Theory of Sampling Contribution to Multivariate Image Analysis

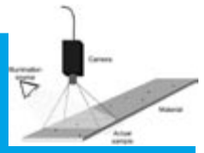
Dahl, Casper Kierulf

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Casper Kierulf Dahl

An Introduction to Image Analytical Sampling



Casper Kierulf Dahl

Theory of Sampling Contributions to Multivariate Image Analysis

with a Special Focus on

Image Analytical Sampling

ISBN 978--87-7606-021-3

ACABS Research Group
Aalborg University Esbjerg

2007

AALBORG UNIVERSITY
ESBJERG

THEORY OF SAMPLING CONTRIBUTION TO MULTIVARIATE IMAGE ANALYSIS

WITH A SPECIAL FOCUS ON

IMAGE ANALYTICAL SAMPLING

PhD Thesis by

CASPER KIERULF DAHL

ACABS Research Group
(Applied Chemometrics, Acoustic Chemometrics, Analytical Chemistry,
Applied Biotechnology, Bioenergi & Sampling)
Aalborg University Esbjerg
Niels Bohrs vej 8
DK-6700 Esbjerg
Denmark

SUPERVISOR
Professor Kim H. Esbensen

Thesis Submitted to the International Doctoral School of Science and Technology,
Aalborg University, Denmark, for the Degree of Doctor of Philosophy.

June 2007

PREFACE

A long – indeed very long – roller coaster ride is nearing its end. And what a ride it has been; it has been bumpy to say at least, but also thrilling, scary, and tiring at times. Luckily, it has never been boring and how could it in the ACABS Research Group.

The ride started back in the spring of 2001, when Professor Kim Esbensen visited Aalborg University Esbjerg as a guest lecturer and held a short presentation on multivariate data analysis. We ended up talking about sampling aspects, which was a rather new research discipline for Kim (and completely unknown to me and the rest of the group). For some reason Lars Petersen and I ended up talking to Kim about his future employment at Aalborg University Esbjerg and about a PhD. scholarship in his new group. This actually ended up in two full scholarships: One full-time scholarship on theoretical and practical sampling at Aalborg University Esbjerg and one full-time Marie Curie scholarship on practical sampling with one year in Esbjerg and two years at HIT/TF in Porsgrunn, Norway. Shortly before going to Norway (after approximately ¾ of a year) the Marie Curie scholarship was abruptly terminated due to the sudden death of the responsible supervisor in Porsgrunn, Professor Sunil de Silva. The scholarship was quickly converted into an additional full-time scholarship at AAUE in Esbjerg, which required quite a lot of sacrifices from both Kim and head of institute Torben Rosenørn. I am indeed grateful for getting this opportunity; thank you Torben and Kim!

The ride is about to end and hopefully all the essential parts of the journey are documented in this thesis. No one could have imagined the length and the challenges involved in this ride. Typically, a PhD scholarship offers many challenges and this should always be expected even though it can set you back to a certain level. The challenges for this scholarship included; building a new research group including building a suitable imaging facility, losing access to this facility due to a sudden renovation of the laboratories, re-creation of an image laboratory due to both the laboratory renovation and software problems, and cancellation of papers due to author disputes – only to mention a few of the challenges one should expect during a scholarship. However, more uncommon challenges were more or less crucial parts of this journey: marriage, the birth of two wonderful girls, bereavement of close family, illness, getting a new job, and moving the entire immediate family.

Many have contributed directly or indirectly to the work behind this thesis and I would like to thank everyone involved for making this dream come true. I will name a few, but the list is in reality more or less endless.

First, and most importantly, I would like to thank Kim for giving me this opportunity. As mentioned above, it has certainly been challenging, yet I have learned a lot from the dear professor. I would say most of my scientific skills come from many hours of discussions with Kim, so naturally I owe Kim a great deal for everything. The scientific education is one aspect of this educational development; however, the cultural aspect was equally as essential and at least as interesting thanks to Kim.

I would also like to thank my fellow colleague since 1996, Lars Petersen, for being a good friend and laboratory partner through the many years. The first 1½ years of this scholarship were spent primarily with Lars (and Kim) interpreting and deducing the many aspects of TOS (pure theory – a different challenge), imaging and analysis of agglomerated samples (unexplored, but indeed interesting and sticky territory – Paper I), sorting and shoveling of PP pellets (the aquarium experiment – SIGH) and the tedious mass reduction experiment of various mass reduction devices/methods (documented in Paper V).

Another great contributor to this thesis is the entire ACABS Research Group, which during the six year period included: Lars Houmøller, Hans Henrik Friis-Pedersen, Jonas Sjøland, Johnny Madsen, James Burger, Carina Lomborg, Jens Bo Holm-Nielsen, and Peter Paasch Mortensen. Thank you all for everything.

I also would like to thank Anne Cole for introducing me to the many aspects of the English/American language. There is still room for improvements on my side, but you have without a doubt contributed much to my language development throughout the entire scholarship, which has been effectively combined with the practical experience of writing scientific and peer reviewed papers.

The entire team of the AAUE-staff has been unbelievably helpful and inspiring throughout my 10 years at the AAUE campus. Thank you all for making my stay as both a student and an employee at AAUE so interesting, fun, and relaxing.

The entire team at Rationel Kornservice A/S provided unbelievable and unselfish help during this scholarship. I thank RAKO for supporting me and ACABS, and for the great insight into practical sampling solutions.

I could never have completed this long ride without the help and support of everyone around me including my wonderful family. But ultimately I could not have done this without my dear wife, Laisa, who gave me the necessary space and support to finish this great opportunity. I am truly grateful to her and this thesis is dedicated to my loving Laisa.

ABSTRACT

The work behind this thesis was inspired by my interest for two different scientific main areas: The Theory of Sampling (TOS) and multivariate data analysis.

This PhD scholarship started out with focus on the practical application of multivariate data analytical techniques – specifically on the application of the image analytical technique Multivariate AMT Regression (MAR) on various substances/mixtures. A study on moisture-induced agglomeration was followed by an educational study of TOS, which ended up by establishing the foundation for understanding and applying TOS at Aalborg University Esbjerg – a work carried out together with my colleagues Kim H. Esbensen and Lars Petersen. This didactic study was followed by a series of practical mass reduction experiments documenting various aspects of TOS in practice as well as theoretically. With these deductive and practical studies completed, this PhD turned towards application and understanding of MAR.

After a while with several practical imaging studies on various substances/mixtures, the focus of this PhD slowly changed into a research area, which amalgamates the Theory of Sampling (TOS) and image analysis into a new field termed Image Analytical Sampling (IAS).

The introduction of this thesis serves as a first comprehensive delimitation of Image Analytical Sampling, but also of all the many disciplines behind IAS. The introduction is followed by five papers:

I. C.K. Dahl (50%), L. Petersen (40%) & K.H. Esbensen (10%), "*Image Analytical Characterization of Powder Agglomeration by AMT-Regression (Angle Measure Technique) – Development and Method Validation*". In preparation.

II. J.B. Holm-Nielsen (50%), C.K. Dahl (40%) & K.H. Esbensen (10%), "*Representative Sampling for Process Analytical Characterization of Heterogeneous Bioslurry Systems – a Reference Study of Sampling Issues in PAT*", Chemometrics and Intelligent Laboratory Systems Volume 83 (2006) pp. 114-126.

III. C.K. Dahl (80%), P. Minkkinen (10%) & K.H. Esbensen (10%), "*Image Analytical Monitoring of Paper Quality – a Feasibility Study*", TAPPI JOURNAL, Volume 5 Issue 11 (2006) pp. 18-24.

IV. C.K. Dahl (75%) & K.H. Esbensen (25%), "*Image Analytical Determination of Particle Size Distribution of Natural and Bulk Aggregates*", Chemometrics and Intelligent Laboratory Systems, Volume 89 (2007) issue 1 pp. 9-25.

V. L. Petersen (50%), C.K. Dahl (40%) & K.H. Esbensen (10%), "*Representative Mass Reduction in Sampling – a critical Survey of Techniques and Hardware*", Chemometrics and Intelligent Laboratory Systems, Volume 74 (2004) issue 1 pp. 95-114.

SYNOPSIS

Arbejdet bag denne afhandling er inspireret af min interesse for to forskellige videnskabelige discipliner: The Theory of Sampling (TOS) og the Angle Measure Technique (AMT).

Denne PhD startede med fokus på praktisk anvendelse af billedanalytiske teknikker – primært Multivariate AMT Regression (MAR) – på forskellige substanser/blandinger. Første forsøg var et fugt-induceret agglomereringseksperiment, hvilket blev efterfulgt af et teoretisk studium af TOS. Dette studium blev udført sammen med mine kollegaer Kim H. Esbensen og Lars Petersen og endte ud som et fundament for forståelse og anvendelse af TOS på Aalborg Universitet Esbjerg. Studiet blev efterfulgt af en række praktiske forsøg, der dokumenterede forskellige praktiske og teoretiske aspekter af TOS. Kim og Lars fortsatte herefter med at arbejde med TOS, mens mit arbejde drejede i retning af anvendelse og forståelse af MAR.

Efter at have gennemført adskillige praktiske billedanalytiske studier på forskellige substanser/blandinger, blev fokus for afhandlingen ændret til at introducere en ny videnskabelig disciplin, som sammensmelter TOS og billedanalyse til billedanalytisk prøvetagning – Image Analytical Sampling (IAS).

Introduktionen til denne afhandling fungerer som en første indledning til Billedanalytisk Prøvetagning (Image Analytical Sampling), samt som en introduktion til de forskellige teknikker bag IAS. Introduktionen efterfølges af fem artikler:

I. C.K. Dahl (50%), L. Petersen (40%) & K.H. Esbensen (10%), "*Image Analytical Characterization of Powder Agglomeration by AMT-Regression (Angle Measure Technique) – Development and Method Validation*". Under forberedelse.

II. J.B. Holm-Nielsen (50%), C.K. Dahl (40%) & K.H. Esbensen (10%), "*Representative Sampling for Process Analytical Characterization of Heterogeneous Bioslurry Systems – a Reference Study of Sampling Issues in PAT*", Chemometrics and Intelligent Laboratory Systems Volume 83 (2006) pp. 114-126.

III. C.K. Dahl (80%), P. Minkkinen (10%) & K.H. Esbensen (10%), "*Image Analytical Monitoring of Paper Quality – a Feasibility Study*", TAPPI JOURNAL, Volume 5 Issue 11 (2006) pp. 18-24.

IV. C.K. Dahl (75%) & K.H. Esbensen (25%), "*Image Analytical Determination of Particle Size Distribution of Natural and Bulk Aggregates*", Chemometrics and Intelligent Laboratory Systems, Volume 89 (2007) issue 1 pp. 9-25.

V. L. Petersen (50%), C.K. Dahl (40%) & K.H. Esbensen (10%), "*Representative Mass Reduction in Sampling – a critical Survey of Techniques and Hardware*", Chemometrics and Intelligent Laboratory Systems, Volume 74 (2004) issue 1 pp. 95-114.

CONTENTS

1 INTRODUCTION

The need for industrial online process or product monitoring is ever increasing to maximize either process output or product quality (or both simultaneously). This monitoring is often carried out by use various image analytical techniques. These techniques do in general all suffice from a serious assumption:

> *The samples underneath the camera are tacitly assumed to be representative of the lot product quality in general and 2-D image acquired of the sample surface is assumed to be representative of the 3-D bulk sample.*

Both assumptions are only semi-valid in limited cases and can not be considered applicable in practice. One sample/image can never fully represent the quality of a whole lot! With regards to the second assumption (concerning 2-D image vs. 3-D sample representativity), this assumption is only valid for non-segregating materials. What is the solution to this image analytical problem? Structurally correct sampling as laid down in the Theory of Sampling (TOS).

Combining image analysis and TOS provides the ultimate solution for many online process/product monitoring tasks by taking the best from both worlds and providing representative image analysis. This new field is introduced in this thesis and the methodology is termed Image Analytical Sampling (IAS).

IAS can also be considered a process analytical technology (PAT) per the official FDA definition:

Process Analytical Technology is:

- a system for designing, analyzing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality.

[*Source: FDA's "Proces Analytical Technology (PAT) Initiative" http://www.fda.gov/Cder/OPS/PAT.htm*]

The main area for this thesis is combining representative, i.e. structurally correct sampling, of particular materials with proxy methods – specifically with image analysis. The theoretical background and the applicability of this amalgamated PAT technique is exemplified through a specific image analytical technique termed Multivariate AMT Regression (MAR) pioneered by Jun Huang. MAR uses a specific image preprocessing procedure called the Angle Measure Technique (AMT) whose output is so-called AMT spectra, which typically are analyzed by use of traditional multivariate regression methods – in this thesis multivariate calibration is only exemplified through partial least squares regression, PLS-R.

2 MULTIVARIATE DATA ANALYSIS

Many different types of multivariate data analysis exist in chemometrics. They are typically divided into three distinct groups [1]:

- Data description
- Discrimination and classification
- Regression and prediction

This thesis has had its main chemometric focus on multivariate regression, which has been carried out on unfolded[a] and preprocessed[b] images using Partial Least Squares regression (PLS-R).

The objective of multivariate calibration is to establish a regression model of empirical (X,Y) relations for prediction purposes. A calibration model can be illustrated as shown in Figure 1, which focuses on the fact that the X-variables always *are* and the Y-variable(s) *can* be multivariate.

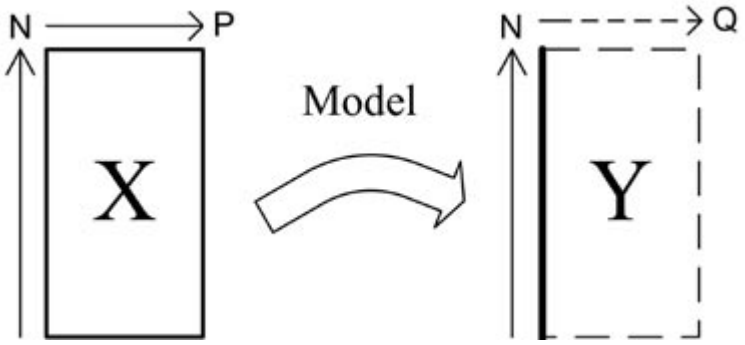


Figure 1. Establishing a multivariate calibration model. N represents the number of objects (samples), while P and Q represent the variables in X and Y, respectively.

In many cases a linear regression relationship between the X and Y data exists (shown here for one y-variable):

---

[a] See chapter 3.1.
[b] Preprocessed using the Angle Measure Technique (chapter 2.3).

$$\mathbf{y=X\beta+E}$$

with **β** being the regression coefficients for all P **X**-variables and **E** the errors (unexplained variance).

2.1 PARTIAL LEAST SQUARES REGRESSION

PLS regression is a powerful generalized regression technique that surpasses standard statistical calibration concepts such as principal component regression (PCR) and multiple linear regression (MLR) [1,2]. The objective of PLS regression is to predict the dependent variable(-s) (**Y**) from the independent variables (**X**), and the technique does - as the only method - allow an active influence of the Y-information in the decomposition of the X-matrix by only seeking the part of X that is relevant (i.e. correlated) for the description of Y.

The procedure for estimating the model parameters is typically built on the so-called non-linear iterative partial least squares algorithm (NIPALS) – an algorithm based on iterative convergence to underlying (hidden) correlated latent components in both (X,Y) spaces. This algorithm was originally formulated by Herman Wold, and was later further developed in chemometrics by Svante Wold [3,4] and others.

The NIPALS algorithm is used in PLS regression in two different versions: PLS1 and PLS2. PLS1 is used for predicting a single y-variable, while PLS2 is used for predicting several y-variables simultaneously, taking advantage of correlations between the Y-variables. This can, however, be disadvantageous, if one or more Y-variables are not correlated with the other. PLS2 is for this reason used primarily for explorative analysis of Y-variable relationships. It can be especially difficult to get an overview of all Y-variables from individual PLS1 models alone; Paper III deals with no less than 42 Y-variables.

PLS-R is methodologically superior in its use of Y-data structure directly for decomposing the X-matrix hereby ensuring that only the interesting Y-correlated X-variation is put into the regression model (to be used for future predictions of **Y**), while the uninteresting non-correlated X-variation is filtered out (and put into an error matrix, **E**). There is also an error matrix for Y termed F.

This regression approach is based on transforming the original multi-dimensional X-data space into an embedded coordinate system with fewer dimensions lying centrally in the data – illustrated in Figure 2. The data elements from the original data space are projected onto this Y-guided PLS component space which captures the main variation in the data in the sense of simultaneously seeking the maximum **X-Y** covariance [2]. The direction corresponding to this maximum within the data (X) is represented as PLS component number 1 (PC1) shown on Figure 2, while the direction corresponding to the same criterion in the orthogonal residual X-space will be constructed as PLS component number 2 (PC2) and so on.

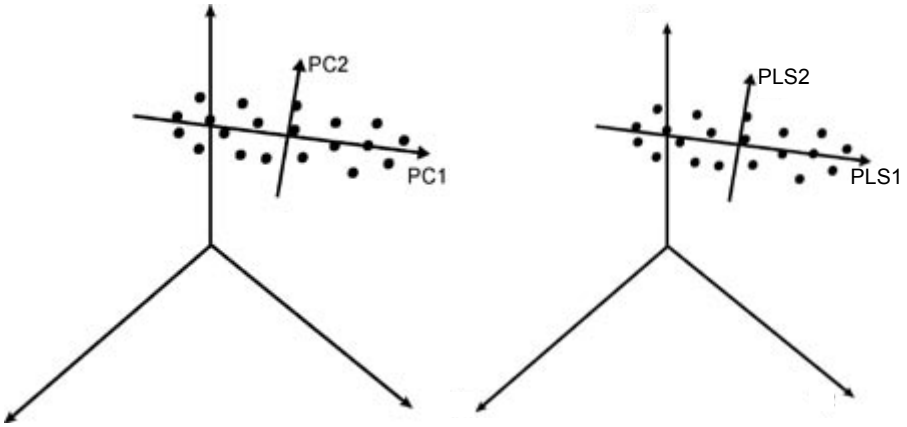


Figure 2. **Original three-dimensional X and Y data set represented by new two-dimensional embedded PLS component space. This projection feature of PLS-regression can be generalized also to higher-order dimensions (in both X- and Y-spaces).**

Each direction of PLS components can be understood as a linear combination of **P** unity vectors (**P** original variables in the X-space), specified by **P** direction-coefficients called loadings, p_{ka} (k is the variable index, and a is the component index). In PLS three different set of loadings exists: **P**, **Q** and **W**. The P-loadings express the relationship between the PLS components and the original X-axes, while **Q**-loadings express the similar relationship between PLS components and the original Y-axes. The loading weights (**W**) represent the effective loadings

which characterize the maximum (X,Y) covariance - see Figure 3 for an overview illustration. W expresses the core model relationships of the PLS model.

Both the X and Y scores (collected into matrices T and U) are found by projecting each data point (object) onto the pertinent PLS component expressing the distance from the projected footprint on the component to the center of the model. The residual can be found as the perpendicular projection distance from an object to a component.

All data can hence be described by two interdependent bilinear models:

$$\mathbf{X} = \sum_a \mathbf{T}_a \cdot \mathbf{P}_a^{\mathrm{T}} + \mathbf{E}$$
$$\mathbf{Y} = \sum_a \mathbf{U}_a \cdot \mathbf{Q}_a^{\mathrm{T}} + \mathbf{F}$$

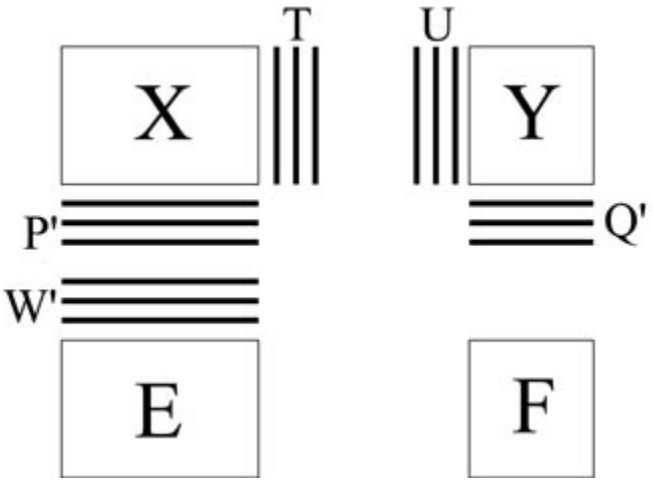


Figure 3. Schematic overview of the elements in a PLS-regression model.

