

"

" no

"

">"no"

"

"CUTTER CLEARANCE:

"

"The absence of sticky material in the cutter is correct.

"

"

"

"What is the cutter driving system?"

"

" electric

"

" hydraulic

"

" pneumatic

"

" magnetic

"

" manual

"

">"electric"

"

"CUTTER DRIVING SYSTEM:

"

"The cutter driving system is correct.

"

"

"

"Are all the cutter idle positions located outside the stream?"

"

" yes

"

" no

"

">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"
"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">10
"
"What is the largest particle diameter in cm?
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct.
"
"
"
"What is the cutter length in cm?
"
">3
"
"What is the stream thickness in cm?
"
">4
"
"CUTTER LENGTH:
"
"The cutter length is not correct.
"
"ADVICE: The cutter length should be higher than 4 cm.
"
"
"
"Is the cutter path normal to the stream?
"
" yes
"
" no

"
">"yes"
"
"CUTTER PATH:
"
"The cutter path is correct.
"
"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.
"
"
"
"What is the cutter speed in m/s?
"
">0.6
"
"Is the cutter speed constant?
"
" yes
"
" no
"
">"yes"
"
"CUTTER SPEED:
"

"The cutter speed is correct.

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"----- END -----

"

"

"

E10. CASE STUDY NUMBER 10

"----- SAMPLING PROCESS INSPECTOR -----

"

"

"

"What is your sampling scheme?

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"

"SAMPLING SCHEME:

"

"The sampling scheme is correct.

"
"
"

"What is your cutter geometry?"

"

" rectangular

"

" square

"

" circular

"

" triangular

"

" trapezoidal

"

" other

"

">"rectangular"

"

"CUTTER GEOMETRY:

"

"The cutter geometry is correct.

"

"

"

"Is the cutter resistant to the violent impact of large particles?"

"

" yes

"

" no

"

">"yes"

"

"CUTTER STRENGTH:

"

"The cutter strength is correct.

"

"

"

"Is the cutter obstructed by sticky materials such as fines?"

"

" yes

"

" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"
"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"
"

"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">10
"
"What is the largest particle diameter in cm?
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct.
"
"
"
"What is the cutter length in cm?
"
">10
"
"What is the stream thickness in cm?
"
">5
"
"CUTTER LENGTH:
"
"The cutter length is correct.
"
"
"
"Is the cutter path normal to the stream?
"
" yes
"
" no
"
">"no"
"
"CUTTER PATH:
"
"The cutter path is not correct.
"
"ADVICE: The cutter path should be normal to the stream.
"

"
"
"What is the cutter width in cm?
"
">2
"
"CUTTER WIDTH:
"
"The cutter width is correct.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.
"
"
"
"What is the cutter speed in m/s?
"
">0.6
"
"Is the cutter speed constant?
"
" yes
"
" no
"
">"yes"
"
"CUTTER SPEED:
"
"The cutter speed is correct.
"
"
"
"GENERAL CONCLUSION:
"
"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"----- END -----

"

"

"

E11. CASE STUDY NUMBER 11

"----- SAMPLING PROCESS INSPECTOR -----

"

"

"

"What is your sampling scheme?

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"

"SAMPLING SCHEME:

"

"The sampling scheme is correct.

"

"

"

"What is your cutter geometry?

"

" rectangular

"

" square

"

" circular

"

" triangular

"

" trapezoidal

"

" other

"

">"rectangular"

"

"CUTTER GEOMETRY:

"

"The cutter geometry is correct.

"

"

"

"Is the cutter resistant to the violent impact of large particles?"

"

" yes

"

" no

"

">"yes"

"

"CUTTER STRENGTH:

"

"The cutter strength is correct.

"

"

"

"Is the cutter obstructed by sticky materials such as fines?"

"

" yes

"

" no

"

">"no"

"

"CUTTER CLEARANCE:

"

"The absence of sticky material in the cutter is correct.

"

"

"

"What is the cutter driving system?

"

" electric

"

" hydraulic

"

" pneumatic

"

" magnetic

"

" manual

"

">"electric"

"

"CUTTER DRIVING SYSTEM:

"

"The cutter driving system is correct.

"

"

"

"Are all the cutter idle positions located outside the stream?