2.1.1 THE PLS1 NIPALS ALGORITHM

Most applications in this thesis deal with only one y-variable at a time (PLS1), which makes the u-vector and y-vector identical. The general PLS2 NIPALS algorithm is reduced to the simple procedure below:

1. $\mathbf{w}_f = \mathbf{X}_f^T \mathbf{y}_f \Big/ \left| \mathbf{X}_f^T \mathbf{y}_f \right|$ (normalization of $\mathbf{w}$)
2. $\mathbf{t}_f = \mathbf{X}_f \mathbf{w}_f$
3. $q_f = \mathbf{t}_f^T \mathbf{y}_f \big/ \mathbf{t}_f^T \mathbf{t}_f$
4. $\mathbf{p}_f = \mathbf{X}_f^T \mathbf{t}_f \big/ \mathbf{t}_f^T \mathbf{t}_f$
5. $\mathbf{X}_{f+1} = \mathbf{X}_f - \mathbf{t}_f \mathbf{p}_f^T$ and $\mathbf{y}_{f+1} = \mathbf{y}_f - \mathbf{q}_f \mathbf{t}_f$
6. $f = f + 1$

Following several iterative steps using the above procedure, it is possible to obtain a stable (converged) model that explains as much as possible of the Y-variance using only the correlated X-variance.

The questions that naturally arise are: Can this model be used to predict future observations? How can this model be used to predict new Y-values directly from new X-data alone? What is the accuracy and precision of such predictions?

These questions can only be answered after a suitable *validation* of the prediction model performance.

2.2 VALIDATION

The purpose of validation is to provide a reliable assessment of the prediction performance of an already established model. This way it is possible to substantiate whether the model will work in the future on similar data sets. A second aspect of validation is finding the optimal model complexity – i.e. finding the optimal number of model components.

Finding the optimal model complexity is essential in multivariate regression. This task is often based on finding the lowest residual validation variance (the minimal prediction error), which can be seen as Root Mean Square Error of Prediction (RMSEP):

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{y}_{i,val} - y_{i,val} \right)^2}{n}}$$

with $\hat{y}_{i,val}$ being the predicted y-value in the validation and $y_{i,val}$ the actual y-value.

As shown on Figure 4, the model complexity can be visualized as a function of two parts: The modeling error ($\hat{y}_{cal} - y_{cal}$), which decreases when model complexity is increased[c], and the prediction error coming from estimating the regression coefficients (noise), which increases with increasing model complexity.

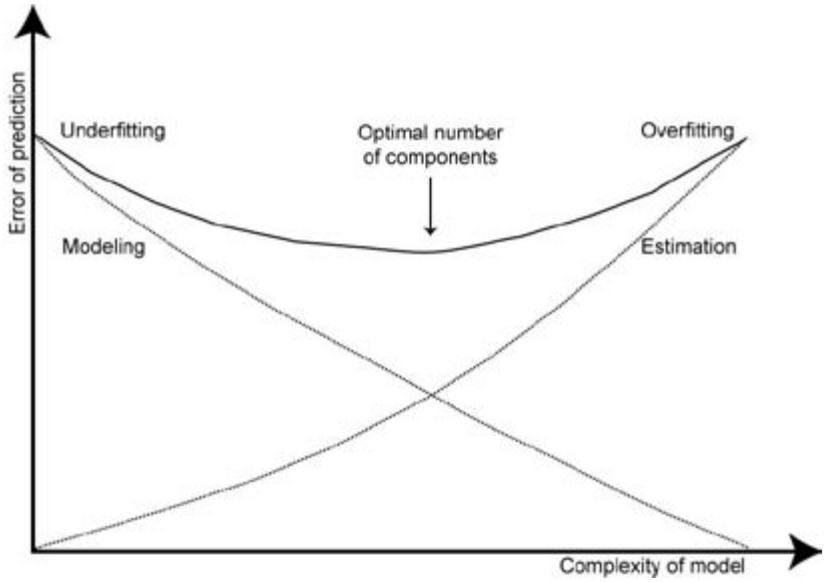


Figure 4. Illustration of the optimal number of components through a minimization of the compound modeling and prediction error.

The methodology described above for finding the optimal number of components is *the* well-established practice. First of all, it should be noted that this clear "V-rule" does not always apply; it depends on the data structure (as well as outlier removal). Secondly, all previous knowledge should be used in determining the optimal number of components. It is an all too easy task to minimize the

[c] Given that the calibration object are representative for the new objects to be predicted.

prediction error beyond the realistic by outlier removal or overfitting, but awareness should be paid to the estimation of uncertainties around the different validation methods; most validation procedures are often over-optimistic resulting in an unrealistic low prediction error.

Having selected the optimal model complexity, assessing the prediction performance of a model is the next task. It can be rather easy and straightforward if the proper prerequisites are met. This is, however, very seldom the case. Without proper consideration of all errors involved in the model building there is no valid validation. The only way to know if a model will truly give realistic predictions of future observations is to test its performance on a completely new and independent data set [1]. This data set specifically needs to be obtained in a way, so it is representative of future situations, as well as it is extracted in accordance to the principles of structurally correct sampling[d] for being representative of the bulk material, from which it was initially sampled. There is a direct relationship between unrecognized TOS-errors and typical chemometric validation issues, which need further attention than what is usually considered in chemometrics.

2.2.1 VALIDATION PROCEDURES

Within the chemometric society at least three well-know validation methods are in use:

- Test set validation
- Cross validation
- Leverage correction

Each method has its advantages and drawbacks, which will be elaborated on below.

TEST SET VALIDATION

Test set validation is used for estimating a realistic measure of the predictive ability of a model on a new, independent data set. The "price" of this validation approach is that two individual data sets are required: one for establishing the

[d] As defined by the Theory of Sampling (TOS) [6,28-30].

model and one for testing its predictive ability (aptly called the calibration and the test set, respectively). Thus two independent drawings from the parent population should be made, in which the statistical and physical (TOS) sampling conditions should be as similar as possible. In order to further the most realistic estimate of the future prediction accuracy and precision, the key issue is that both data sets should be representative of all future situations in which the prediction model is to be used [1,5]. If both data sets indeed have been extracted separately and under identical conditions, the difference between these sets should only consist of sampling errors and material variation [1,6].

Test set validation determines both model complexity and prediction error estimation simultaneously. This method has the obvious advantage that the validation procedure is fast and the computer requirements are low, since the method only requires one calculation step for validation (opposite to cross validation – see further below).

Procedure (Test Set Validation)
A model is first made using the calibration (often also called the modeling) data set, which is to be validated with regards to the optimal number of PLS-components. For both purposes the differences between the actual (measured or reference) and the predicted Y-values are calculated for each validation object.

Figure 5 shows an example of a predicted vs. measured graphical plot for paper roughness (example taken from Paper III).

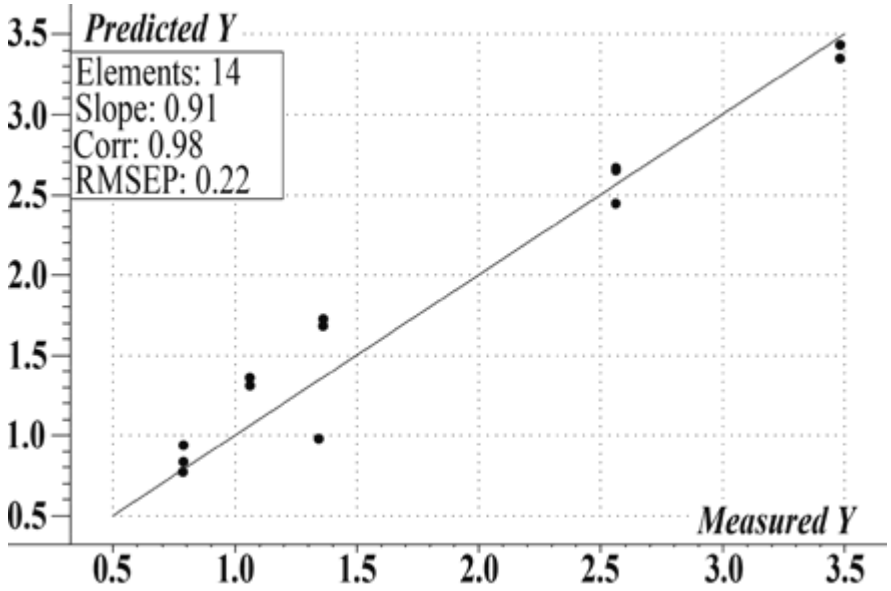


Figure 5. Example of predicted vs. measured plot. The measured (reference) Y-values are plotted along the x-axis, while the predicted Y-values are plotted along the y-axis. A standard y|x regression line is fitted to these pairs of data from which the slope and RMSEP statistics (amongst other) can be calculated.

This type of plot furthers an illuminative overview of the prediction performance: the degree to which a fitted regression line has a slope as close to 1.0 as possible is a measure of the overall prediction accuracy, while the "spread around the model" (around the fitted regression line) gives rise to the RMSEP measure of the prediction precision.

The correlation coefficient, r^2, is another, alternative measure of the precision, proportional to RMSEP:

$$r^2 \propto 1 - (\left|\text{RMSEP}\right|)^2$$

CROSS VALIDATION

Sometimes it may not be possible (and/or desirable) to get as many samples as required for a test set validation. For such situations cross validation *can* be used, although at a price (TOS-sampling variance for the second data set – standing in for all future data sets – is left out). Thus it can never be *as* realistic, since it is

only a simulation of test set validation. The principle behind cross validation is simply that a model is made from part of the data set, while the rest is used for testing. When determining model dimensionality this should also be taken into account, especially with regards to overfitting.
Two different versions of cross validation are typically used: segmented and full cross validation.

Segmented cross validation is used whenever a relative large number of samples are available. The number of segments ranges between 2 and n, with n being full cross validation. The segments function as virtual, temporary small test sets. Selecting the appropriate number of segments can be difficult and it would be nice if universally applicable rule of thumb could be put up. This is not possible, however, since such a selection must be based on the size of the data set, as well as a priori knowledge of the data structure itself.
If two independent data sets cannot be obtained, cross validation must be used; the second best situation is then to obtain a calibration data set being twice as large as necessary for modeling. In that case it is possible to use "test set switch"[e], a segmented cross validation approach, where the data set is simply divided into two with modeling of the first part and subsequent validation using the second part and vice versa. In the end, the prediction error is found as the mean error from these two validations. This approach is (only) formally similar to test set validation, since this is also missing the random (TOS) sampling variations that only can be captured using an individually extracted second data set.
When it is not possible to get such a large data set; in these situations "real" cross validation most come into play.

Full cross validation, or leave one out cross validation, can also be considered a segmented cross validation approach – only with one sample in each segment. Full cross validation is used when the data set is extremely small or whenever it is not desirable to use segments. This is, however, seldom the case, only when the span of the data set will be severely affected by putting specific samples into one segment. Using full cross validation on large data sets will eventually lead to over-optimistic results, because the left out sample will on average give a significantly smaller (simulated) sampling variance contribution for increased data set sizes.

---

[e] A singularly bad term only kept for historical reasons; there is manifestly no test set present!

When cross validation is used on measurement replicates, one should select the cross validation type carefully. Running full cross validation with replicates in the data set will inevitably result in an artificial low validation variance due to the validation of samples very similar to those already included in the model!

LEVERAGE CORRECTION

With leverage correction a direct copy of the training set serves as a test set with a leverage correction for each individual object prediction error estimates. Leverage correction can therefore also be considered a weighted, one segment cross validation.
This validation approach may be used in the initial modeling stages instead of cross validation or test set validation, but never in the final validation of a multivariate calibration model, since leverage correction is only an approximate prediction error assessment and since this validation procedure almost always gives over-optimistic predictions. The only reason for this method being so widespread in method validation is that it is extremely fast, because the corrections involved can be computed simultaneously with the modeling stages. For this reason an initial leverage corrected model will be exactly identical to the test set or cross validated versions, only the validations are different. However, since the validation is supposed to give information about the optimal number of components, ultimately alternative validation procedures *may* end up as different models.

SUMMARIZING

Leverage corrected validation and cross validation use the same set of samples (calibration data set) both for establishing the calibration model and for testing its predictive abilities – though differently. Both these approaches unfortunately often lead to over-optimistic predictions [1].

Test set validation, on the other hand, is the only option for estimating a realistic measure of the predictive ability of a model, since a completely new data set is used in this validation procedure. One universal validation method for predictive purposes does exist; test set validation with samples extracted according to TOS.
The one and only case where it is methodically sound to use other validation approaches, is in comparative studies, where different versions of a prediction model are to be compared. In this special context it seems reasonable to use a

well-reflected, identical version of cross validation. However, *the* final model should always be tested on a completely independent data set.

For all purposes, the data sets should be as realistic as possible in order to be representative for current and future situations. They must also span the X and Y space as widely as necessary for the final validated model to do justice to the prediction purposes. This goes hand in hand with the fact, that the span of a calibration model only rarely can be expanded with confidence by extrapolation.

2.3 THE ANGLE MEASURE TECHNIQUE (AMT)

The Angle Measure Technique (AMT) is a signal characterization method originally developed to characterize the complexity of 1-dimensional data. Although this generic method can be used on all one-dimensional measurement series [5], within chemometrics it has so far been used mostly for extracting textural features from images. This thesis deals only with image analytical AMT applications.

AMT is potentially able to transform a two-dimensional image into a one-dimensional complexity spectrum without losing essential textural information [4,10], although the specifics of the unfolding operations involved (see chapter 3.1), can be important. Provided the image is essentially isotropic, a condition certainly fulfilled by for example paper master roll or sand images, an important simplifying pretreatment transform from the image domain to the scale domain is unfolding the image.

A typical color image consists of three (or more) individual gray-scale images (multi-spectral imagery consist of many more channels); while AMT works on univariate (grey-level) images (see chapter 3.1 for an introduction to imagery). Thus a critical step preceding AMT is selection of an appropriate channel, or several channels when each is carrying specific information. Often the loading relationships from a PCA decomposition of the spectral space (MIA) can be used for this purpose. Papers I and IV covers all aspects regarding channel selection.

Each channel is geometrically two-dimensional with I x J pixels and is separately *unfolded* [7,8] before being subjected to AMT. This unfolding procedure can be carried out in several different ways; primarily as linear unfolding, spiral unfolding or snake-wise unfolding. These techniques are all illustrated in Figure

10. The unfolding eventually leads to a one-dimensional discrete measurement series.

After affixing a "connecting line" between each individual datum (pixel), the measurement series is often referred to as a "digitized line" in the AMT parlance; all digitized lines are routinely subjected directly to AMT. Figure 6 shows a conceptual illustration of the AMT transform on a 1-D measurement series (in which the digitalization unit is so small that the connecting line appears smooth and continuous, but this is only a matter of scale) .

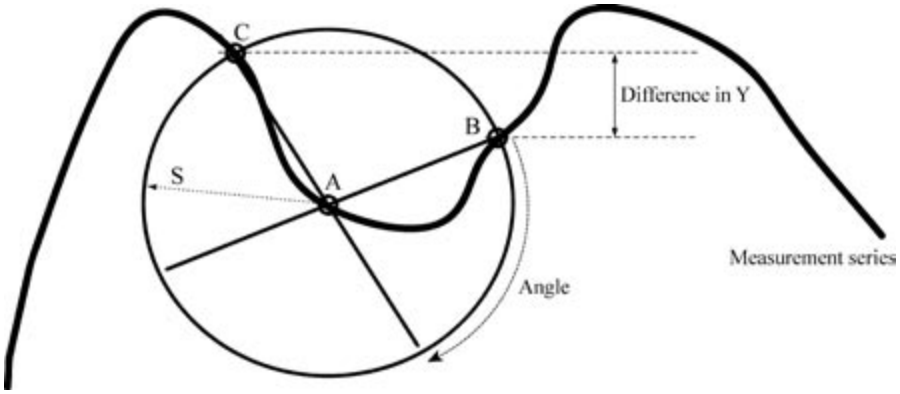


Figure 6. The concept behind the AMT transform. A circle centered on A, with contemporary radius S, intersects the measurement series/digitized line in two points: B and C, which are used to define both AMT's trigonometric "Angle Measure" as well as the Y-direction difference measure of complexity of the series.

A number of center points, A, are randomly chosen along the entire measurement series. This number is highly dependent on the specific context [8-10]. For each such point A as origo, a circle with radius S is drawn. Radius S is the contemporary scale of interest; it can be viewed as a scale parameter with which to gauge the complexity of the measurement series. This circle will intersect the measurement series in two points: B and C, which defines both the angle CAB (for computational reasons the supplement to angle CAB is always used) as well as the difference in Y measure of complexity for the series. An overall grand average for both angle as well as the Y-direction complexity measure is calculated (over all the pertinent A-points distributed across the measurement series), called the "mean angle", MA, and the Mean Y-Direction measure, MDY, respectively.

The contemporary scale is now increased from scale S to S+1 and the average MA and MDY measures are again calculated for all A points - and so forth. The scale of interest will run in the interval [1, N/2], where N is the total number of data in the measurements series; often a smaller scale range will be significant.

With a sufficiently large number of both initial points A, as well as corresponding scales (250 to 500 is typically used, but depending on the textures and other features in the images to be characterized, it can be increased with no problem), it is always possible to obtain statistically robust mean angles, MA, and mean difference in Y, MDY.
These operations constitute a transformation from the two-dimensional image domain, via the unfolded measurement series, to a new scale-domain: For each measurement series, one can construct a MA (and/or MDY) versus scale relationship[f], which is most often depicted in a simple two-dimensional scatter-plot:

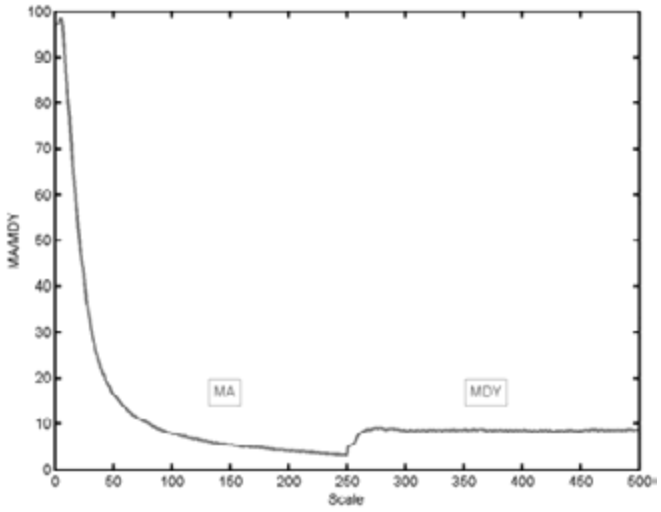


Figure 7. MA/MDY-scatter plot for paper sample (type star with matt surface finishing – see Paper III)[g].

---

[f] A so-called complexity spectrum.
[g] This scatter plot can be considered a spectrum, but *only* as a visual analogy as the X-axis is not phenomelogically equivalent to e.g. a wavelength interval in spectroscopy.

AMT complexity spectra are to be viewed as the mean angular (or vertical) change as a function of the set of increasing scales S. Therefore AMT can be viewed as a new domain transform, which characterizes the data series complexity as a function of scale [5,8].

Thus AMT functions analogously to other domain transforms, e.g. the well-known Fourier transform, but where the Fourier transform returns a frequency transform, AMT returns what could be termed a complexity spectrum (for each image in this case).

In general, AMT complexity spectra constitute a synoptic complexity characterization of any measurement series as a function of the MA/MDY versus scale relationship. The complexity measures (MA and MDY, respectively) are able to vary freely with scale, thus enabling AMT a much larger potential to express essential complexity relationships than for example fractal descriptors, which rigidly assume a constant fractional dimension (measure of complexity) for all scales (sic). Although AMT spectra can be used (and often is) with an aim of interpretation of the intrinsic complexity and scale-dependencies in their own regimen, direct post-processing of AMT-spectra has been much used the last 10 years [5,8-17] related to the very powerful chemometric multivariate calibration possibilities, most often in the form of PLS-regression.

3 IMAGE ANALYTICAL TECHNIQUES

Images are most-often analyzed by traditional image analysis, Multivariate Image Analysis (MIA), or Multivariate AMT Regression (MAR). However, prior to any usage of these analytical techniques, it is often necessary to preprocess the images using image processing techniques.

3.1 THE IMAGE

Essentially, a grayscale image may be viewed as a 2-D array of pixels with row- and column-indices [I,J] specifying the row and column in the array of pixel values. Each pixel can be described by an intensity value and an X- and Y-position in the spatial (image) plane - see Figure 8.

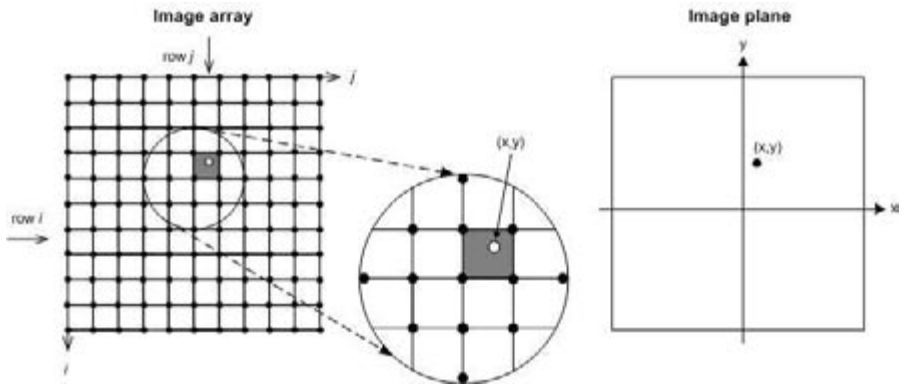


Figure 8. Presentation of relationship between image array indices and image plane coordinates (spatial image). A normal image is represented in the image plane, while the computer considers the digitized image as an image array.

If the image consists of more than one image channel and hence is multivariate - multitemporal with several images of the same motive acquired at various times or hyperspectral with several images at different wavelengths – it can simply be viewed as a stack of image planes. The image array hence becomes defined as [I,J,K] with K being the number of planes (channels) within the image. A typical 3-D color image (RGB) is digitally represented as shown in Figure 9.

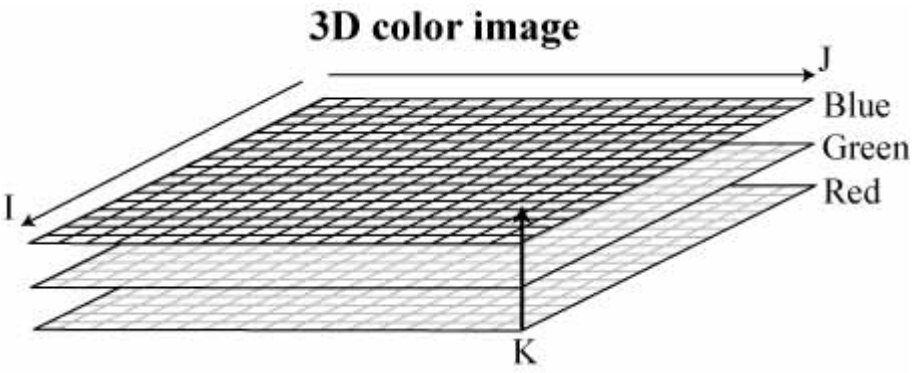


Figure 9. Digitized 3-D color image with image array indices (I,J,K).

UNFOLDING

In certain situations it may be useful to reduce the image rank, for instance by changing a three-way image array into a two-way array, which is typically done by simple reorganization from the [I,J,K] array into a [I×J,K] array. Figure 10 illustrates such reorganization – also know as vectorizing, reshaping, or unfolding – of a 2-D image (grayscale).

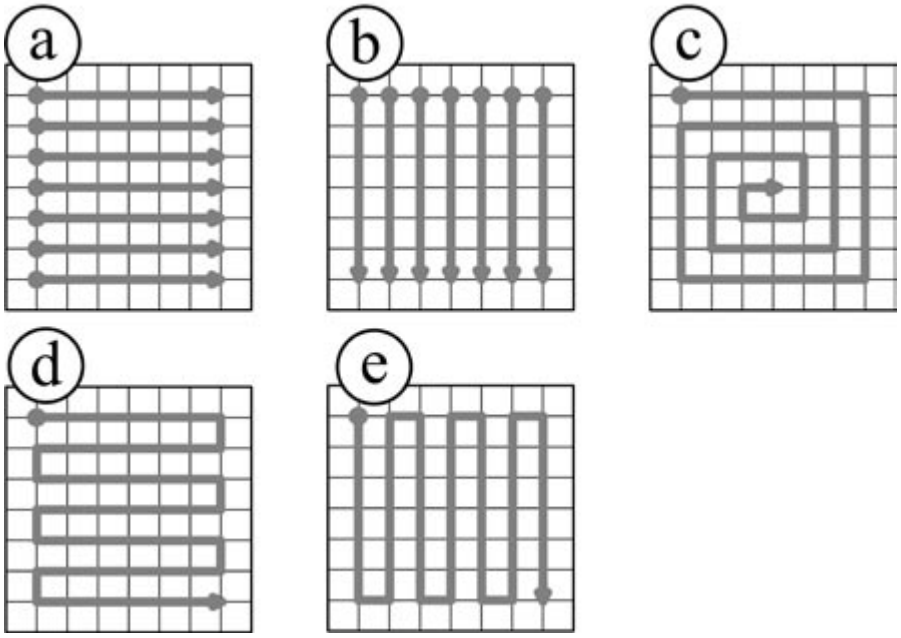


Figure 10. Unfolding a two-dimensional image. Figures a and b shows *the* classical row- or column-wise unfolding technique´(colloquially known as "Chop chop"), figure c spiral unfolding, whilst figures d and e shows a snake like unfolding technique.

Traditionally, unfolding is carried out by vector-wise concatenation (row-by-row or column-by-column juxtaposing) as illustrated in Figure 10a and b. However, other unfolding schemes also exist, which eventually also leads to one-dimensional discrete measurement series. Each such measurement series (unfolded image) can now be viewed as a digitized spectrum. A K-spectral image has dimensions IxJ with K individual unfolded image planes.

3.2 IMAGE (PRE-)PROCESSING

Under ideal conditions an image can be used directly in image analysis. However, image parameters such as insufficient contrast, lack of focus, artifacts, and shading effects very often prevent straightforward image analysis. In such situations image preprocessing can be used to compensate for the plethora of image deficiencies and hence allows better analysis of features of interest.

3.2.1 SHADING EFFECTS

In some cases images may need some type of shading correction. This correction takes away errors caused by the optical system (such as inhomogeneous illumination). Figure 11 illustrates how illumination gradients can be observed.

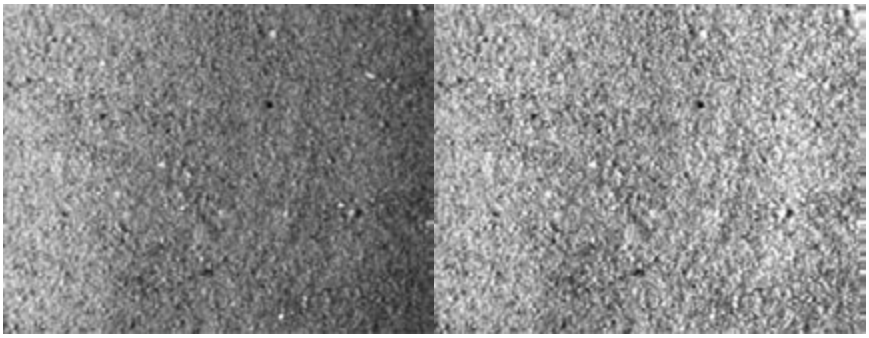


Figure 11. Sand sample with (left) and without (right) illumination gradient. Both images show the exact same sample.

Several different techniques based on different algorithms a.o. can be used to remove shadings [18-23]. In Paper IV a new technique was introduced, which is capable of estimating and removing illumination gradients generated by point source illumination – Figure 11 illustrated a sample with the illumination gradient and the exact same sample – only with the gradient removed by the new gradient removal technique.

The gradient seems to be linear in the horizontal direction and non-existent in the vertical direction. This makes is possible to calculate the average pixel intensity per column in the 2-D image, resulting in a function as illustrated in Figure 12.

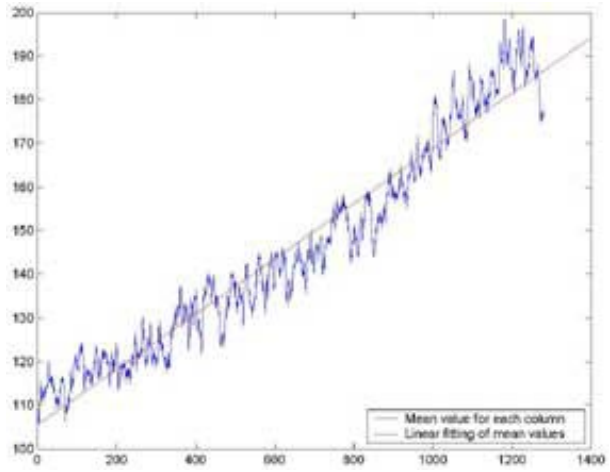


Figure 12. The mean pixel intensity (y-axis) for each column in the image (x-axis), which makes it possible to find a fitted linear regression line.

By use of least squares regression a linear function describing the illumination gradient can be obtained. This function can be computed into a correction image by assuming that the gradient only exists in the horizontal direction, as illustrated in Figure 13.

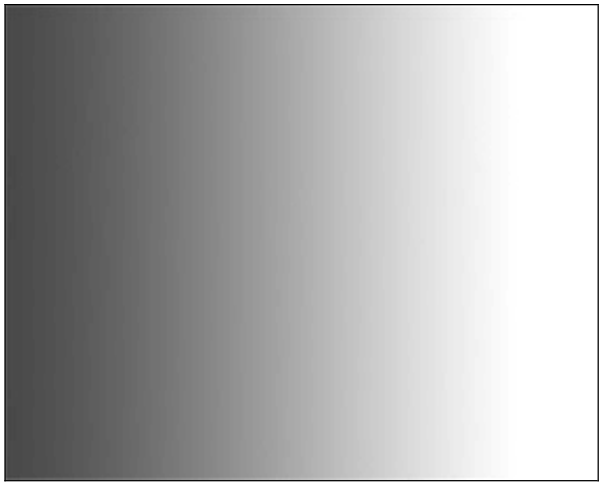


Figure 13. Example showing a linear gradient correction image. The linear gradient function shown in Figure 12 can be extended into the above image by row-wise duplication with all rows being identical.

If the correction image and the original (but shaded) image are divided, a gradient corrected image can be computed (the right picture in Figure 11).

3.2.2 IMAGE CONTRAST

Some images might have very low initial contrast primarily caused by non-optimal illumination and camera exposure. Such limitations might hinder proper use of image analysis and eventually leads to bad or non-interpretative results from analysis.
The proper solution to inadequate image contrast is reacquiring the images; however, often this is neither possible nor desirable. In such cases an intensity adjustment can be useful, because this remaps an image's intensity values to a new range – typically to cover the full dynamic range - thus leading to an improved image contrast.

In addition to having several shortcomings with regards to unwanted effects, an image might also have incorrect luminosity. Images often lack contrast due to either low or high luminosity. This is very often the case for images corrected by the gradient removal procedure described above.

If the steepness of the illumination gradient[h] function is too high, an intensity adjustment can be a necessity. If such an adjustment is not applied, the image often becomes too bright. This follows from a simple division of the pixel value by the linear gradient function, since a pixel value around maximum (typically 0/255) is often obtained.

A histogram[i] for a gradient image and a corrected image would display the differences caused by the gradient correction – see Figure 14.

[h] Procedure described above.
[i] An image histogram is a chart that shows the distribution of intensities in a grayscale image.

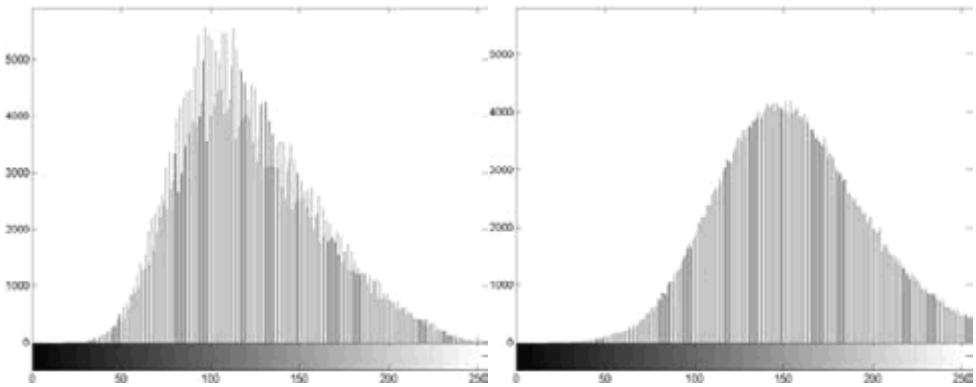


Figure 14. Differences in grayscale distribution with and without a gradient correction (for the images shown in Figure 11). The left image shows the histogram for the original, uncorrected image, while the right image shows the smoothed histogram for the corrected image.

From Figure 14 it can be seen, that the gradient correction also reduces the amount of noise. Sometimes the distribution of the pixels is also changed, resulting in a change in overall brightness. These changes are a consequence of the gradient correction procedure. Such changes do certainly affect any regression model, but do not necessarily result in worse models.

3.3 IMAGE ANALYSIS

3.3.1 IMAGE ANALYSIS

Traditional image analysis is widely used within the industrial sector for qualitative and quantitative characterization in many different areas. Information is extracted from images - usually from grayscale (2-D) or color images (3-D) - by means of various digital image processing techniques – typically in either the spatial or the frequency domain. Techniques such as thresholding, segmentation, and various filters are very common within image analysis [24-25].

3.3.2 MULTIVARIATE IMAGE ANALYSIS (MIA)

While traditional image analysis most often works in the spatial or frequency domain for 2-D (grayscale) or 3-D (color) image arrays, the spectral or time

domain become a higher priority when images become multivariate or multitemporal. When this is the case Multivariate Image Analysis (MIA) becomes very useful [26-27].

Multivariate image analysis is primarily used for explorative analysis and works by decomposing the unfolded image space into a principal component space – a latent variable space – by principal component analysis (PCA).

3.3.3 MULTIVARIATE AMT REGRESSION (MAR)

Multivariate AMT Regression (MAR) is a method for multivariate calibration between AMT complexity spectra (X) and a dependent variable (Y), for example denoting reference analytical results.
This approach was pioneered by Jun Huang in a series of works, which have been a major inspiration for the present work [5,14-17]. The regression equation in general is given as:

$$\hat{b} = A^{+}y$$

with A^{+} denoting the generalized inverse of the AMT spectra. Any multivariate regression method can be applied. Figure 15 shows the entire process behind MAR using Partial Least Squares (PLS) as regression method [5,15].

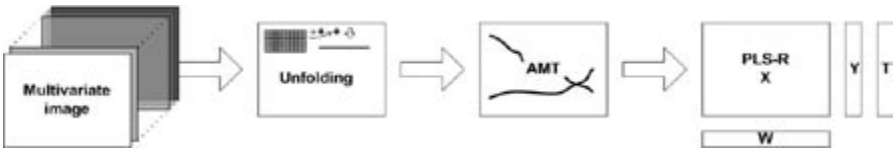


Figure 15. Procedure for Multivariate AMT Regression (MAR).

Each image plane is processed individually with unfolding, AMT, and subsequent PLS regression.

4 IMAGE ANALYSIS AND SAMPLING

Image Analysis (IA) is a widely applied industrial technique for process and product control. Many image analytical approaches and techniques are based on an unsubstantiated *assumption* that the samples (images) examined are *representative* from the bulk sample underneath the imaged surface area. This is an assumption, which is almost *never* realistic however.

To understand why this is so one needs to understand and appreciate the fundamental issues involved for representative sampling – enter the Theory of Sampling (TOS). In the following TOS is first reviewed to the extent needed here, and subsequently applied to the image analysis scenario.

4.1 THE THEORY OF SAMPLING (TOS)

The sole purpose of sampling is to reduce the mass of a lot and obtain a sample which is still representing the initial material with sufficient accuracy and precision (TOS gives strict definitions on these attributes, see further below). Any sampling – which leads to an experimental estimate of the true lot content – forms the basis for the product quality that ultimately is desired to be both consistent and satisfactory. It is always easy to extract an(y) indiscriminate lot portion and to analyze it with the best analytical method available. Nonetheless, any such extraction and the subsequent analytical result, will in practice be worthless if the sampling process is not representative, i.e. the particular lot portion is but a specimen[j].

Representativity is intimately related to the concept of 'heterogeneity', as will be presented below.

[j] According to TOS incorrectly extracted, and hence non-representative mass-reduced lot portions should be termed 'specimens' - to clearly distinguish these from truly representative 'samples'.

4.1.1 HETEROGENEITY

If we look at a heap (e.g. stockpile, railroad or truck load etc.) of material at a certain distance, it might at first glance appear 'homogeneous', i.e. all constituents[k] are apparently identical. As we move closer to the heap however, it will be obvious at some point that the material in reality is *not* homogeneous; Visual differences from some parts of the heap to another will often be observable. Moving down to the level of each fragment one will discover that the material indeed is heterogeneous. All fragments will in practice *never* be strictly identical; different composition, shape, size, density, etc. occurs. Within sampling, we might as well remove the term 'homogeneous' from our vocabulary, since this is a theoretical limiting case only, alas never encountered in reality for any naturally occurring material.

Representative samples can only be obtained by following the specific guidelines laid down in TOS. This is a rigorous, comprehensive theory developed by Dr. *Pierre Gy*, which describes all aspects of how to establish a representative sampling process from which the results can be demonstrated to be samples with the desired characteristics w.r.t. the lot. TOS makes use of a number of well-defined sampling errors – which will be duly explained below.
When we are sampling we are replacing the true lot content a_L with its estimator a_i. This leads to a selection error:

$$SE = \frac{a_i - a_L}{a_L} \quad [28,29]$$

The representativity, r, of the sampling process is defined as the sum of the mean SE[l] and the variance of SE:

$$r^2(SE) = m^2(SE) + s^2(SE) \quad [28,29]$$

with *r* being the representativity of the selection error, *m* the mean, and s^2 the variance.

---

[k] Constituent: The smallest elements that can be considered to be indivisible during the selection process. In sampling also known as *fragments*, which can be grains, particles, molecules, ions.
[l] $m^2(SE)$ is the bias.

According to TOS, representativeness is characterized by the absence of bias (accuracy) and by the particular sampling variance pertaining to the sampling process involved (TOS prefers to speak of reproducibility). Having reached an unbiased sampling process (a *correct* sampling process), it is now up to the sampler to secure that the remaining sampling variance (reproducibility) reaches a sufficiently low level – no matter how heterogeneous the material is. This is the task of sampling process optimization. It cannot be emphasized enough however, that the prerequisite for all successful sampling is that the particular process contemplated must be accurate, must be unbiased. It is not possible for any sampling process, once performed, to correct for sampling bias, as this is in principle un-estimable [28-30].

The primary objective for any sampling process is therefore to fall in line with the necessary TOS specifications needed in order for the results of the process to be unbiased.

A material can be heterogeneous at two distinctive different levels:

- Constitutional
- Distributional

CONSTITUTIONAL HETEROGENEITY (CH)

The constitutional heterogeneity (CH) is an intrinsic property of the material under consideration and units[m] making up any such material are considered unalterable and indivisible and thus stable in a physical and chemical environment. Indeed the units must be stable also w.r.t. the sampling process itself; i.e. the sampling process should not lead to fragmentation of any of these units. If it does (for example when sampling friable materials), TOS conveniently speaks of 'fragments' as the smallest stable inseparable units, allowing us to contemplate all types of sampling process and materials.