"

" yes

"

" no

"

">"yes"

"

"CUTTER IDLE POSITIONS:

"

"The cutter idle positions are correct.

"

"

"

"What is the vertical distance, in cm, between the stream and the cutter?

"

">10

"

"What is the largest particle diameter in cm?

"

">0.05

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is correct.

"

"

"

"What is the cutter length in cm?

"

">10

"

"What is the stream thickness in cm?

"

">5

"

"CUTTER LENGTH:

"

"The cutter length is correct.

"

"

"

"Is the cutter path normal to the stream?

"

" yes

"

" no

"

">"yes"

"

"CUTTER PATH:

"

"The cutter path is correct.

"

"

"

"What is the cutter width in cm?

"

">0.5

"

"CUTTER WIDTH:

"

"The cutter width is not correct.

"
"ADVICE: The cutter width should be higher than 1.150 cm.
"
"
"
"What is the cutter depth in cm?
"
">5
"
"CUTTER DEPTH:
"
"The cutter depth is correct.
"
"
"
"What is the cutter speed in m/s?
"
">0.6
"
"Is the cutter speed constant?
"
" yes
"
" no
"
">"yes"
"
"CUTTER SPEED:
"
"The cutter speed is correct.
"
"
"
"GENERAL CONCLUSION:
"
"The sampling process is not correct.
"
"Thus the sampling process is biased.
"
"This means that the sampling error average is different from zero.
"
"To ensure a correct sampling you should act according to the advices given previously.
"

"
"
"
"
"
"
"

----- END -----

E12. CASE STUDY NUMBER 12

"----- SAMPLING PROCESS INSPECTOR -----"

"
"
"

"What is your sampling scheme?"

"

" taking-the-whole-stream-part-of-the-time

"

" taking-part-of-the-stream-all-of-the-time

"

" taking-part-of-the-stream-part-of-the-time

"

">"taking-the-whole-stream-part-of-the-time"

"

"SAMPLING SCHEME:"

"

"The sampling scheme is correct.

"

"

"

"What is your cutter geometry?"

"

" rectangular

"

" square

"

" circular

"

" triangular

"

" trapezoidal

"
" other
"
">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"
"
"
"Is the cutter resistant to the violent impact of large particles?
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct.
"
"
"
"Is the cutter obstructed by sticky materials such as fines?
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct.
"
"
"
"What is the cutter driving system?
"
" electric
"

" hydraulic

"

" pneumatic

"

" magnetic

"

" manual

"

">"electric"

"

"CUTTER DRIVING SYSTEM:

"

"The cutter driving system is correct.

"

"

"

"Are all the cutter idle positions located outside the stream?"

"

" yes

"

" no

"

">"yes"

"

"CUTTER IDLE POSITIONS:

"

"The cutter idle positions are correct.

"

"

"

"What is the vertical distance, in cm, between the stream and the cutter?"

"

">10

"

"What is the largest particle diameter in cm?"

"

">0.05

"

"VERTICAL DISTANCE:

"

"The vertical distance between the stream and the cutter is correct.

"

"

"CUTTER DEPTH:

"

"The cutter depth is not correct.

"

"ADVICE: The cutter depth should be higher than 1.150 cm.

"

"

"

"What is the cutter speed in m/s?

"

">0.6

"

"Is the cutter speed constant?

"

" yes

"

" no

"

">"yes"

"

"CUTTER SPEED:

"

"The cutter speed is correct.

"

"

"

"GENERAL CONCLUSION:

"

"The sampling process is not correct.

"

"Thus the sampling process is biased.

"

"This means that the sampling error average is different from zero.

"

"To ensure a correct sampling you should act according to the advices given previously.

"

"

"

"-----END-----"

"

"

"

"

E13. CASE STUDY NUMBER 13

"----- SAMPLING PROCESS INSPECTOR -----"
"
"
"
"What is your sampling scheme?
"
" taking-the-whole-stream-part-of-the-time
"
" taking-part-of-the-stream-all-of-the-time
"
" taking-part-of-the-stream-part-of-the-time
"
">"taking-the-whole-stream-part-of-the-time"
"
"SAMPLING SCHEME:
"
"The sampling scheme is correct.
"
"
"
"What is your cutter geometry?
"
" rectangular
"
" square
"
" circular
"
" triangular
"
" trapezoidal
"
" other
"
">"rectangular"
"
"CUTTER GEOMETRY:
"
"The cutter geometry is correct.
"

"
"
"Is the cutter resistant to the violent impact of large particles?"
"
" yes
"
" no
"
">"yes"
"
"CUTTER STRENGTH:
"
"The cutter strength is correct."
"
"
"
"Is the cutter obstructed by sticky materials such as fines?"
"
" yes
"
" no
"
">"no"
"
"CUTTER CLEARANCE:
"
"The absence of sticky material in the cutter is correct."
"
"
"
"What is the cutter driving system?"
"
" electric
"
" hydraulic
"
" pneumatic
"
" magnetic
"
" manual
"
">"electric"

"
"CUTTER DRIVING SYSTEM:
"
"The cutter driving system is correct.
"
"
"
"Are all the cutter idle positions located outside the stream?
"
" yes
"
" no
"
">"yes"
"
"CUTTER IDLE POSITIONS:
"
"The cutter idle positions are correct.
"
"
"
"What is the vertical distance, in cm, between the stream and the cutter?
"
">10
"
"What is the largest particle diameter in cm?
"
">0.02
"
"VERTICAL DISTANCE:
"
"The vertical distance between the stream and the cutter is correct.
"
"
"
"What is the cutter length in cm?
"
">10
"
"What is the stream thickness in cm?
"
">5
"

"CUTTER LENGTH:

"

"The cutter length is correct.

"

"

"

"Is the cutter path normal to the stream?

"

" yes

"

" no

"

">"yes"

"

"CUTTER PATH:

"

"The cutter path is correct.

"

"

"

"What is the cutter width in cm?

"

">2

"

"CUTTER WIDTH:

"

"The cutter width is correct.

"

"

"

"What is the cutter depth in cm?

"

">5

"

"CUTTER DEPTH:

"

"The cutter depth is not correct.

"

"ADVICE: The cutter depth should be higher than 1.150 cm.

"

"

"

"What is the cutter speed in m/s?

"
">1
"
"Is the cutter speed constant?
"
" yes
"
" no
"
">"yes"
"
"CUTTER SPEED:
"
"The cutter speed is not correct.
"
"ADVICE: The cutter speed should be higher than 0.6 m/s and lower than 0.866038 m/s.
"
"
"
"GENERAL CONCLUSION:
"
"The sampling process is not correct.
"
"Thus the sampling process is biased.
"
"This means that the sampling error average is different from zero.
"
"To ensure a correct sampling you should act according to the advices given previously.
"
"
"
"----- END -----"
"
"
"

APPENDIX F
KNOWLEDGE BASE OF
THE EXPERT SYSTEM SAMPLING ERROR EVALUATOR

/*----- Sampling Error Evaluator Knowledge Base
Copyright (C) 1998 Chefi Ketata -----*/

initialdata = [welcome,increment-mass-displayed,sampling-error1-displayed,sampling-period-displayed,sampling-error-displayed,minimum-increment-number-displayed,maximum-sampling-period-displayed,maximum-sampling-error-displayed,evaluation-end].

/*----- Shape Factor F -----*/

question(particle-shape) = 'What is the shape of the particles?'.
legalvals(particle-shape) = [cube,sphere,flake,nugget,needle].

automaticmenu(particle-shape).

if particle-shape = cube
then shape-factor = 1.

if particle-shape = sphere
then shape-factor = 0.523.

if particle-shape = flake
then shape-factor = 0.1.

if particle-shape = nugget
then shape-factor = 0.2.

if particle-shape = needle
then shape-factor = number(1,10).

/*----- Minimum Particle Size -----*/

question(minimum-particle-size) = 'What is the minimum particle size in cm?'.
legalvals(minimum-particle-size) = number.

/*----- Maximum Particle Size -----*/

question(maximum-particle-size) = 'What is the maximum particle size in cm?'.
 legalvals(maximum-particle-size) = number.

/*----- Maximum -----*/

if D0 < D1 and
 D1 = D
 then maximum(D0,D1) = D.

if D1 < D0 and
 D0 = D
 then maximum(D0,D1) = D.

/*----- Ratio -----*/

if D0/D1 = R
 then ratio(D0,D1) = R.