A lot (material) has a homogeneous constitution only when all units are strictly identical, while the constitution of a material (lot) is heterogeneous if all units are not identical. However, all units do not necessarily have to be completely identical, as the understanding of constitution heterogeneity depends on whether

[m] Definition of unit

we for example are considering one specific material property or all of them simultaneously.

The amount of heterogeneity carried by one fragment, F_i, is defined as:

$$h_i = \frac{(a_i - a_L)}{a_L} \cdot \frac{M_i}{M_{\bar{i}}} \text{ [28, p.65]}$$

The constitutional heterogeneity of all N_F fragments in a lot can be described as:

$$CH_L = s^2(h_i) = \frac{1}{N_F}\sum_i h_i^2 = N_F \sum_i \frac{(a_i - a_L)^2}{a_L^2} \cdot \frac{M_i^2}{M_L^2} \text{ [29, p. 67]}$$

The crucial conceptual element is that CH_L is defined as the variance of all fragment heterogeneity contributions making up the entire lot.

If one lot portion is extracted alone (below to be termed a 'grab sample'), this can never constitute a representative sample by itself, due to the intrinsic constitutional heterogeneity displayed by any heterogeneous material at large. Any two such specimens can never give rise to identical analytical results, indeed any repeated sampling, say 10 times, per force must give rise to a distribution of analytical results. The specific sampling error arising from this non-constant material heterogeneity is termed the Fundamental Sampling Error (FSE). Constitutional heterogeneity is a material characteristic related to the differences in composition *between* the fundamental, smallest inseparable fragments making up the total lot. Since this heterogeneity is intrinsic, mixing and 'homogenization' has no effect on it.

Distributional heterogeneity, however, is affected by mixing.

DISTRIBUTIONAL HETEROGENEITY (DH)

The distributional heterogeneity (DH) can be observed when comparing different *groups-of-neighboring fragments*, while only differences between these individual units were considered regarding the constitutional heterogeneity. Distributional heterogeneity is a concern when we sample *groups of fragments* (DH is nil if 'samples' consist of single fragments only). In practice samples of course never consist of singular fragments only, and DH can therefore be viewed

– and used – to characterize also the added uncertainty stemming from the practical necessity of sampling *some* volume – always consisting of more than one fragment. This sampling volume shall be known as the *increment* volume (or the *increment* mass) in the following.

This heterogeneity depends on the constitution heterogeneity, but it especially reflects the specific *spatial distribution* of the fragments, but more importantly of the groups (increments) in the lot. Furthermore the size and shape of the lot also affects DH_L, since this aspect of the lot heterogeneity is affected by the omnipresent gravitational force pervading our environment. By definition, a lot has a homogeneous constitution only when all groups of a strictly similar size have strictly identical composition, which of course can never be the case – since even the compositional heterogeneity is never zero for naturally occurring materials. If the lot is subjected to mixing and homogenization however, DH_L[n] is reduced; by contrast segregation leads to an increase in DH_L.

The heterogeneity carried by a group-of-fragments within the lot can be described by an identical equation as for the fragments. Observe however the critical step-up in scale, from fragments to groups:

$$h_n = \frac{(a_n - a_L)}{a_L} \cdot \frac{M_n}{M_{\bar{n}}} \quad [29, p.65]$$

where M_n is now the mass of the group and $M_{\bar{n}}$ the average mass of all groups in the lot (for a theoretical analysis, we may assumed these quantities are fully known). The distributional heterogeneity, DH_L, is now defined as the variance of the heterogeneity between all groups, $s^2(h_n)$.

DH_L can also be expressed as a direct function of the constitutional heterogeneity (CH) and of the spatial, distributional heterogeneity, which is characterized using Y, a grouping factor (equal to the average number of fragments in each group in the lot) and Z, a segregation factor (1 for fully segregated material and 0 for completely homogenized material):

---

[n] DH_L is the lot Distribution Heterogeneity.

$$DH_L = s^2(h_n) = CH_L \cdot \frac{1+YZ}{1+Y} \quad [29, p.65]$$

From the above formula it can be seen that distribution heterogeneity only exists if the lot constitution indeed were homogeneous. Furthermore it can be seen that $CH_L \geq DH_L$, i.e. the distribution heterogeneity is always smaller than the constitutional heterogeneity. With all materials being essentially heterogeneous, sampling operations will thus inevitably result in a compound sampling error. The added error (exceeding FSE) associated with the spatial distribution of the groups-of-fragments within the lot material is termed the Grouping and Segregation Error (GSE).

DH, as delineated above, not only depends on both CH and the spatial distribution of the units (characterized by *Y* and *Z*), but also on the lot dimensionality.

HETEROGENEITY AND LOT DIMENSIONALITY

Strictly speaking, of course all lots display a three-dimension nature, but one or two of these dimensions can in practice often be regarded as much less important; indeed they may at times be completely disregarded provided the increments involved span this/these dimensions completely (and correctly) - the effective lot dimensionality will thus be reduced. TOS classifies lots as three-, two-, one-, or zero-dimensional bodies.

Zero-Dimensional Lot

Zero-dimensional lots can at first be difficult to comprehend based on their apparent non-existing dimensionality in reality, but this term is simply meant to signify typical statistical populations consisting of a large number of non-ordered, non-correlated, and random units, fragments or groups, depending on the scale contemplated. A zero-dimensional body could be a large number of truckloads, railroad cars, or barrels, which have lost their internal relationship or ordering (geometric or chronologic) and thus must be regarded as random and non-correlated units. In practice, any heap of particulate, free-flowing material which can be manipulated in the sampling process would be classified as a zero-dimensional body.

Sampling of a zero-dimensional lot material will vary as a consequence of:

- random fluctuations between fragments – and between groups

- errors generated as a results of the sampling extraction of groups[o]
- grouping and segregation (the spatial distribution of heterogeneity)

Assuming that sampling is carried out correctly ("correctness" is defined in chapter 4.1.2), the errors associated with the material alone are:

$$CSE_1 = FSE + GSE$$

This heterogeneity, h_1, in zero-dimensional lots can be classified as *small-scale* fluctuations covering the specific constitution of fragments and groups. In zero-dimensional lots we therefore only deal with a comparison of fragments and groups of fragments, while we also focus on long(er)-range variations for one-dimensional lots.

One-Dimensional Lot

These lots are typically very long with width- and depth-dimensions of vanishing importance. Such lots are either elongated, continuous material streams or non-random and discontinuous units. Examples include series of truckloads and barrels as well as materials on conveyor belts and in pipes.

Industrial (manufacturing, processing) one-dimensional lots are often generated by chronological operations, wherefore they will reflect processing fluctuations – always including the small-scale fluctuations (h_1), but also from *long-range* (h_2) and *cyclic* (h_3) fluctuations. These lot fluctuations on all scales lead to the following full complement of errors for correct sampling in one-dimensional lot case:

$$CSE = CSE_1 + CSE_2 + CSE_3$$

Two- and Three-Dimensional Lots

The dimensions in three-dimensional lots are all of equal importance and the dimensionality can hence not be reduced. Examples of three-dimensional lots are geological occurrences, e.g. mineral deposits and large, isolated, piles such as shiploads, stockpiles which are too big to be manipulated without excessive efforts.

---

[o] These errors are aptly termed 'sampling process errors'.

For two-dimensional lots the depth-dimension can be considered of vanishing importance. Two-dimensional lots are often flattened three-dimensional lots such as flat mineral deposits or the content of a flat truckload; other examples would be from the environmental realm, for example when taking soils samples of constant depth (or from a constant depth-interval) for all increments.

.

Two- and three-dimensional lots are not easy to sample and typically also not feasible to sample correctly [28,29]. For large material lots – such as mineral deposits – geostatistics provides the beginning of a solution, but for small material lots such an evaluation is not feasible [29]. Since handling of two- and three-dimensional lots is strictly speaking not included in TOS, this thesis will henceforth not deal with lots having more than one dimension.

4.1.2 STRUCTURALLY CORRECT SAMPLING

As sampling from two- and three-dimensional lots are most-often not feasible, ultimately the sampling operator should try to convert such lots into zero- or one-dimensional lots. These conversions are all-too-often not carried out in practice, because it is argued that it is too expensive and/or difficult to perform a lot-dimensionality conversion. What is often forgotten in this process is the fact that we have no control of what we are analyzing when the sampling is not carried out with the sole purpose of getting a structurally correct sample.

> *If handed a sample it is impossible to tell if it is representative of the lot material from which it was extracted, or not. Consequently, all efforts should go into making the sampling process representative.*

THE SEVEN SAMPLING ERRORS

The sampling process is an error-generating process by itself at every sampling stage, including materials handling operations. These numerous errors can all be characterized using TOS in a very effective, systematic fashion.

Correct Sampling Errors (CSE)

Thus far all errors described arise from the heterogeneous nature of the material (FSE+GSE). These two errors are summarized as the *correct sampling errors* (CSE), and they result from a point-selection process in which FSE originates from the intrinsic nature of each fragment, and GSE originates from the

distributional heterogeneity, which, again, is caused by the practical necessity of sampling groups of fragments instead of individual fragments.

$$CSE = FSE + GSE$$

Incorrect Sampling Errors (ISE)

Since sampling essentially is a forced (groups-of-fragments) statistical selection process, we need to understand how to perform this selection without making compromises with regards to the sample representativity. Thus far we have tacitly assumed that we can always follow the basic principle of structurally correct sample extraction, i.e. all fragments of the lot are submitted to the selection process with an equal probability P of being selected.

The discrete nature of the fragments makes it difficult to obey this rule when performing the actual physical sampling however. The sampling process results in a materialized *increment*, which is that group of fragments which have been extracted from the lot in a single operation. This materialization is achieved by first defined the increment to be extracted – an operation termed *increment delimitation*. Having defined the delimitation of the increment, the increment must then be physically *extracted*. These two operations are error-generating processes leading to errors which are termed the Increment Delimitation Error, IDE, and the Increment Extraction Error, IEE respectively.

Increment Delimitation Error (IDE)

A structurally correct increment must be delineated properly. Dealing with a one-dimensional lot this means the increment must consist of a cross-section defined by two parallel (or curvi-parallel) planes. This ensures that no part of the lot is represented in higher or lower portions than any other in successive increments, as illustrated in Figure 16.

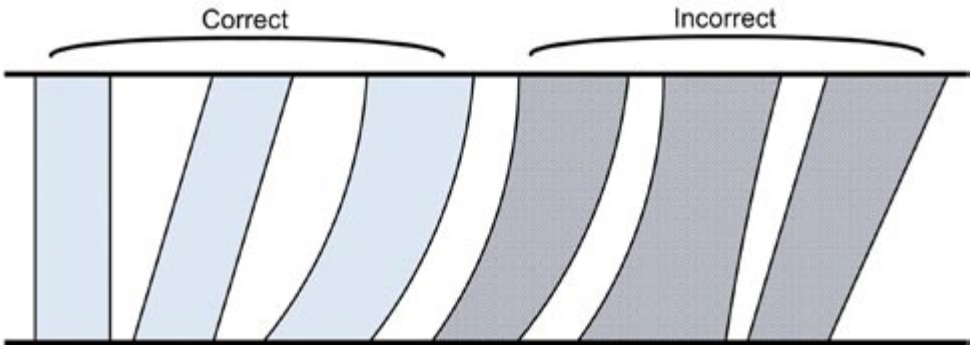


Figure 16. Correct and incorrect increment delimitation. The first three examples are all delimited correctly with parallel sampling boundaries, whilst the latter three all over-represents some part of the 1-D lot.

Increment Extraction Error (IEE)

Having delimited the prospective increments area, the material located within these boundaries must be extracted. In order to avoid generating an additional sampling error performing this operation, known as the Increment Extraction Error, IEE, it has to be ensured that nothing else than what is defined in the correctly delineated area is actually extracted. Even though a fragment has been defined as belonging partly to the increment and partly to the remainder of the lot, this fragment must not be divided in any way as it in TOS is considered indivisible during the selection process. In practice this means that fragments having their center of gravity included in the delimitated increment must be included in the extraction process and vice versa. This principle is known as the rebounding rule [28]. The actual extraction of an increment from a correctly delimited area is exemplified in Figure 17.

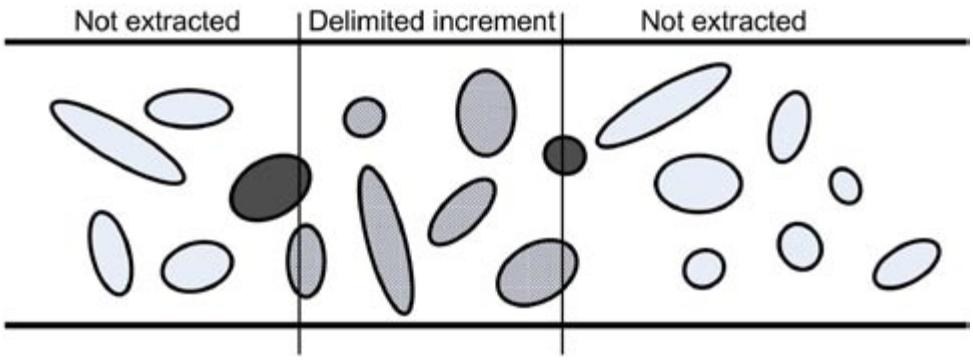


Figure 17. Increment delimitation and correct materialization. White fragments are not part of the delimitation, while fragments marked as checkered are all selected as belonging inside the increment. The gray fragments are on the other hand not selected – because their center of gravity falls outside the boundaries defined by the delimited increment.

In this illustration, four fragments are crossed by the delimitated boundaries of the increment; two are gray and two are checkered. The center of gravity for the two gray fragments fall outside the marked increment and should consequently not belong to the extracted increment, while the two checkered fragments correctly falls inside the increment.

IEE quantifies the errors associated with the physical increment extraction of the delineated area. Any deviation from the rebounding (or the center-of-gravity) rule will inevitable lead to an extraction error, which, according to Pitard [28], can be a "*substantial error with many types of materials*".

Increment Preparation Error (IPE)

The materialization procedure of the increment does not only include a delimitation and extraction of the actual increment, but also a handling of the increments between the sampling stages as well as all of the processes commonly known in the analytical laboratory as 'sample handling and preparation'. The errors which may occur here are all grouped under the Increment Preparation Error term, IPE. This includes errors incorrectly occurring during operations such as mixing, comminution, and transportation. Further handling operations generates additional errors also belonging to the Increment Preparation Error IPE, which could be caused as a consequence of loss of material, contamination, deliberate alteration.

Incorrect Sampling Error

The materialization error – known collectively as the Incorrect Sampling Errors (ISE) – is simply the sum of IDE, IEE, and IPE:

$$ISE = IDE + IEE + IPE$$

The CSE model alone does not take these errors into account. A correct sampling procedure is consequently defined by the demand:

$$IDE = IEE = IPE = 0$$

It is clear that elimination of ISE is very high on the agenda when striving to achieve an accurate, bias-free sampling process – in fact, complete elimination of ISE is the very criterion for a sampling process which is in statistical control; ISE := 0.

IDE and IEE are random variables, whilst IPE is non-random being generated accidentally and is hence non-statistical in nature

Total Sampling Error (TSE)

Thus far we have focused on how to perform representative sampling according to TOS, which has introduced both the correct and incorrect sampling errors, CSE and ISE. This combines to the Total Sampling Error, TSE:

$$TSE = CSE + ISE$$

which, in total, represent all the following errors - for each sampling stage:

$$TSE = FE + GSE + CSE_2 + CSE_3 + IDE + IEE + IPE$$

It should be carefully noted however that errors CSE_2 and CSE_3 are of concern only for one-dimensional lots, since both the cyclic heterogeneity, CSE_3 and the long-range heterogeneity, CSE_2, have no meaningful counterparts in zero-dimensional lots - and have more elaborate meanings in both 2-D and 3-D lots, which will have to be installed specifically for the cases at hand here.

Global Estimation Error

Thus an analytical result for sample S includes the analytical errors, Total Analytical Error, TAE, as well as all sampling errors stemming from the whole sample extraction procedure, TSE. For completion, the sum of these two errors is termed Global Estimation Error, GEE, and hence includes all errors associated with any given sample S:

$$GEE = TSE + TAE$$

In total, using N sampling stages, the following errors will be associated to a sample S:

$$GEE = TAE + TSE \Leftrightarrow$$
$$GEE = TAE + \sum_{n=1}^{N} (FE + GSE + CSE_2 + CSE_3 + IDE + IEE + IPE)$$

This thesis will not deal with TAE as analytical errors have been amply dealt with in the pertinent literature [31-34].

The connection between all the zero-dimensional errors can be visualized as shown in Figure 18.

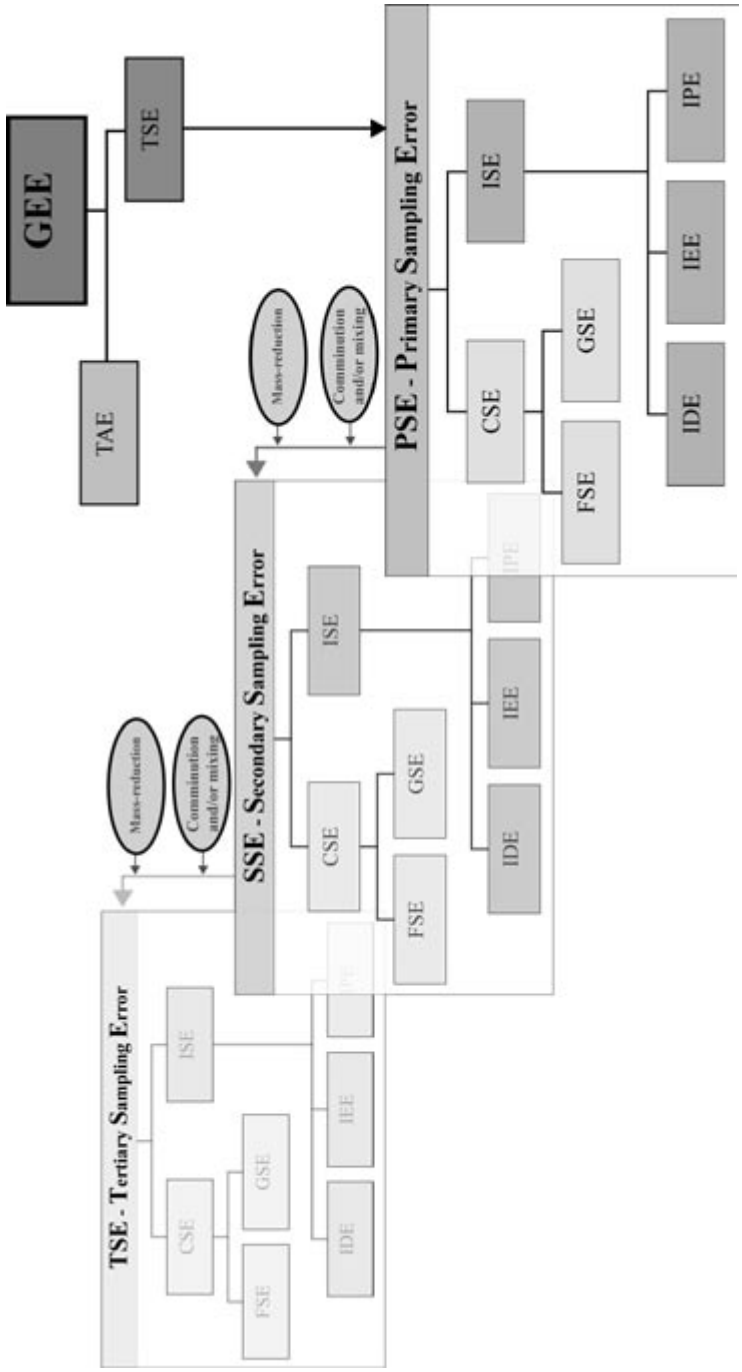


Figure 18. Visualization of the five zero-dimensional sampling errors. Several sampling stages and their individual, sampling error generating operations are often necessary to reach the final, optimal sample size.