/*----- Equal -----*/

if ratio(D0,D1) = R and
 R * 1 = R
 then equal(D0,D1) = D0.

/*----- Granulometric Factor G -----*/

if minimum-particle-size = D0 and
 maximum-particle-size = D1 and
 maximum(D0,D1) = D1 and
 ratio(D1,D0) = R and
 R > 4
 then size-range = large.

if size-range = large
 then material-state = non-calibrated.

if material-state = non-calibrated
 then granulometric-factor = 0.25.

if minimum-particle-size = D0 and
 maximum-particle-size = D1 and
 maximum(D0,D1) = D1 and

ratio(D1,D0) = R and
 R >= 2 and
 R <= 4
 then size-range = medium.

if size-range = medium
 then material-state = artificially-calibrated.

if material-state = artificially-calibrated
 then granulometric-factor = 0.5.

if minimum-particle-size = D0 and
 maximum-particle-size = D1 and
 maximum(D0,D1) = D1 and
 ratio(D1,D0) = R and
 R > 1 and
 R < 2
 then size-range = small.

if size-range = small
 then material-state = naturally-calibrated.

if material-state = naturally-calibrated
 then granulometric-factor = 0.75.

if minimum-particle-size = D0 and
 maximum-particle-size = D1 and
 D1 = D0
 then size-range = uniform.

if size-range = uniform
 then material-state = uniformly-calibrated.

if material-state = uniformly-calibrated
 then granulometric-factor = 1.

/*----- Composition of the Critical Component -----*/

question(critical-component-composition) = 'What is the composition of the critical stream
 component in %?'.
 legalvals(critical-component-composition) = number.

/*----- Density of the Critical Component -----*/

question(critical-component-density) = 'What is the density of the critical stream component in g/cm³?'

legalvals(critical-component-density) = number.

/*----- Density of the Noncritical Components -----*/

question(noncritical-components-density) = 'What is the density of the noncritical stream components in g/cm³?'

legalvals(noncritical-components-density) = number.

/*----- Mineralogical Factor C -----*/

if critical-component-composition = AL and
 AL/100 = AAL and
 critical-component-density = CD and
 noncritical-components-density = ND and
 $(1-AAL)*((CD*(1-AAL)/AAL) + ND) = C$
 then mineralogical-factor = C.

/*----- Constitution Heterogeneity -----*/

question(material-constitution) = 'How is the constitution of the stream material?'.
 legalvals(material-constitution) = [extremely-heterogeneous,very-heterogeneous,heterogeneous,average,homogeneous,very-homogeneous,uniform].

automaticmenu(material-constitution).

if material-constitution = extremely-heterogeneous
 then liberation-factor1 = 1.

if material-constitution = very-heterogeneous
 then liberation-factor1 = 0.8.

if material-constitution = heterogeneous
 then liberation-factor1 = 0.4.

if material-constitution = average
 then liberation-factor1 = 0.2.

if material-constitution = homogeneous
 then liberation-factor1 = 0.1.

```
if material-constitution = very-homogeneous
then liberation-factor1 = 0.05.
```

```
if material-constitution = uniform
then liberation-factor1 = 0.
```

```
/*----- Liberation Size -----*/
```

```
question(liberation-size) = 'What is the liberation size in cm?'.
legalvals(liberation-size) = number.
```

```
/*----- Liberation Factor L -----*/
```

```
if maximum-particle-size = D1 and
   liberation-size = DL and
   sqrt(DL/D1) = L
then liberation-factor2 = L.
```

```
if liberation-factor1 = L1 and
   liberation-factor2 = L2 and
   maximum(L1,L2) = L
then liberation-factor = L.
```

```
if liberation-factor1 = L1 and
   liberation-factor2 = L2 and
   equal(L1,L2) = L
then liberation-factor = L.
```

```
/*----- Constitution heterogeneity CH -----*/
```

```
if (shape-factor = F and
    (granulometric-factor = G and
     (maximum-particle-size = D and
      (mineralogical-factor = C and
       (liberation-factor = L and
        (F * G * C * L * D * D * D = CH))))))
then constitution-heterogeneity = CH.
```

```
/*----- Lot Mass -----*/
```

```
question(lot-mass) = 'What is the lot mass in t?'.
legalvals(lot-mass) = number.
```

```
/*----- Solid Flowrate -----*/
```

```
question(stream-solid-flowrate) = 'What is the solid flowrate of the stream in t/h?'.
legalvals(stream-solid-flowrate) = number.
```

```
/*----- Cutter Width -----*/
```

```
question(cutter-width) = 'What is the cutter width in cm?'.
legalvals(cutter-width) = number.
```

```
/*----- Cutter Speed -----*/
```

```
question(cutter-speed) = 'What is the cutter speed in m/s?'.
legalvals(cutter-speed) = number.
```

```
/*----- Increment Mass -----*/
```

```
if stream-solid-flowrate = F and
   cutter-width = W and
   cutter-speed = S and
    $F*W/(0.36*S) = IM$ 
then increment-mass = IM.
```

```
/*----- Increment Mass Displayed -----*/
```

```
if increment-mass = IM
then conclusion0 = ['INCREMENT MASS:',nl,'The mass of an increment equals ',IM,'
g.',nl,nl].
```

```
if increment-mass = IM and
   conclusion0 = CONCLUSION
then display(CONCLUSION).
```

```
if increment-mass = IM and
   conclusion0 = CONCLUSION and
   display(CONCLUSION)
then increment-mass-displayed.
```

```
/*----- Sample Increments Number -----*/
```

```
question(increments-number) = 'What is the number of increments extracted for the
sample analysis?'
```

legalvals(increments-number) = number.

/*----- Sample Mass -----*/

if increment-mass = IM and
 increments-number = N and
 $IM * N = SM$
 then sample-mass = SM.

/*----- Sampling Period -----*/

if lot-mass = LM and
 stream-solid-flowrate = SF and
 increments-number = IN and
 $60 * LM / (SF * IN) = P$
 then sampling-period = P.

/*----- Sampling Period Displayed -----*/

if increments-number = N and
 sampling-period = P
 then conclusion1 = ['SAMPLING PERIOD:',nl,'The sampling period equals ',P,' min for
 ',N,' sample increments.',nl,nl].

if sampling-period = P and
 conclusion1 = CONCLUSION
 then display(CONCLUSION).

if sampling-period = P and
 conclusion1 = CONCLUSION and
 display(CONCLUSION)
 then sampling-period-displayed.

/*----- Fundamental Error FE for One Increment-----*/

if constitution-heterogeneity = CH and
 lot-mass = LM1 and
 $LM1 * 1000000 = LM$ and
 increment-mass = IM and
 $((1 / IM) - (1 / LM)) * CH = FE$
 then fundamental-error1 = FE.

/*----- Fundamental Error FE -----*/

if constitution-heterogeneity = CH and
 lot-mass = LM1 and
 LM1 * 1000000 = LM and
 sample-mass = SM and
 ((1 / SM) - (1 / LM)) * CH = FE
 then fundamental-error = FE.

/*----- Distribution Heterogeneity -----*/

question(material-distribution) = 'How is the distribution of the stream material?'.
 legalvals(material-distribution) = [extremely-heterogeneous, very-
 heterogeneous, heterogeneous, average, homogeneous, very-homogeneous, uniform].

automaticmenu(material-distribution).

if material-distribution = extremely-heterogeneous
 then distribution-factor = 1.

if material-distribution = very-heterogeneous
 then distribution-factor = 0.9.

if material-distribution = heterogeneous
 then distribution-factor = 0.7.

if material-distribution = average
 then distribution-factor = 0.5.

if material-distribution = homogeneous
 then distribution-factor = 0.3.

if material-distribution = very-homogeneous
 then distribution-factor = 0.1.

if material-distribution = uniform
 then distribution-factor = 0.

/*----- Distribution Error DE for One Increment -----*/

if constitution-heterogeneity = CH and
 lot-mass = LM1 and
 LM1 * 1000000 = LM and

```

increment-mass = IM and
((1 / IM) - (1 / LM)) * CH = FE and
fundamental-error1 = FE and
distribution-factor = DF and
DF*FE = DE
then distribution-error1 = DE.