This diagrammatic overview may at first sight appear complicated, but in practice all of these errors can either be minimized or fully eliminated by following the simple principles and practical guidelines for structurally correct (unbiased) sampling.

4.2 IMAGE ANALYTICAL SAMPLING

Instead of physical sampling, there are many situations in today's industrial practice, where it is desirable to rely on indirect sampling, also called sampling-by-proxy. The reason behind this is the need and desire for real-time, on-line (at-line) quality inspection or quality control etc. Many process times are so low that immediate feed-back from the products will be extremely beneficial for process control purposes.

Using image analysis as a means for indirect sampling as well as for analysis is in many cases perfect for on-line purposes. This way it is possible to rely exclusively on on-site calibrated image analytical models instead of direct physical sampling with subsequent off-site analysis. This opens up for the realm of applied Image Analytical Sampling (IAS), which is based on a 2-D rendition of the product/process surface, as supported by a relevant image acquisition system, suitably calibrated. In the following introduction to the IAS approach, we discuss generic imaging systems without technical details, while focusing on the IAS principles and parameters. IAS is aimed at analyzing, understanding, and making use of the imaging process (and the associated image analysis) – as seen from the perspective of the general Theory of Sampling (TOS).

Using an image analytical technique gives a lot of advantages, but it also has a number of disadvantages:

Advantages; image analysis is:

- Fast
- Inexpensive
- Simple
- Easy to automate
- Remote sensing
- Non-invasive

Disadvantages; image analysis is:

- Based on empirical reference data – these are subject to traditional TOS and TAE errors
- Sensitive to change in production parameters (image recording parameters …)

The fact that this methodology is dependent on reference sampling/empirical data when establishing prediction models makes it especially sensitive to changes in production/material parameters.

The IAS approach can be applied to all image analytical areas and several different imaging techniques can be used; for instance traditional grey-level image analysis, Multivariate AMT[p] Regression (MAR) [5,15], and Multivariate Image Regression (MIR) [26,35,36]. In addition, the images can be pre-processed using different techniques such as Fast Fourier Transform (FFT) [5,7] and Wavelet compression [5,7] – AMT can also be viewed as a preprocessing technique.

In this thesis the well-described image analytical technique MAR will be used to illustrate potential application areas for a new combined MAR/IAS methodology, which is mainly applied in the solids producing/processing industry sectors. The MAR approach can be used to quantify most product parameters, which are correlated to surface properties.

This is facilitated by data analysis of signals from problem-dependent (1-channel, multi-channel, VIS, NIR) camera sources. MAR works by extracting relevant information from accurate and precise camera data with subsequent multivariate calibration. Spectral image data (X) relate to the surface of the solid/product. Provided a satisfactory calibration can be set with regards to the reference results (Y), e.g. PLS-regression or similar, image analytical signals can then be used for prediction of the corresponding Y-data directly from new images thus making sampling with offline analysis unnecessary.

[p] Angle Measure Technique (see Chapter 2.3)

4.2.1 TOS ON IMAGE ANALYSIS: IMAGE ANALYTICAL SAMPLING (IAS)

IAS thus has the potential to replace actual sample extraction with subsequent analysis. The most important differences between these two approaches are:

- IAS is an *indirect analytical technique* and does not physically alter the material – IAS does not require a physical sample extraction.
- Physical sample extraction naturally involves analyzing in depth, analyzing the bulk material, while IAS per force focuses on the material surface alone – X-ray and ϒ-ray (neutron) analysis excluded. This thesis only covers optical (VIS and NIR spectroscopy) imaging.

Indirect image sampling methods is for example often used for analyzing compositional mixtures, particle size distributions, particle shape, color, physical characteristics a.o. Such methods can only yield good results for prediction if they are based on representative physical samples - for instance from a moving stream or an elongated lot. This process is denoted as Physical Reference Sampling, PRS, for which TOS stipulates all the necessary and complete requirements for such increments to be representative samples.

Estimating the analytical content(s) in such samples necessitates subsequent off-line analysis, which ultimately permits establishment of a useful[q] calibration model. Having established a satisfactory model should eventually result in a fully automated, remote-sensing monitoring facility with no need for further physical sampling!

In order to characterize any lot material, the necessary sampling procedures should be considered first. With a well-defined sampling scheme one has to decide whether to aim at sampling with time-consuming off-line analysis or whether to use a proxy sampling methodology such as image analysis (Figure 19).

[q] Useful is here defined as an accurate and precise prediction model as tradition in chemometrics.

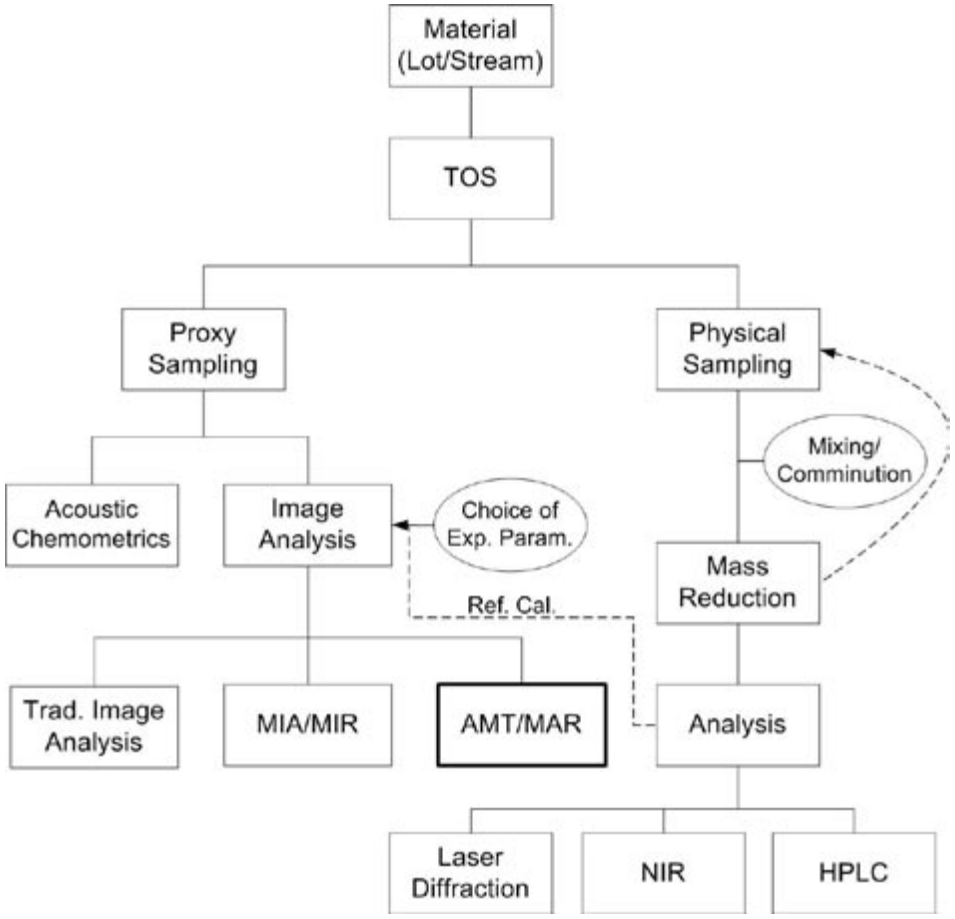


Figure 19. Overview of Image Analytical Sampling and multivariate calibration: IAS relationships with Theory of Sampling (TOS), image analysis, chemometrics and reference analysis.

IAS is critically dependent on representative reference sampling to establish a predictive model. TOS stipulates how to perform "structurally correct" sampling, representative sampling, which often includes several mass reduction steps, comminution/grinding, and homogenization (mixing) of the sample material. Eventually we end up with a sample representative of the lot material in a size suitable for analysis. Such samples therefore form the required basis for the reference calibration and are hence critical for the validity and reliability (accuracy and precision) of the final prediction model.

Pitted against this background, the cornerstone of the image sampling process, IAS, will be the definition of the imaging area, the Field-Of-View (FOV) and the image acquisition. A finite imaging FOV can for most lot dimensionalities be well-defined following general TOS-guidelines and the imaging process itself should be fairly easy to carry out in practice. Nevertheless, optimization of the imaging equipment with the sole purpose of obtaining optimized predictability of any material property is a realm of its own. The essential optimization areas are addressed and described below (see section 4.3.1 and Paper IV).

THE "STRUCTURALLY CORRECT" IMAGE

The imaging area is in most cases easily defined, as TOS delineates how to obtain representative samples – for all lot dimensionalities. In general, it has no relevance to address three-dimensional lots with any imaging technique, as the depth dimension always will be undetectable to some extent – regardless of the hardware (VIS, NIR, a.o).

IAS does not apply to three-dimensional lots

There might be conceived a few exceptions to this statement – mostly connected to materials with low density and/or high transparency – but in general, if IAS works on a material with three apparent dimensions, at least one dimension can be regarded as of secondary importance and the lot dimensionality is hence reduced in practice.

Zero-, one-, and two-dimensional lots can all potentially be described using an IAS technique. Image-wise these dimensions are all identical, because the image – hence the sample – will always display the 2-D surface delimited in the imaging area. Ultimately this means that each image can be considered two-dimensional, while all samples and lots can be classified as either zero- or one-dimensional – depending on the application.

If the lot is zero-dimensional the image should consist of the whole increment surface[r], while images in one- (conveyor belt, pipeline) and two-dimensional lots always consist of but fractional surface sampling areas selected according to a systematic or stratified random sampling scheme. The sampling areas for one-

[r] As zero-dimensional lots pr. definition consists of discrete units.

dimensional lots should furthermore be proscribed to contain the whole cross-section area of the 1-D process stream as illustrated in Figure 16 and Figure 20.

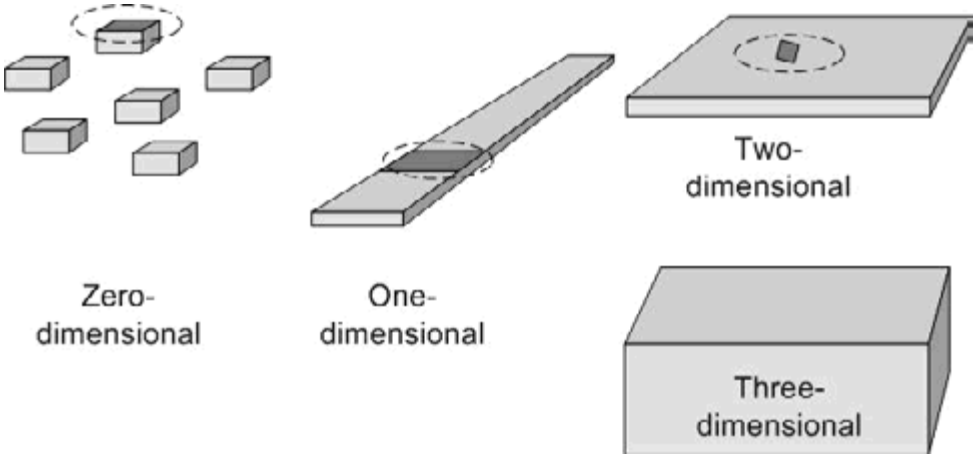


Figure 20. Imaging area versus lot dimensionality. Imaging areas are marked in dark gray for zero-, one-, and two-dimensional lots. Three-dimensional lots are not feasible to study using IAS and are therefore not marked.

All sampling schemes should initially include a heterogeneity characterization of the lot material. This characterization can then be used for defining optimal sample size and sampling frequency. Full definition of the various 1-D sampling schemes is beyond the scope of this introduction and referral is made to the pertinent introductions and explanations found in the TOS-literature [28,29,37].

The camera design very nearly always pre-defines the size and shape of the increment / measurement area, because the camera's optical system defines a sharply defined regular measurement area, the Field-Of-View. This area can however easily be narrowed or expanded – either through a simple imaging crop procedure effectively reducing the available data or through the use of several cameras/images for expansion of the measurement area – see Figure 21 for exemplification of various sampling schemes. Furthermore, Paper 4 details how extended sample sizes can be used in practice and how this influences model representativity.

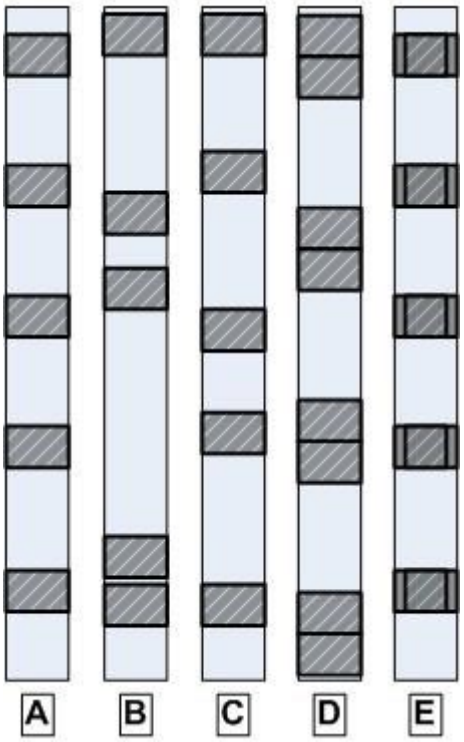


Figure 21. Illustration of various sampling schemes *vs.* imaging FOVs. Modeled imaging areas are hatched, while camera field-of-view is gray. Image sampling can be A) systematic, B) random, C) stratified random, D) expanded systematic, E) reduced systematic. Combination modes can also be conceived for special purposes.

Now, the imaging world is not quite this simple when models are to be established, as bulk samples never can be regarded as fully flat (with no depth dimension) and often there has been accepted a certain lack of correspondence between the top surface of the physical reference sample and the camera's optical footprint. This physical discrepancy between the image and the reference signals will necessarily contribute to producing model – and prediction errors which ultimately show up as unnecessarily large RMSEP.

> *The image is almost never fully representative of the bulk sample beneath it!*

This a critical success factor, not always sufficiently contemplated: How to reconcile the inherent mismatch between these different supports for the X-space (image) and the Y-space?

THE "REPRESENTATIVE IMAGE"

This mismatch comes to fore because the spectral image data (X) in practice almost always relate *only* to the surface of the solid/product[s]. In order to assume that this surface image is representative, it *must* be assumed that the two-dimensional area captured by the camera's field-of-view always represents the true grade of the three-dimensional bulk sample. This can only be true *if* no heterogeneity is present vertically, expressed by either grouping or the segregation. In other words, the bulk sample must be homogeneous vertically and any variation has to be described by the 2-D surface only. This will of course never occur in practice for most naturally occurring or produced materials, as FSE/GSE is always present in any bulk sample hereof.

On the other hand, there does also exist a certain class of products for which this depth-heterogeneity issue is either so reduced or practically eliminated, that this point is not relevant. Examples include potato chips, corn flakes, pizza surface ingredient density mapping, flakes, shards etc. It is a fair assumption though, that this class of specific products is distinctly smaller than its complement – but in any event the present IAS introduction only deals with the general case, in which physical reference samples must be extracted in order to establish an image analytical prediction model.

PHYSICAL REFERENCE SAMPLING, PRS

The physical reference samples, PRS's, are easy to identify and most-often also to extract with the sampling practices well covered by TOS. A procedure for identification and extraction of representative samples can always be established by rigorous use of TOS. Any such sampling scheme involves optimal sample location, number and frequency, and size(mass) of the increments to be taken. Extraction should make use of all the relevant TOS-principles needed to provide representative samples for analysis.

---

[s] Excluding the camera-sources which *can* in fact look below the upper material surface.

Physical extraction of samples often involves several other steps before a final sample can be obtained. The first extraction often leads to larger samples than needed or desired. Depending on the material being sampled, grinding/comminution and/or homogenization may need to be performed, which is often followed by a mass reduction step. This mass reduction can be either another physical extraction or a mass reduction step by use of – for instance – a riffle splitter (see paper 5 for an introduction to mass reduction). Following the final mass reduction step having reached a final sample of a given size (weight, volume), this is subjected to analysis being any analytical method.

Using *proxy*-methods[t] it is only necessary to physically extract and analyze a set of calibration samples. But proxy-sampling only allows characterization of the set of reference samples if based on a proper multivariate calibration in which the calibration data set has been sufficiently controlled in both X- and Y-space. Establishing an image-PLS model is comparatively straight forward, as is proper validation provided it follows the rules set forward for test set validation [1]. With IAS this proviso becomes particularly important, because of the potential 2-D (image) vs. 3-D support mismatch however. This issue forms a specific new element in IAS w.r.t. traditional physical TOS.

4.3 SAMPLE REPRESENTATIVITY IN IMAGE ANALYSIS

The critical issue in IAS is related to the degree of match, or mismatch, between IAS's 2-D rendition of the surface of a product or process stream relative to its bulk 3-D characteristics. Materials which are not segregated in their natural state (high-viscosity bulks, slurries, solids), can be directly imaged – and IAS characterizations will in general be representative. Materials prone to for example gravity segregation obviously need to be homogenized before presentation to the camera and the only possible counteraction for this is to install a mechanical homogenization device before presentation to the camera. This must take place immediately before image acquisition, because of the severe risk for recurrence of the same transportation-induced segregation etc. However, it is important to note that due to the inherent inhomogeneous nature of most materials, it will not be possible to obtain completely homogeneous samples even though the most

[t] E.g. acoustic chemometrics or image analysis.

thorough mixing is applied (a natural *minimum* residual heterogeneity always exists, cfr. TOS). Be this as it may, many mechanical engineering solutions can be conceived for homogenization (but not treated further here as they can be considered fully "doable", which will suffice in this conceptual analysis).

To minimize this effect a large number of *replicating increments* must be used – replicating increments can either be the result of acquiring new images along the 1-D lot ("translational replicates"), or can be effectuated by rotating the lot -, or the individual samples. The latter can come about either by physically rotating the samples, or by a circular translation of the camera and/or the illumination sources (Figure 22)

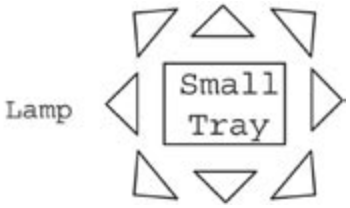


Figure 22. Creation of rotational replicated increments for improved sample representativity. The sample material can either be rotated or several illumination sources can be used.

When using low-angle structured light, changing the illumination or the viewing angle effectively acts as opening up a new view of the surface of the material ("rotational replicates"). Both kinds of replication carry their own pro's and con's. Rotational replication can be comparatively easy to set up, but it also carries its own penalty: no matter how many rotational replicates, they all but shared the same, singular analytical Y-reference value - it is effectively only the X-increments which are replicated. But this approach does open up for a very effective averaging procedure. By contrast, translational replication allow individual calibration with a fresh reference sample, but now each of these will in turn correspond to a new residual heterogeneity manifestation. Balancing these two opposing issues is a central issue when designing an optimal (i.e. appropriate and effective) IAS-system.

As a case in point (related to IA characterization of a multi-component mixture for example), acquiring only one image with a fixed illumination and camera angle (Figure 23) will often result in lack of sufficient information for the task at hand.

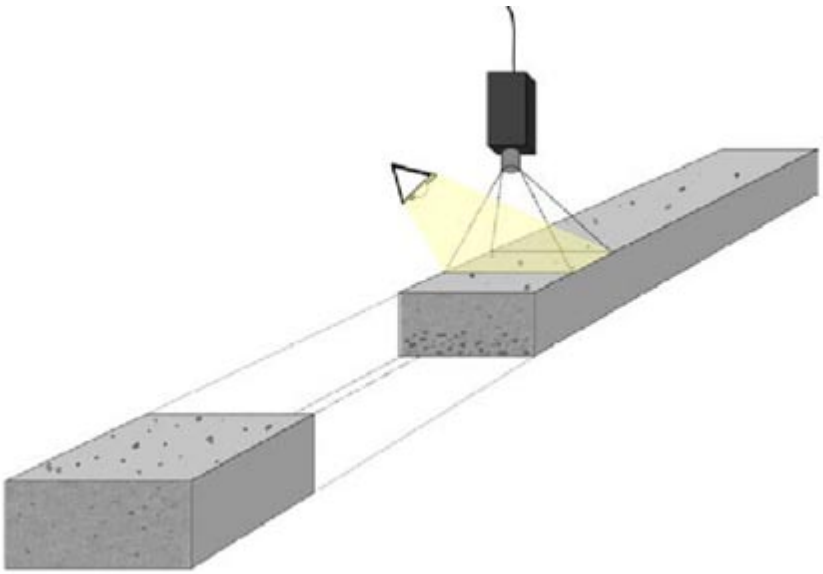


Figure 23. Principle of IAS applied to a moving process string - with possible vertical segregation or not!