```

```

/*----- Distribution Error DE -----*/

```

```

if constitution-heterogeneity = CH and
lot-mass = LM1 and
LM1*1000000 = LM and
sample-mass = SM and
((1 / SM) - (1 / LM)) * CH = FE and
fundamental-error = FE and
distribution-factor = DF and
DF*FE = DE
then distribution-error = DE.

```

```

/*----- Sampling Error SE for One Increment -----*/

```

```

if fundamental-error1 = FE and
distribution-error1 = DE and
FE+DE = SE
then sampling-error1 = SE.

```

```

if sampling-error1 = SE and
SE*100 = SEPERCENT
then conclusion-se1 = ['INCREMENT SAMPLING ERROR:',nl,'The sampling error
variance equals ',SEPERCENT,'%,' for one sample increment.',nl,nl].

```

```

/*----- Sampling Error SE Displayed for One Increment -----
*/

```

```

if sampling-error1 = SE and
conclusion-se1 = CONCLUSION
then display(CONCLUSION).

```

```

if sampling-error1 = SE and
conclusion-se1 = CONCLUSION and
display(CONCLUSION)
then sampling-error1-displayed.

```

/*----- Sampling Error SE -----*/

if fundamental-error = FE and
 distribution-error = DE and
 FE+DE = SE
 then sampling-error = SE.

if increments-number = N and
 sampling-error = SE and
 SE*100 = SEPERCENT
 then conclusion = ['SAMPLING ERROR:',nl,'The sampling error variance equals
 ',SEPERCENT,' %,' for ',N,' sample increments.',nl,nl].

/*----- Sampling Error SE Displayed -----*/

if sampling-error = SE and
 conclusion = CONCLUSION
 then display(CONCLUSION).

if sampling-error = SE and
 conclusion = CONCLUSION and
 display(CONCLUSION)
 then sampling-error-displayed.

/*----- Sampling Error Variance Ratio -----*/

question(sampling-error-ratio) = 'What is the ratio, in %, of the sampling error variance
 for the minimum number of increments to the sampling error variance for one increment?'.
 legalvals(sampling-error-ratio) = number(0,100).

/*----- Minimum Increment Number -----*/

if constitution-heterogeneity = CH and
 lot-mass = LM1 and
 increment-mass = IM and
 LM1*1000000 = LM and
 distribution-factor = DF and
 sampling-error1 = SE and
 sampling-error-ratio = R0 and
 R0/100 = R and
 $1/(IM*((R*SE)/(CH*(1+DF)))-(1/LM))) = N1$ and
 N1+1 = N2 and
 fix(N2) = N

then minimum-increment-number = N.

/*----- Minimum Increment Number Displayed -----*/

if minimum-increment-number = N and
 sampling-error-ratio = R
 then conclusion-min = ['MINIMUM INCREMENT NUMBER:',nl,'The minimum
 increment number equals ',N,' for a sampling error variance ratio of ',R,' %.',nl,nl].

if minimum-increment-number = N and
 conclusion-min = CONCLUSION
 then display(CONCLUSION).

if minimum-increment-number = N and
 conclusion-min = CONCLUSION and
 display(CONCLUSION)
 then minimum-increment-number-displayed.

/*----- Minimum Sample Mass -----*/

if minimum-increment-number = N and
 increment-mass = IM and
 IM * N = M
 then minimum-sample-mass = M.

/*----- Sampling Period for Minimum Increment Number -----
 --*/

if lot-mass = LM and
 stream-solid-flowrate = SF and
 minimum-increment-number = IN and
 $60 * LM / (SF * IN) = P$
 then maximum-sampling-period = P.

/*----- Maximum Sampling Period Displayed -----*/

if minimum-increment-number = N and
 maximum-sampling-period = P
 then conclusion-msp = ['MAXIMUM SAMPLING PERIOD:',nl,'The sampling period
 equals ',P,' min for ',N,' sample increments.',nl,nl].

```

if maximum-sampling-period = P and
  conclusion-msp = CONCLUSION
then display(CONCLUSION).

```

```

if maximum-sampling-period = P and
  conclusion-msp = CONCLUSION and
  display(CONCLUSION)
then maximum-sampling-period-displayed.

```

```

/*----- Maximum Fundamental Error FE -----*/

```

```

if constitution-heterogeneity = CH and
  lot-mass = LM1 and
  LM1*1000000 = LM and
  minimum-sample-mass = SM and
  ((1 / SM) - (1 / LM)) * CH = FE
then maximum-fundamental-error = FE.

```

```

/*----- Maximum Distribution Error DE -----*/

```

```

if constitution-heterogeneity = CH and
  lot-mass = LM1 and
  LM1*1000000 = LM and
  minimum-sample-mass = SM and
  ((1 / SM) - (1 / LM)) * CH = FE and
  maximum-fundamental-error = FE and
  distribution-factor = DF and
  DF*FE = DE
then maximum-distribution-error = DE.

```

```

/*----- Maximum Sampling Error SE -----*/

```

```

if maximum-fundamental-error = FE and
  maximum-distribution-error = DE and
  FE+DE = SE
then maximum-sampling-error = SE.

```

```

if minimum-increment-number = N and
  maximum-sampling-error = SE and
  SE*100 = SEPERCENT
then conclusion-mse = ['MAXIMUM SAMPLING ERROR:',nl,'The sampling error
variance equals ',SEPERCENT,' %,' for ',N,' sample increments.',nl,nl].

```

```
/*----- Maximum Sampling Error SE Displayed -----*/
```

```
if maximum-sampling-error = SE and
  conclusion-mse = CONCLUSION
then display(CONCLUSION).
```

```
if maximum-sampling-error = SE and
  conclusion-mse = CONCLUSION and
  display(CONCLUSION)
then maximum-sampling-error-displayed.
```

```
/*----- Welcome -----*/
```

```
if display(['----- SAMPLING ERROR EVALUATOR -----
-----',nl,nl])
then welcome.
```

```
/*----- Evaluation End -----*/
```

```
if display(['----- END -----
-----',nl,nl])
then evaluation-end.
```

APPENDIX G
VALIDATION OF
THE EXPERT SYSTEM SAMPLING ERROR EVALUATOR

G1. CASE STUDY NUMBER 1

"----- SAMPLING ERROR EVALUATOR -----

"
"
"
"What is the solid flowrate of the stream in t/h?
"
">10
"
"What is the cutter width in cm?
"
">2
"
"What is the cutter speed in m/s?
"
">0.6
"
"INCREMENT MASS:
"
"The mass of an increment equals 92.59260 g.
"
"
"
"What is the shape of the particles?
"
" cube
"
" sphere
"
" flake
"
" nugget
"
" needle

"
">"sphere"
"
"What is the minimum particle size in cm?
"
">0.001
"
"What is the maximum particle size in cm?
"
">0.01
"
"What is the composition of the critical stream component in %?
"
">20
"
"What is the density of the critical stream component in g/cm³?
"
">4.5
"
"What is the density of the noncritical stream components in g/cm³?
"
">2.8
"
"How is the constitution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"heterogeneous"
"
"What is the liberation size in cm?
"

">0.01
"
"What is the lot mass in t?
"
">100
"
"How is the distribution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"average"
"
"INCREMENT SAMPLING ERROR:
"
"The sampling error variance equals 3.5246e-006 % for one sample increment.
"
"
"
"What is the number of increments extracted for the sample analysis?
"
">10
"
"SAMPLING PERIOD:
"
"The sampling period equals 60.0 min for 10 sample increments.
"
"
"
"SAMPLING ERROR:
"
"The sampling error variance equals 3.52457e-007 % for 10 sample increments.

"

"

"

"What is the ratio, in %, of the sampling error variance for the minimum number of increments to the sampling error variance for one increment?"

"

">10

"

"MINIMUM INCREMENT NUMBER:

"

"The minimum increment number equals 11 for a sampling error variance ratio of 10 %.

"

"

"

"MAXIMUM SAMPLING PERIOD:

"

"The sampling period equals 54.54550 min for 11 sample increments.

"

"

"

"MAXIMUM SAMPLING ERROR:

"

"The sampling error variance equals 3.20415e-007 % for 11 sample increments.

"

"

"

"----- END -----"

"

"

"

G2. CASE STUDY NUMBER 2

"----- SAMPLING ERROR EVALUATOR -----"

"

"

"

"What is the solid flowrate of the stream in t/h?"

"

">20

"

"What is the cutter width in cm?

"

">2

"

"What is the cutter speed in m/s?

"

">0.6

"

"INCREMENT MASS:

"

"The mass of an increment equals 185.1850 g.