It is nearly always beneficial to use several lamps – especially with regards to MAR. Using several light sources, but not at the same time, will allow presentation of new surface aspects of a sample:

- Particles that were hidden using a specific combination of illumination and camera angle (in shadows) will now appear.
- Particle shapes will be better delineated, also through their "shadow image", which can be of special importance for characterization of non-spherical particles.

Based on a scheme designed with multiple illumination sources, e.g. for a specific industrial setting, an analogue situation can be *simulated* in the laboratory, meaning that the sample can be rotated a number of times corresponding to the number of fixed illumination sources. This is the setup used in many of ACABS pilot studies (see Figure 22 and Figure 23).

Sample homogenization through mixing and/or other augmentation of the sample size/sample representativity through replicating increments are two ways to obtain improved prediction models. Nonetheless, IA prediction models can also be significantly affected by the imaging equipment and the settings of this. In fact, a wealth of parameters *can* be important to a specific, certain extent, but these are all highly problem-dependent.

4.3.1 IMAGING PARAMETERS OF IMPORTANCE IN IAS

Many image acquisition parameters are of critical importance for obtaining reasonable prediction models, because bad equipment and poor experimental settings can seriously degrade the correlation between image and material parameters of interest. Special focus should be on imaging device (camera) and illumination source/settings as these two factors are instrumental for finding useful correlations between the X and Y spaces.

CAMERA

Different camera solutions greatly affect the possibility for successful prediction models, through different options for resolution, color depth, wavelength, number of cameras etc.

The imaging equipment – camera and lens – must be selected especially for the purpose at hand. A simple TOS characterization reveals that the required sample size (camera "footprint") is phenomenological proportional to the heterogeneity of the material to be examined. With very heterogeneous material relative large image field-of-views, or many image increments, is necessary in order to obtain a sufficient representativity [9].

Selecting a camera appropriate for the tasks at hand can however be difficult. Efforts regarding this selection should be put into many different areas such as:

- Camera type (digital/analogue)
- Chip type (CMOS/CCD)
- Lens
- Filter(s)
- Camera resolution
- Image channel(s) (BW/RGB/Hyper spectral)

Most cameras used today are digital, minimizing the risk of introducing artifacts in the digitization process, which can be serious impediments to quantitative measurements. Some cameras are analogue, directly attached to a frame-grabber, which digitizes the pictures hereby functioning as an A/D-converter. This might introduce noise that should be eliminated or minimized afterwards with filtering techniques. Using even a low-end digital chip camera is as good as using a conventional analogue camera [22].

Image compression might pose an even larger problem than the digitization process. This is especially a problem for conventional digital cameras. Image compression techniques, such as JPEG, discard exact pixel information and use approximations instead, which of course might have severe effects on quantitative image analysis, since image details are lost. Higher-end consumer and professional cameras all offer image storage without lossy compression. These cameras also offer higher fidelity with the original colors showing subtle differences and finer details. This can be observed, especially in the high-brightness and high-saturation regions [22].

In some cases using an ordinary lens is not sufficient for reaching a desired field of view (or desired effective pixel size), since a 1:1 ratio maximally can be reached. It can thus be necessary to use other solutions; if the objective is inverted using a bellows solution, a (limited) magnification can be reached. If such an enlargement still is too small, a combination of a camera and a microscope can be used.

The final image resolution and color depth should in general be high. Of course, it is not always advantageous to have a really high resolution, since it directly affects computational requirements – the higher the resolution, the higher the computational load. In many cases an exceptionally high resolution is also unnecessary, since requirements for resolution depth and particle size are intimately connected, and this automatically will define the requirements for image resolution. Besides, an unnecessary high image resolution will introduce a great deal of redundant information in the images, since new features of no interest might be brought forward as a consequence of the change in scale. Such redundant information will never improve quantifications – it will only complicate prediction models and perhaps also make for computational problems, at least in on-line application scenarios.

Figure 24 shows three different cameras ranging from a simple, off-the-shelf 3-CCD color camera (medium resolution), over a high-resolution multiple-channel,, to a full multi-channel, line-scan camera.

Figure 24. Examples of different camera types: Left: Sony 3-CCD camera. Centre: QImaging b/w camera with color wheel. Right: Camera X line-scan camera (BBcom).

In many industrial inspection and monitoring situations it may be satisfactory to use an off-the-shelf 1-CCD (black/white) or CMOS camera. The problem with single chip cameras is that the spatial – and the radiometric resolutions – are closely tied together. For quantification of large particles the resolution is often above what is satisfactory, while it is rather expensive to get a camera with a satisfactory resolution for fine particles such as powders used in pharmaceuticals, especially if the need calls for a full multi-spectral facility. However, many other correlated features needs to be observed when selecting camera and illumination solutions.

ILLUMINATION

Choosing optimal settings for illumination cannot be emphasized enough, as this is what determines the optical quality of the digitized imagery. Both unilateral and bilateral illumination modes can be employed, as well as omni-directional light (Figure 25).

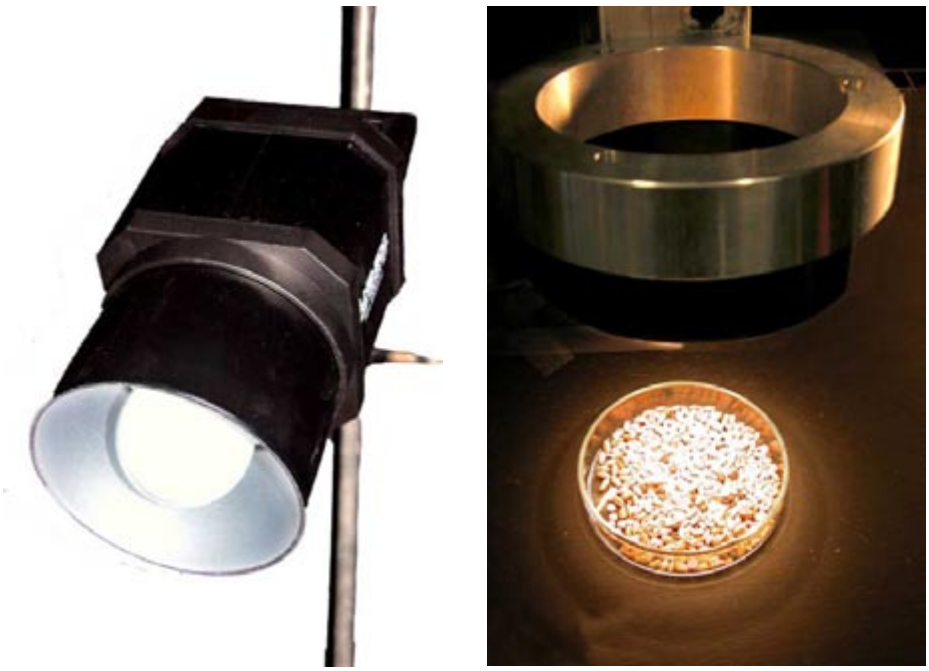


Figure 25. Left: Standard light source for both uni- and bilateral illumination (type: LHS-500). Right: Omni-directional ring-light (Schott-Fostec).

Choosing optimal light sources for image analytical techniques is of course highly dependent upon the specific application – and material – context. Material-dependent features such as reflection/absorption always have to be taken into account. Choice of illumination source and camera is also closely related; a typical lamp in the visible electromagnetic range reaches approximately 700 nm, while the near-infrared area starts around 800 nm. But these are highly–problem dependent issues, for which reason no firm general guidelines can be given. If the illumination settings are not optimal, i.e. not giving a clear contrast in the gray-scale levels for each color channel, MAR will for example not work to its full capacity. Contrasts can sometimes be enhanced by use of image enhancing and processing techniques.

OTHER OPTIMIZATION ISSUES

It is also possible to enhance the surface features (for instance color treatment) by more traditional means, such as filters or by modifying the images using low-level

image processing to improve image contrast, edge-definitions etc., but this latter always adds heavily to the real-time computational load.

Getting the right equipment and finding the optimal settings for this is unfortunately not the only necessary steps to consider for obtaining the optimal prediction models. As mentioned above segregation might occur and requires homogenization prior to imaging, but also the sample/image delimitation, the image acquisition, and the reference sample extraction has to be carefully considered.

An essential part of obtaining representative samples – be that physical samples or the virtual image footprints – is understanding and appreciation of the numerous sampling errors didactically elaborated in TOS (see chapter 4.3.2).

4.3.2 TOS SAMPLING ERRORS IN IAS

All the seven sampling errors comprising the total sampling error, TSE, are also applicable to IAS, but the cyclic and long range heterogeneities are only a concern for one-dimensional lot types and can be characterized as described in the TOS literature [6,28-30].

THE CORRECT SAMPLING ERRORS IN IAS

Even with dominant focus and efforts on optimization of the sampling procedures and the correct image acquisition and definition, sampling errors will still influence the prediction models as the correct sampling errors to some extent are unavoidable – not only in image analytical sampling, but in sampling in general.

Because we are normally not imaging only one fragment we have to introduce the grouping and segregation error, GSE. GSE can be minimized by either reduction of the fundamental sampling error, the grouping parameter, Y, or the segregation parameter, Z. Minimization of FSE is treated below, whereas the grouping parameter only can be eliminated by imaging (analyzing) smaller and smaller FOVs (in the limit actually imaging fragments one by one) but there is a natural limit to this approach. The segregation parameter can only be reduced by subjecting the lot, or the pertinent process stream segment to thorough mixing prior to sampling/imaging. GSE can on the other hand not be completely eliminated, as a residual heterogeneity – the minimum distribution heterogeneity,

MPE – will exist *regardless* of the quality of the mixing. This variance of the Minimum Practical Error (MPE) is defined as:

$$s^2(\text{MPE}) = s^2(\text{FSE}) + \text{YZ}s^2(\text{FSE}) \quad [29 \text{ p. } 77]$$

The residual distribution heterogeneity is the consequence of the omnipresent gravity acting on the fragment characteristics, which ultimately leads to an asymptotic GSE-limit always significantly above zero (obviously always material dependent).

In addition to the pervasive omnipresent distribution heterogeneity, the material also has an intrinsic constitution heterogeneity characterized by the fundamental sampling error, FSE. The fundamental sampling error variance can be well approximated by:

$$s^2(\text{FSE}) = \frac{\text{C}d^3}{M_S} \quad [29 \text{ p. } 73] :$$

with C representing a material specific constant, d the average top particle size d_{95}, and M_S the sample mass.

As can be deduced from this formula, this error can only be minimized through comminution/crushing of the lot material to a smaller average particle size, d, or by increasing the sample mass, M_S , which is much more easily done simply by acquiring more images; the sample size increase is only limited by digital image computational power!
A general reduction of the particle size will, however, change the foundation for IAS, as such reduction might also require a change in camera towards higher resolution; ultimately to distinguish between the individual fragments in the actual image analysis. As mentioned in chapter 4.3.1 a higher resolution *should* at least give identical prediction models, but in most cases result in even better predictions.

These correct sampling errors will in total never cancel out, but mixing prior to imaging, composite images[u], and ultimately reduction of the average particle size in the lot *will* effectively minimize GSE and FSE also in IAS.

It is a quite distinct other matter, that the original reason for image analysis in process control most likely is a desire for passive, remote sensing with absolutely no interference with the material stream being imaged. The possibility for such drastic intervention as having to crush the material will in all likelihood not be a frequent option.

INCORRECT SAMPLING ERRORS IN IAS

Within sampling the incorrect sampling errors can be worked on until being practically eliminated given specific conditions, but there may have to be a dedicated amount of work involved, Petersen & Esbensen [6].

In IAS applications IDE is not present if the whole cross-sectional sample/product is imaged, as the cameras field-of-view automatically defines a sample which equally represents the width of the process stream. However, *if* the whole cross-sectional sample is not imaged, the case of reduced imaging (case E in Figure 21), IDE will be present.

IEE, the extraction error, is practically impossible to avoid in IAS (as images most often only reflects the sample surface). The imaging virtually extracts only a fraction of the delimited sample area (unless the sample in the imaging area per default consists of a mono-like layer of sample material). But IEE can be reduced through reduction of the sample depth. However, IEE will always be present in imaging studies, as the imaging process divides fragments/particles on the sample boundaries, whilst fragments are considered indivisible in TOS.

The incorrect preparation error, IPE, is just as applicable to IAS as for regular TOS. For specific details see references Pitard and Gy [28,29].

The incorrect sampling errors in IAS can only be avoided if imaging the entire sample/cross section of the process stream and if the sample depth is virtually equal to zero (mono-layer sample depth). In practice IEE is unavoidable, but the magnitude of this error greatly depends on the specific material characteristics.

---

[u] Composite images: Images of several increments combined into one, larger image.

All three incorrect sampling errors, but especially IEE, will be subject to significant changes in their individual manifestation levels to a degree which is strongly correlated to the specific material and process involved. In fact, nothing much more can be said in general as to their influence – a specific analysis will always be called for, for each individual material and imaging scenario.

5 DISCUSSION AND CONCLUSION

The critical issue in IAS is centered on to which degree it is possible to acquire images which are fully representative of the bulk material under investigation. In most cases the representativity issue is tightly connected to segregation and how to guarantee sufficient homogenization of the material before image acquisition. This is just another formulation of the representativity issue: For all reference samples, which are physically extracted from the lot material, the surface cross-section (the camera footprint) must represent a sufficiently homogenized surface rendition of the 3-D bulk material.

For IA-cameras the old dictum is particularly true: WYSIWYG (What You See Is What You Get). But in the present context the issue is even worse: WTCDNSWNSUITIAPM (What The Camera Does Not See Will Not Show Up In The Image Analytical Prediction Model). This critical issue is unfortunately not always properly recognized in image analysis.

The conceptual IAS analysis above delineated the individual factors involved, both the image acquisition issues - representing the X-signals in a multivariate calibration context – and the reference sampling issues, which makes up the Y-data. It was stressed that while the latter issue has an obvious, direct relation to the TOS as a critical prerequisite, so does the X-side. An analysis of image acquisition and low-level image processing in the light of TOS was therefore carried out in detail. This resulted in specification of the critical factors and issues to be aware of when contemplating situations where chemometric multivariate calibration comes to the fore, i.e. when image analytical prediction models are to be invoked. These issues are treated in depth in the individual studies behind the collected papers making up this thesis.

5.1 APPLICATION POSSIBILITIES

The principles behind the IAS approach delineated above is applicable for on-line monitoring in many industrial sectors, among others the food and beverage-, feed -, mining -, energy -, polymer -, pharmaceutical, as well as the powder and pulp industries – indeed the solids producing/processing process industry at large.

Sector	Examples of IAS applications
Food & beverage	Process control in manufacturing of food & feeds
Polymer films	Process control of polymer films and coatings
Medical	Correlation between imagery and reference sampling (e.g. blood vs. NIR)
Powder industry	Aggregates, powders, mixtures
Paper & pulp	Quality control in lumber mills or process control in pulp and paper manufacturing. The textile industry is very similar.
Commercial sheet & surface products	Classification of steel or aluminum surface quality. Control of surface appearance of consumer products, wood boards, paper texture (correlation with material characteristics)
Pharmaceuticals	Raw material inspection, mixtures, tablets, liquids
Energy production	Fuels, wood pellets/chips, coal, oil, biofuel characterization

Figure 26. Possible PAT areas for IAS application.

5.2 FUTURE WORK

This PhD has touched many different scientific as well as technical areas. However, on an overall work-basis two of these areas stand out: Applied multivariate data analysis and Image Analytical Sampling. This being so because the applied multivariate data analytical technique Multivariate AMT Regression has been used to show the potential of both the specific methodology and the applicability of Image Analytical Sampling.

With two *so* different techniques this calls for two different evaluations of future working areas:

Multivariate AMT Regression
The applicability and theoretical study/optimization of the algorithms behind AMT/MAR has undergone thorough research and in my mind MAR is at a developmental stage, which makes the technique ready for industrial application. Slow computers have earlier stopped this technique from being implemented industrially, but with the ever faster computers this problem is getting smaller and smaller.

Most technical/imaging parameters regarding AMT/MAR should always be determined based on the specific material investigated and on the application context. On the other hand general guidelines/rules of thumb might be established, however, this would require a full-time PhD scholarship looking at correlations between most experimental parameters and different materials/substances. Indeed interesting, but also very time consuming and my best guess is that in the end it would be very difficult to put up guidelines in general.

As all image analytical applications should take structurally correct sampling into account prior to and during implementation, Image Analytical Sampling should always come into consideration.

Image Analytical Sampling
This area is much needed for the imaging industry and the first task at hand should be to deduce and write a complete tutorial on both the theory and the practical aspects behind Image Analytical Sampling. A significant part of such a tutorial can be found in this the introduction and in the four image analytical applications described in Papers I to IV.

REFERENCES

1. Esbensen K.H., *Multivariate Data Analysis - In Practice*, (5 edn) CAMO ASA, 2002.

2. Höskuldsson A., *Prediction Methods in Science and Technology*, Thor Publishing, 1996.

3. Martens H. and Næs T., *Multivariate Calibration*, John Wiley and Sons, 1989.

4. Jöreskog K. G. and Wold H., *Systems under indirect observation: causality, structure, prediction*, North-Holland, 1982 pp. 263–270.

5. Huang J., *Developments in Applied Chemometrics - AMT, Acoustic Chemometrics and N-way Analysis*, PhD thesis, HIT 2001.

6. Petersen L., Minkkinen P., and Esbensen K.H., *Representative Sampling for Reliable Data Analysis: Theory of Sampling*, Chemometrics and Intelligent Laboratory Systems, 77 (2005) pp. 261-277.

7. Alsberg B.K. and Remseth B.G., *Multivariate Analysis of Surfaces: Folding and Interpretation,* Journal of Chemometrics, 6 (1992) pp. 135-160.

8. Esbensen K.H., Hjelmen K.H., and Kvaal K., *The AMT Approach in Chemometrics – First Forays*, Journal of Chemometrics, 10 (1996) pp. 569-590.

9. Mortensen P.P. and Esbensen K.H., *Optimization of the Angle Measure Technique for Image Analytical Sampling of Particulate Matter*, Chemometrics and Intelligent Laboratory Systems, 75 (2005) pp. 219-229.

10. Andrle R., *The Angle Measure Technique: a New Method for Characterizing the Complexity of Geomorphic Lines,* Mathematical Geology, 26 (1994) pp. 83-97.

11. Dahl C.K. and Esbensen K.H., *Image Analytical Determination of Particle Size Distribution of Natural and Bulk Aggregates*, Chemometrics and Intelligent Laboratory Systems. Accepted for publication May 2007.

12. Dahl C.K, Petersen L., and Esbensen K.H., *Image Analytical Characterization of Powder Agglomeration by AMT-regression (Angle Measure Technique). Part 1: Development and Method Validation*, In prep.

13. Dahl C.K, Minkkinen P., and Esbensen K.H., *Image Analytical Monitoring of Paper Quality – a Feasibility Study*, TAPPI JOURNAL November 2006.

14. Huang J. and Esbensen K.H., *Applications of AMT (Angle Measure Technique) in Image Analysis Part 1: A New Methodology for In-situ Powder Characterization*, Chemometrics and Intelligent Laboratory Systems, 54 (2000) pp. 1-19.

15. Huang J. and Esbensen K.H., *Applications of Angle Measure Technique (AMT) in Image Analysis – Part II: Prediction of Powder Functional Properties and Mixing Components Using Multivariate AMT Regression (MAR)*, Chemometrics and Intelligent Laboratory Systems, 57 (2001) pp. 37-56.

16. Huang J., Wium H., Qvist K.B., and Esbensen K.H., *Multi-way Methods in Image Analysis - Relationships and Applications*, Chemometrics and Intelligent Laboratory Systems, 66 (2003) pp. 141-158.

17. Huang J., Ose S., de Silva S., and Esbensen K.H., *Non-invasive Monitoring of Powder Breakage During Pneumatic Transportation Using Acoustic Chemometrics*, Journal of Powder Technology, 129 (2003) pp. 130-138.

18. Zuech N. and Miller R. K., *Machine Vision*, Kluwer Academic, 1989.

19. Jain R., Kasturi R., and Schunck B.G., *Machine Vision*, McGraw-Hill, 1995.

20. Zamperoni P., *Wasserzeichenextraktion aus Digitalisierten Bildern mit Methoden der Digitalen Bildsignalverarbeitung*, Chemische Technologie der Cellulose, 43 (1989) pp.133-143.

21. The Quantitative Imaging Group, Delft University, *Shading Correction.* http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip-Shading-2.html, 27 Aug. 2004.

22. Russ J.C., *The Image Processing Handbook*, (4 edn) CRC Press, 2002.

23. Gonzalez R.C, Woods R.E., and Eddins S.L., *Digital Image Processing using Matlab*, Pearson Prentice Hall, 2004.

24. Soille P., *Morphological Image Analysis*, (2 edn) Springer, 2003.

25. de Jong S.M. and van der Meer F.D, *Remote Sensing Image Analysis: Including the Spatial Domain*, Springer, 2004.

26. P. Geladi and H. Grahn, Multivariate Image Analysis, John Wiley and Sons Ltd. 1996.

27. Esbensen K.H. and Geladi P., *Strategy of Multivariate Image Analysis*, Chemometrics and Intelligent Laboratory Systems, 7 (1989) pp. 67-86.

28. Pitard F.F., *Pierre Gy's Sampling Theory and Sampling Practice*, (2 edn) CRC Press Ltd., 1993.

29. Gy P., *Sampling for Analytical Purposes*, John Wiley and Sons Ltd., 1996.

30. Smith P.L., *A Primer for Sampling Solids, Liquids, and Gases*, Siam, 2001.

31. Skoog D.A., Holler J.F., and West D.M., *Fundamentals of Analytical Chemistry*, (7 edn) Harcourt Brace College Publishers, 2005.

32. Kenkel J., *Analytical Chemistry for Technicians*, (3 edn) CRC Press, 2002

33. Christian G.D., *Analytical Chemistry*, (6 edn) John Wiley and Sons Ltd., 2003

34. Patnaik P., *Dean's Analytical Chemistry Handbook*, (2 edn) McGraw-Hill, 2004.

35. Lied T.T., Geladi P., and Esbensen K.H., *Multivariate Image Regression (MIR): Implementation of Image PLSR – First Forays*, Journal of Chemometrics, 14 (2000) pp. 585-598.

36. Lied T.T. and Esbensen K.H., *Principles of MIR, Multivariate Image Regression - I: Regression Typology and Representative Application Studies*, Chemometrics and Intelligent Laboratory Systems, 58 (2001) pp. 213-226.

37. Petersen L. and Esbensen K.H., *Representative process sampling for reliable data analysis - a tutorial*, Journal of Chemometrics, 19 (2006) pp. 625-647.

Paper I

Casper K. Dahl*, Lars Petersen and Kim H. Esbensen

"Image Analytical Characterization of Powder Agglomeration by AMT-regression (Angle Measure Technique) – Development and Method Validation"

In Preperation

Image Analytical Characterization of Powder Agglomeration by AMT-regression (Angle Measure Technique) Development and method validation

Casper K. Dahl *, Lars Petersen and Kim H. Esbensen

Applied Chemometrics, Analytical Chemistry and Sampling Research Group (ACACSRG), Aalborg University Esbjerg, Denmark

This paper is dedicated to the memory of

Sunil de Silva
Professor, research director, Ph.D.
Telemark University College, Department of Process Technology &
POSTEC (Powder Science and Technology) Research Group at
Telemark Technological R&D Centre (Tel-Tek),
Porsgrunn Norway

who abruptly passed away April 29th 2002, at the height of his academic career –
we shall thoroughly miss his scholarly expertise, friendship and leadership

ABSTRACT:

Fifteen different powders have been subject to controlled water-induced agglomeration and subsequent quantitative image analysis and multivariate prediction modeling. The degree of induced agglomeration varies from dry powder (minimum agglomeration) through water-saturated powders, quantified by a powder-invariant agglomeration index. The powders range from large-sized particulates of many different provenances to small-size pharmaceutical powders such as talc, including seven POSTEC-selected representative industrial powders (cement, dolomite, alumina, talc, cellulose, clay, and sand). The technique uses the Angle Measure Technique (AMT) transformation on digital images obtained by a modified R/G/B camera. The channel with the highest contrast (green) was used throughout the experimental series, characterized by unilateral low-angle illumination. AMT successfully captures the essential texture characteristics of the imaged powder surfaces, which can be quantitatively correlated to the innate moisture content. Most of the resulting models show very good correlation between predicted and reference agglomeration index values, indicating promising aspects for on-line control/characterization (prediction) of moisture and the degree of agglomeration of powders in various industrial processes that can be subjected to image analytical process monitoring.

Keywords: Angle Measure Technique, AMT, Powder Agglomeration, Image Analysis, Multivariate Data Analysis, Prediction, Moisture Content.

1 INTRODUCTION

Powder-handling industries place great emphasis on extensive quality control. One aspect of powder quality control, which has not yet received proper considerations, concerns agglomeration. Quantification of, and changes in, the superficial appearance of powders is of great practical interest in several particulate matter industries.

Powders and aggregate mixtures are frequently subject to agglomeration induced by moisture, condensing water, electrostatic forces amongst other. This is an often occurring, serious problem in several industrial sectors and can be a costly everyday problem

for many process industries and powder-based goods manufacturers. It would be highly desirable if a quantitative agglomeration index (and thus a potential new process operation parameter) could be derived, preferably on-line and if possible by a non-intrusive technique (thus solving a difficult sampling and sample handling problem). This could also furnish a means for quantification of the innate moisture content of raw material by indirect calibration based on agglomeration behavior.

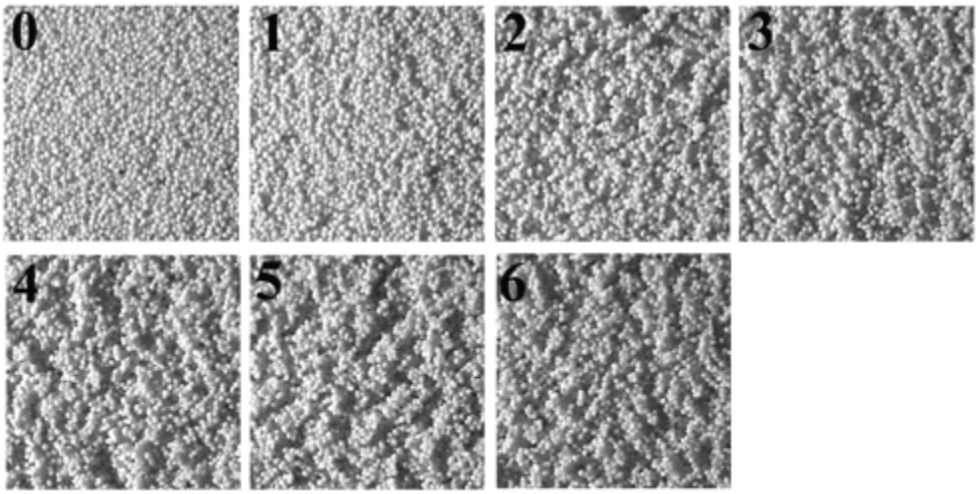


Figure 1. Generic illustration of powder agglomeration. Progressively increasing moisture additions to dry powder (0). 1–6 represent equal increments of added moisture. This work develops a facility for quantitative characterization of either agglomeration or moisture content (see text for details).

1.1 Objective of Study

The objective of this work is to develop – and validate – a non-intrusive, fast, and inexpensive technique for quantifying the degree of agglomeration (Fig. 1) for a comprehensive range of industry-relevant powders. The present work will solely address water-induced agglomeration because this is by far the most prevalent in practical industrial settings (condensation, transport, and/or storage under varying moisture conditions). While electrostatic agglomeration is also rather common, its effects are typically several orders of magnitude less than those of water-induced agglomeration and are therefore excluded here [1-4].

1.2 Development of Pilot-study Laboratory Procedures

For non-intrusive powder characterization, image analysis is an obvious candidate. Digital images can be obtained with relative ease in many industrial settings and is thus a straight-forward primary choice for data acquisition. Industrial and technical image acquisition and digital image processing is today a well developed area with many existing applications in different industrial and laboratory environments; a comprehensive recent overview is that of Zeuch [1]. Because of this we here assert any-and-all practical problems related to specific implementations of an appropriate image analytical data acquisition system as mere engineering problems – all of which are *doable* in the relevant implementation context, and we shall not treat these aspects further below.

For the present study we will not be using standard, conventional image processing procedures for analyzing the powder images obtained. Instead, we shall introduce a new texture-characterization procedure, the AMT (Angle Measure Technique), for decoding the specific agglomeration image data structures in combination with a chemometric regression technique termed PLS-R (Partial Least Squares Regression), both of which shall be briefly described below.

Angle Measure Technique (AMT)

AMT is a relatively new preprocessing technique, which has been found useful in the last 7+ years of initial industrial applications [5-14]. The technique has especially seen powerful image-analytical applications.

AMT is a domain-transforming technique, which in the present setting is used for extracting textural features from images and converting these into spectra characterizing the surface rendered in the image. These spectra can subsequently be used in multivariate calibration and prediction, such as PLS-regression (see further below). For a 2D digital image the local surface is transformed into a 1D vector ("AMT-spectrum") carrying information about textural complexity at all existing scales within the image.

AMT works on grey-level (single-channel) images. In this work the green channel is selected throughout (because of superior contrast) and *unfolded* into a row vector, by juxtaposing the individual rows of the image, which in turn is subjected to the AMT transform. The solid curve in Fig. 2 ("hill-and-valley") represents the grey-level values of each pixel making up this concatenated vector. An image measuring for instance 512 x 512 pixels results in an *unfolded* linear array (the X-axis direction in Fig. 2) of 262.214 elements. The Y-axis in Fig. 2 serves only plotting and visualization purposes allowing the pixel grey-level values to be related in magnitude; often a "connecting line" is depicted between all discrete data. The linear array is often called the "measurement series" allowing a uniform description of the AMT method for both 1D data series as well as unfolded 2D arrays etc.

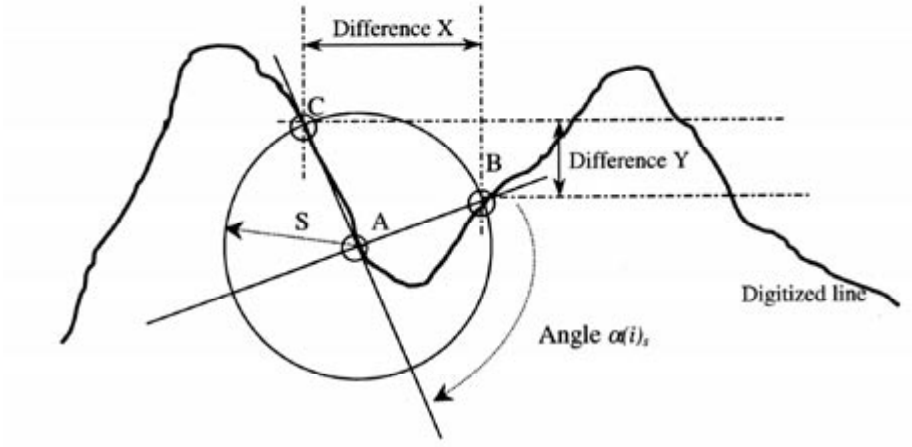


Figure 2. Derivation of the Mean Angle (MA) complexity measure ($\alpha(i)_S$) in AMT, see text. The *complementary* MDY measure is phenomenologically self-explainable; also explained in references [5-7].

The AMT method operates by first choosing a random point from the entire measurement series "A" as the centre of a circle with diameter "S", which is termed the temporary *scale*. Points "B" and "C" are found as the intersections of the circle with the measurement series (i.e. intersection points on the curve made up of the connecting lines between all grey-level values). The *complement* to angle CAB, $\alpha_{(i)}$, is calculated and stored. Differences X and Y are found as the horizontal and vertical distances between points C and B, respectively, see Fig. 2. The selection of point "A" is actually repeated a sufficiently high number of times, 500 or more, and the results are used for calculation of the *Mean Angle* MA (mean of the complementary angles, $\alpha_{(i)}$); a "sufficiently high number" relates to statistical stability for the MA estimate. Following this procedure, the mean differences MDY and MDX are also calculated; very often only MDY is used (MDX is only used for "closed curves"). The random distribution of 500+ "A" points covering the entire data series is a necessary and sufficient condition for the [MA, MDY] complexity measures to represent the complexity at scale S.

Subsequently the value of S is increased by one (S = S+1) and the [MA,MDY] procedure is repeated in full until "S" reaches a (problem dependent) maximum value (often a number equal or smaller than half the horizontal resolution of the image). With the present 512 x 512 images, this results in [MA,MDY]-spectra with 256 data-points corresponding to a half-width of the image. This is very often more than enough to characterize the pertinent textures, because of a problem-dependent choice of the image resolution of the powder aggregate field-of-view, see Fig.1.

Below is presented an illustration of a heavily agglomerated powder image and the corresponding AMT-spectra (Fig. 3). The image is cropped from the central area to avoid influence from the black background surrounding the Petri dish, which is used to contain the powder.

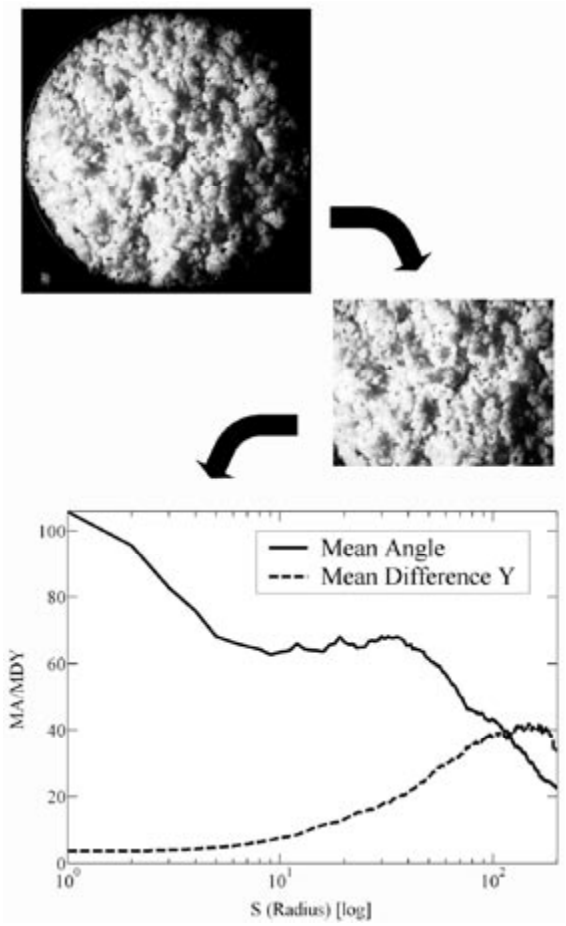


Figure 3. An image of severely agglomerated ground corn is *cropped* and subsequently subject to the AMT-transform (via the AMT-toolbox for MATLAB [9]).

The two spectra, MA and MDY are usually joined into a single row vector consisting of 2 times "S" elements; each element records the appropriate MA and MDY values, respectively. An automated MATLAB routine has been developed for cropping images, averaging data, indexing, transforming and exporting results as a data matrix file, which is compatible with most chemometric (and mathematical/spreadsheet) software packages for post-AMT data-analytical processing.

Fig. 4 provides a schematic overview of the entire process from imaging to the final regression modeling of the AMT-complexity spectra calibrated with regard to a Y-variable of interest. Y is the *independent* data to be modeled on X, which is the dependent (spectral) data. Fig. 4 is a precognition of the specific regression approach used below, PLS-regression, in which T is a matrix of *scores* and W is a matrix of *loading weights*. All these terms are explained in full in the section on Partial Least Squares regression, PLS-R.

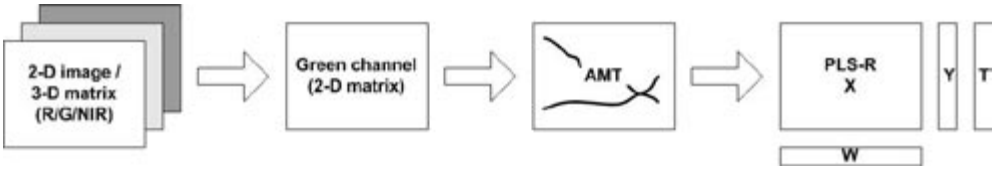


Figure 4. Schematic first overview of the compound AMT/PLS-R transform. The present work made use of a R/G/NIR-camera [6-8,12]. The "best contrast" for all powders turned out to be the green channel.

Why Does AMT Transformation Work for Agglomeration Characterization?
AMT was originally meant for characterizing the complexity of "geomorphologic lines" (e.g. map contours, coastlines, meandering rivers - it was in fact developed by physical geographer Robert Andrle as an alternative to fractal analysis), but has also been shown to describe textures in images very well [5-14].

The success in describing textures in powders, food types amongst other products, is related to using a problem-specific *low-angle illumination*, see e.g. Figs. 1 & 4, fully explained in the published literature, *ibid*. This creates intricate grading shadow variations on the surface imaged by the camera, which records this as complicated light-and-dark patterns because of the interaction of the low-angle illumination with all surface protrusions and hollows. An increasing degree of agglomeration for example results in progressively larger agglomerates and thus more complicated light-dark regions in the image, cfr. Figs. 8 & 9.

The image recording - explained below - results in complex images, which cannot be interpreted directly. Instead, they are used as unfolded "linear texture transform images" as described above (1D vectors or complexity spectra), where dark regions will result in radiometric low grey-level values (in the ordinary R/G/B image each pixel value vary between 1 and 255) while light regions will be represented by high Y-values, Figs. 2 & 3.

There are several reasons for the need of the AMT-transform in the present context. It bears noting that AMT is a much more powerful transformation than for example a simple threshold operator. The light-and-shadow patterns on the sample (powder) surface represent a snapshot of the entire complex surface morphology. The AMT-spectra are able to reveal both the local as well as the scale-dependent complexity upwards from a scale corresponding to the immediate neighboring pixels to half the image size - *simultaneously*. The brightness of each individual pixel represent a very complex interaction between the specific color of the underlying particle (often several pixels in diameter) and reflection of the incident light, due to the geometrical angle of the particle face at the location of the pixel and the unilateral illumination angle as well. The brightness-level, the Y-level of the unfolded linear arrays, is very much a compound measure of intrinsic color and the reflection efficiency, which in turn is a function of the spatial disposition of all particles in question. These relationships are further compounded at the mesoscopic scale, because the surface of the sample is manifestly not level in the pristine state of the powders as presented to the camera (although in half the present experiments this surface was deliberately *scraped*, se further below). These features cannot in any way be decoded by a simple Y-level threshold in neither the raw R/G/B color space, neither in for example L,a,b or the HIS alternative color transformations, nor can they be related to the apparent size-distributions of the dark areas for example.

This latter is because of the dominantly grading relationships between both dark and bright areas – at *all* scales. Thresholding, for example, would amount to defining all "dark areas" by a simple Y cut-off level, which would be but an exceedingly poor simplification of a much more complicated relationship. In fact, the low-angle illuminated imagery of the surface morphology demands use of the entire fidelity in the grey-level intervals recorded [0,255]. By working out the pertinent scale-versus-angle (scale-versus-complexity) relationships for all scales *simultaneously*, AMT is able to render a complete complexity spectrum as a function of *scale*, i.e. able to use all the differential latent texture information present in this type of image. It is precisely this function of scale characteristic, which sets AMT aside as a monumentally more powerful transform than mere fractals [5,14].