"

"

"

"What is the shape of the particles?

"

" cube

"

" sphere

"

" flake

"

" nugget

"

" needle

"

">"sphere"

"

"What is the minimum particle size in cm?

"

">0.002

"

"What is the maximum particle size in cm?

"

">0.02

"

"What is the composition of the critical stream component in %?

"

">89

"

"What is the density of the critical stream component in g/cm³?

"
">4
"
"What is the density of the noncritical stream components in g/cm³?
"
">3
"
"How is the constitution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"average"
"
"What is the liberation size in cm?
"
">0.01
"
"What is the lot mass in t?
"
">80
"
"How is the distribution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"

" homogeneous
"
" very-homogeneous
"
" uniform
"
">"homogeneous"
"
"INCREMENT SAMPLING ERROR:
"
"The sampling error variance equals 1.99579e-007 % for one sample increment.
"
"
"
"What is the number of increments extracted for the sample analysis?
"
">8
"
"SAMPLING PERIOD:
"
"The sampling period equals 30.0 min for 8 sample increments.
"
"
"
"SAMPLING ERROR:
"
"The sampling error variance equals 2.4947e-008 % for 8 sample increments.
"
"
"
"What is the ratio, in %, of the sampling error variance for the minimum number of increments to the sampling error variance for one increment?
"
">10
"
"MINIMUM INCREMENT NUMBER:
"
"The minimum increment number equals 11 for a sampling error variance ratio of 10 %.
"
"
"
"MAXIMUM SAMPLING PERIOD:
"
"

"The sampling period equals 21.81820 min for 11 sample increments.

"

"

"

"MAXIMUM SAMPLING ERROR:

"

"The sampling error variance equals 1.81432e-008 % for 11 sample increments.

"

"

"

"----- END -----

"

"

"

G3. CASE STUDY NUMBER 3

"----- SAMPLING ERROR EVALUATOR -----

"

"

"

"What is the solid flowrate of the stream in t/h?

"

">15

"

"What is the cutter width in cm?

"

">3

"

"What is the cutter speed in m/s?

"

">0.7

"

"INCREMENT MASS:

"

"The mass of an increment equals 178.5710 g.

"

"

"

"What is the shape of the particles?

"

" cube

"

" sphere

"

" flake

"

" nugget

"

" needle

"

">"sphere"

"

"What is the minimum particle size in cm?

"

">0.001

"

"What is the maximum particle size in cm?

"

">0.05

"

"What is the composition of the critical stream component in %?

"

">70

"

"What is the density of the critical stream component in g/cm³?

"

">5

"

"What is the density of the noncritical stream components in g/cm³?

"

">2

"

"How is the constitution of the stream material?

"

" extremely-heterogeneous

"

" very-heterogeneous

"

" heterogeneous

"

" average

"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"extremely-heterogeneous"
"
"What is the liberation size in cm?
"
">0.01
"
"What is the lot mass in t?
"
">50
"
"How is the distribution of the stream material?
"
" extremely-heterogeneous
"
" very-heterogeneous
"
" heterogeneous
"
" average
"
" homogeneous
"
" very-homogeneous
"
" uniform
"
">"very-homogeneous"
"
"INCREMENT SAMPLING ERROR:
"
"The sampling error variance equals 1.25127e-005 % for one sample increment.
"
"
"
"What is the number of increments extracted for the sample analysis?
"

">5

"

"SAMPLING PERIOD:

"

"The sampling period equals 40.0 min for 5 sample increments.

"

"

"

"SAMPLING ERROR:

"

"The sampling error variance equals 2.50251e-006 % for 5 sample increments.

"

"

"

"What is the ratio, in %, of the sampling error variance for the minimum number of increments to the sampling error variance for one increment?

"

">5

"

"MINIMUM INCREMENT NUMBER:

"

"The minimum increment number equals 21 for a sampling error variance ratio of 5 %.

"

"

"

"MAXIMUM SAMPLING PERIOD:

"

"The sampling period equals 9.523810 min for 21 sample increments.

"

"

"

"MAXIMUM SAMPLING ERROR:

"

"The sampling error variance equals 5.95802e-007 % for 21 sample increments.

"

"

"

"----- END -----

"

"

"