From extensive AMT experience the powder-and-agglomeration-state-dependent patterns of dark and light regions have been shown to be qualitatively, indeed quantitatively, related to the degree of agglomeration. This can only come about by the use of a multivariate regression approach such as PLS-R, however.

Partial Least Squares-Regression (PLS-R)

We here only describe the phenomenology of the PLS-R method in brief. Recently another chemometric paper making extensive use of PLS-R was published in this journal [15], in which PLS is explained in somewhat more detail. Extensive descriptions of the theory and practice of PLS-regression (PLS-R) can be found in the chemometric literatures [16,17] and references herein.

PLS-R, Partial Least Squares Regression, is a generalized regression approach, where indirect observations obtained from e.g. instruments (X-data) are related to the interesting properties, the dependent Y-data, in a new augmented fashion. Often the dependent data are laborious, inaccessible or expensive to acquire (analytical chemical data, powder characteristics obtained *in the lab*. etc.), while the indirect observations are much more easily obtainable as measurement/instrumental data (hence fast and inexpensive). In the present context image, and image-derived data, are clearly in the latter, X, category, while the Y-data would be the agglomeration (or experimental moisture addition) data.

<u>The defining PLS-issue</u>: PLS uses the Y-data structure *directly* as a guide for decomposing the X-matrix, thereby ensuring that *only* the interesting, correlated X-variation is put into the regression model, while all the uninteresting non-correlated "X-noise" is filtered out (Y-noise is similarly screened away in this approach). This ensures a statistically optimal future prediction performance of Y, *ibid*.

This regression method is based on representing the original, multi-dimensional data-set (first delineated as a data "point swarm" in the coordinate system made of the p X-variable axes, Fig. 5), with fewer descriptive parameters (latent factors or principal components, PC's) that capture the main variation in the data. Finding these PC-directions is done either by searching for maximum variance directions in a series of orthogonal sub-spaces, or by a least squares minimization approach on the *residuals* (defined below). The maximum variation directions are simultaneously optimized with regard to correlation with the Y-data. Thus the directions depicted as PC1 and PC2 in Fig. 5, are in fact found by this compound optimization criterion; the directions are thus,

more-or-less, tilted with regard to the pure X-variance directions alone (corresponding to the Principal Component decomposition). Higher-order components, PC3+ are derived as the case may be, dependent on the data complexity present.

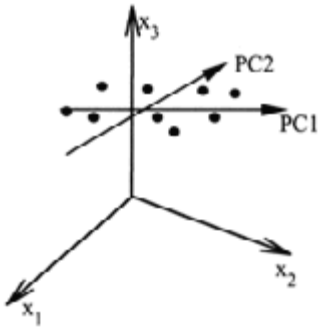


Figure 5. The original data set represented in the p-dimensional coordinate system (made up of p X-variable axes). The PLS-representation makes use of a *projection sub-space* defined by a series of components, here illustrated by a new 2-dimensional PC-space [10,11]. In reality, there is always a very high degree of *dimensionality reduction* with PCA and PLS, for example from 512 X-variables (e.g. spectroscopic data, AMT spectra) to a very few components only, typically 1-5 or so.

The perpendicular *projection distance* from a variable (a point) to a PC is termed the *residual*, e_i. Every PC can be described as a linear combination of the p X-variables. Every PC has a set of coefficients called *loadings*, p_{ka}, which characterizes this direction with regard to the original p X-variable axes, Fig.5 (*k* is a variable index; *a* is a ordering component index for the PC's).

All original [X,Y] data can hence be decomposed as following:

$$X = \sum_A T \cdot P^T + E$$
$$Y = \sum_A U \cdot Q^T + F$$

Where
T is a matrix of *scores* for the X-data; P^T contains the corresponding *loadings*.
U is a matrix of *scores* for the Y-data; Q^T contains the corresponding *loadings*.

The *scores* are found by projecting each point onto a PC and subsequently measuring the distance from that point to the center of the model. All PLS-calibrations in fact results in two set of X-loadings, loadings, P, and loadings weights, W, respectively. The P-loadings express the internal relationship between the X-data and their scores, whilst the loading weights represent the essential guiding of the regression model towards maximum correlation with the Y-data; these latter carry the essential X-Y regression information in the overall PLS regression model.

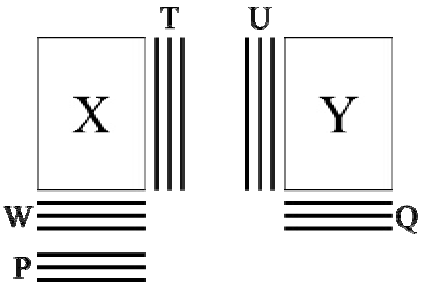


Figure 6. Summary of the matrixes involved in PLS [9,16,17].

Visual inspection of the W-vectors forms a central element of establishing a PLS model. These W-spectra, as they are called, allows direct interpretation of which X-variables contribute most to the model. In a sense, once a model has been established – and suitably validated – the w-relationships contain all the essential model information which can then be directly related to the original AMT-spectral information amongst other. We shall use, and explain, this feature extensively in the data analysis and results sections below

2 POWDER SELECTION CRITERIA

Fifteen carefully selected powders and particulate materials are investigated for the possibility of developing a quantitative model to describe the degree of agglomeration.

The powders are shown in fig. 7:

1. Ground Corn
2. Couscous
3. Paprika
4. Amaranth
5. Red Lenses
6. Brown Rice
7. Clayey Sand
8. Plastic pellets
9. Clay
10. Fine Sand
11. Microcrystalline Cellulose
12. Cement
13. Alumina
14. Talc
15. Micro dolomite (MD100)

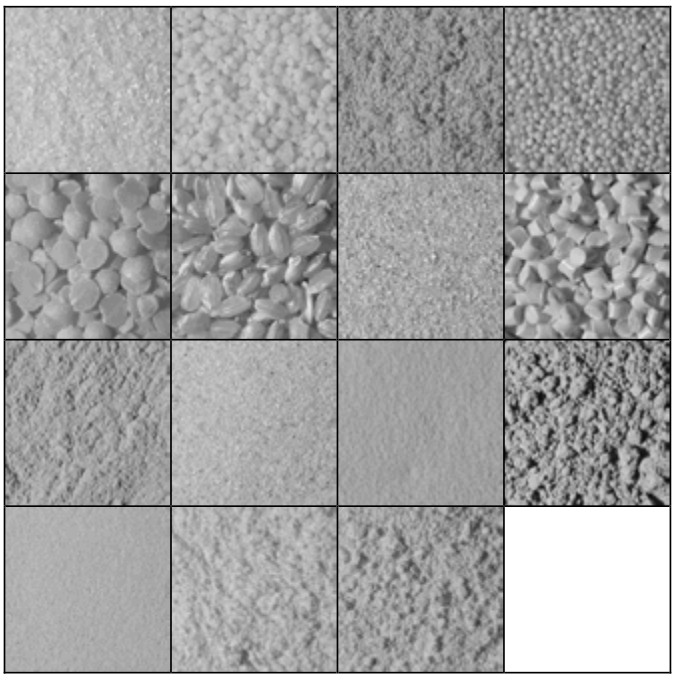


Figure 7. Photograph illustration of all 15 powders. All photographs are to the same scale. The numbers correspond to the images from top left through right bottom. These grey-level illustrations pertain to the *green* channel used throughout this study.

We have put great emphasis on selecting relevant and typical powders, representative both with regard to their use and abundance in the agglomeration-concerned industry as well as from a fundamental powder science and technology point of view, in order – hopefully – to be able to make valid interpretations of the underlying principles at work. Thus the first seven powders were chosen from easily accessible "everyday" powders, but with focus on the widest possible span regarding the basic surface morphological characteristics, as well as their wettability characteristics. Prior to the selection of the final set of powders in the present study, many others were also tested for their ability to agglomerate, by screening tests employing spray-added water (see further below). These first seven powders span a wide range in what could be called the powder agglomeration space.

The remaining eight powders were selected from a repertoire of major, proven industry-related powder types, as delineated by a standard set of powders from the POSTEC research group, Porsgrunn, Norway.

Thus we have good reason to consider the present set of powders as spanning the gamut of the relevant characteristics for the purpose at hand. While we, from a practical point of view, are interested in finding out to what degree it is possible to establish image analytical methods for quantitative agglomeration characterization, we are equally interested in the fundamental powder science and technology aspects of why the established regression models work (planned for an companion paper II), by focusing on more detailed interpretations of the PLS-regression parameters and their relationship to the underlying fundamental AMT and image analytical powder features.

In the present study the one overriding design parameter is the degree to which the selected powders show a relatively easy propensity towards agglomeration (for obvious reasons of industrial relevance). We *assume* that individual particle surface morphology and wettability are among the most influencing physical features involved, because of their direct role as determinants as to the aggregate system behaviors. We have had very little control over possible additional chemical and/or physiological features leading to different water adsorptions/absorptions by the powder particles themselves. Finally we have discarded electrostatic aspects as was delineated above.

3 EXPERIMENTAL

For the first seven powders we performed several sets of pilot experiments in order to orient ourselves with regard to the somewhat elusive agglomeration relationships.

The "maximum amount" of water - until the powder was saturated - was measured for every powder type individually by adding controlled incremental amounts of water, under continuous stirring, until the water began adhering to the sides of the stirring beaker and the sample seemed not to agglomerate further. This amount was equated as: "100% agglomerated, saturated", see Fig. 8. Any additional water resulted only in producing a slurry that became progressively more diluted. Water was added by a modified spray apparatus giving uniform increments of water. The individual maximum amounts of water that could be added to each powder type, were divided into 10% increments (plus the initial dry sample), totaling 11 agglomeration levels for each powder. Relevant absolute increment sizes and total amounts are compared in Table 1.

Rather than concentrating on these maximum absolute amounts of water, we have derived a *relative* water-induced agglomeration *index*, which shall be used in the quantitative modeling studies below. We state that this powder agglomeration quantification is the most relevant measure keeping in mind that we are both interested in the *individual* agglomeration quantification and modeling possibilities for each powder, as well as the underlying fundamental powder science relationships which are related to the relative relationships *between* the powders.

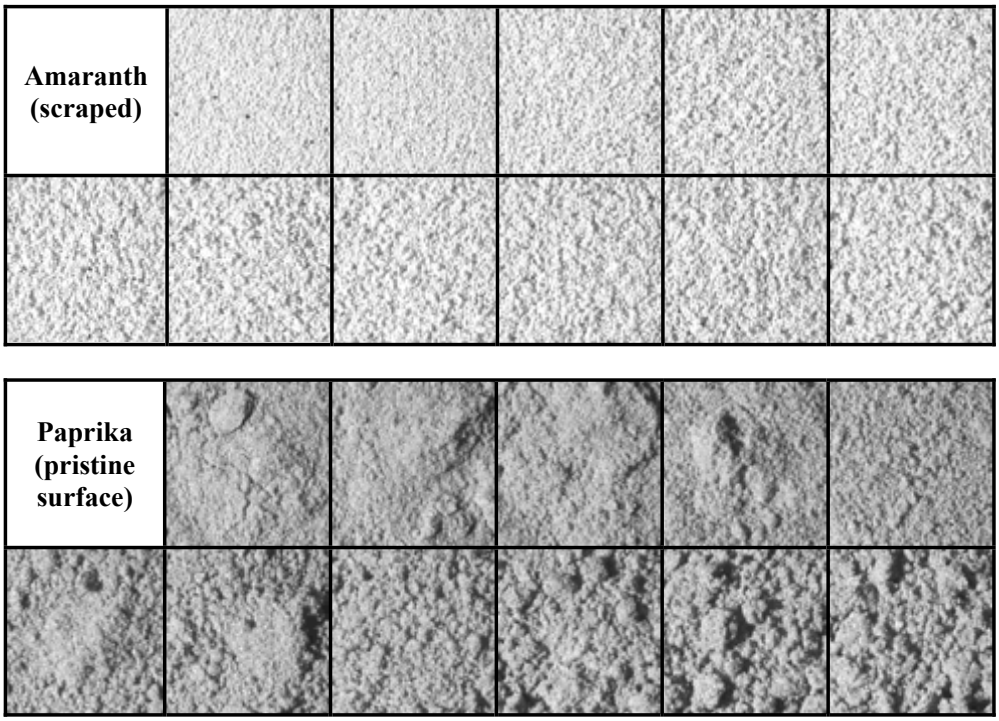


Figure 8. Progressive agglomeration (0 % through 100 %; increments of 10 %) for Amaranth (scraped surface) and Paprika (pristine, non manipulated surface). These illustrations represent two of the 15 agglomerated powders used in the quantitative modeling below. The alternative surface treatments (scraped, pristine) are explained in text immediately below.

Table 1.
Pilot study powders: Absolute amounts of water added.

Component	Max amount of water [No. of increments]	Increment size
Ground Corn	60	6
Couscous	60	6
Paprika	40	4
Amaranth	20	2
Red lenses	10	1
Short brown rice	10	1
Clayey Sand	10	1
Pellets	10	1
Clay	30	3
Sand (0,18 mm)	10	1
Cellulose	30	3
Cement	20	2
Alumina	30	3
Talc	30	3
MD100	20	2

The individual samples were prepared by adding 70 ml of the "dry" (as purchased, no specific drying were performed) powders to a glass beaker with a standardized stirring approach (a laboratory fork) between additions of each increment. After each addition of water and subsequent stirring, the samples were transferred to a Petri dish for photographic recording in one of two different ways producing either a *scraped* or a *pristine*, non-manipulated powder surface.

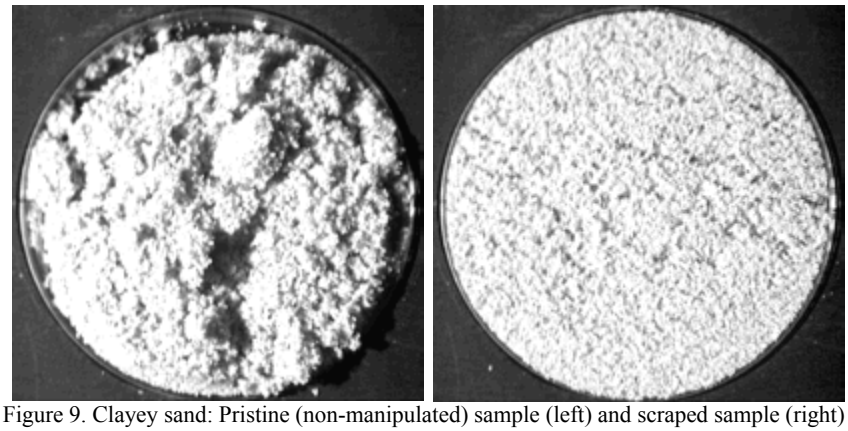


Figure 9. Clayey sand: Pristine (non-manipulated) sample (left) and scraped sample (right)

Rationale: These alternative preparation techniques are designed to *simulate* the two most obvious real-life industrial imaging pre-processes imaginable, namely a conveyor belt transportation past a camera - with or without - a mounted surface-scraping instrument. "Scraping" in this study means that the sample is transferred to a Petri dish in bulk, after which the aggregate surface is scraped by a sharp object to get a broadly planar, even surface (using a straight knife edge in this instance), Fig. 9. Conversely, pristine, or "non-manipulated", means only gently shaking the contents of the dish-

transferals into place (in an as "reproducible" way as possible), to get a surface that is not tampered with in any mechanical fashion. Fig. 9 shows the resulting surface manifestations to good measure. From our initial powder image analytical knowledge, [5-9], it is not known a priori which surface preparation type will lead to optimal image analytical fidelity; hence we here treat both alternatives throughout the modeling below.

The time from addition of the first water increment till the first image was recorded was exactly four minutes. This interval is chosen in order to make any untoward illumination-induced evaporation as reproducible as possible for all powder samples. The Klieg light in Fig. 10 gives off 1000 Watts; thus minor, but possibly non-negligible, water evaporation cannot be ruled out.

The digital imaging camera setup is illustrated in Fig. 10, in which also the low-angle unilateral illumination employed can be observed. This imaging procedure followed closely precursory powder studies by [5-9], in which the advantages of the combined low-angle illumination/AMT texture-characterization is laid out in full detail, although we here performed a completely new illumination angle optimization pertaining to the present set of powders - to be described below.

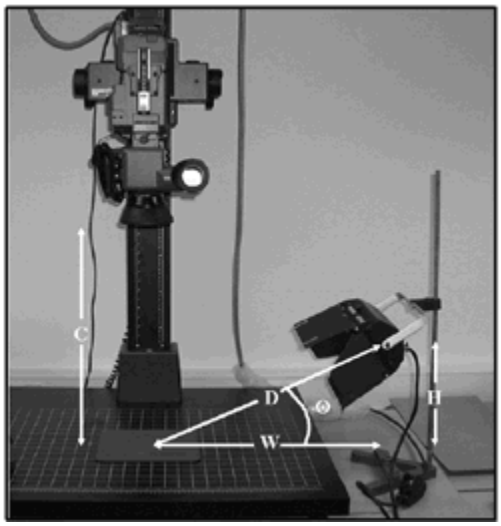


Figure 10. The camera-setup used for image recording. An illumination angle (θ) optimization was performed for an initial selected sub-set of powders (see text for details).

The distance from the light source to the sample "D" is held constant at all angles by varying the distances "H" and "W", respectively, hereby insuring equal light intensity. The distance from the camera to the sample "C" is held constant resulting in an average image resolution for all powders (fixed from a survey pilot experiment).

Replication regimen: Four replicate images were recorded for each powder sample (the Petri dish was rotated 90 degrees between each acquisition – spaced by 15 seconds). A pre-chosen central quadratic area of all Petri dish images were then cropped and subsequently transformed by AMT. The resulting AMT-spectra were imported into "The Unscrambler", a multivariate data analysis software package, and appropriate averages of

these four replicate AMT-spectra were calculated and stored (together with all individual recordings). In the present study, as indeed in all precursor studies, we make use of the significant signal-to-noise improvement available when using averages over four replicates. This type of replication is absolutely necessary when imaging surfaces as irregular as those found in the present context; this has also been amply confirmed by all our earlier powder imaging experiences.

Subsequently we carried out individual PLS-regressions on all 15 series (each consisting of eleven averaged objects) for both the scraped and pristine samples. All PLS-models are here validated by full cross-validation, i.e. leaving one sample out sequentially, while modeling the remaining ten samples and validating by predicting for the left-out object – repeating the procedure circulating out/replacement for all 11 objects for each powder type.

For the present model assessment purpose, we may refrain from the potentially most objective test-set validation (i.e. using a complete new, 11-object validation data set), often considered *mandatory* for valid chemometric regression modeling, because of the limited *comparison objectives* of the model evaluations and interpretations [16].

3.1 Optimal Illumination Angle for Image Acquisition

To optimize image acquisition for the modeling of the degree of agglomeration, four different illumination angles were initially tested. This illumination optimization constitutes one of the most important features related to the (subsequent) practical industry implementation engineering problems otherwise not treated here. This type of optimization is absolutely necessary when imaging surfaces as irregular as those found in the present context.

Based on our earlier powder imaging experiences, the angles tested were as follows:

Table 2.
Alternative illumination angles.

Angle number	Angle [degrees]
1	23
2	35
3	48
4	63

The alternative illumination angles were not tested for all fifteen powders however – rather this was pilot-investigated for a selected subset of only three representative powder types: amaranth, couscous and paprika, selected because of the extreme visual appearance differences. It turned out that the angle 48 degrees was very clearly optimal for this context, as is borne out by the results reported in Table 3, in which 5 out of 6 of the best models relate to this angle. Thus this angle is used in all succeeding experiments.

Caveat: It should be noted that the present objective of relating a widely varying set of different powders constitutes a different and demanding situation from that of implementing an automated image acquisition system for use in a fixed industrial environ-

ment. Such an environment is typically only concerned with a single powder type, which of course calls for a much more constant illumination context.

Evaluation of the regression prediction performances for the different models in Table 3 follows exactly the same procedure (explained in detail below). For now it suffices to note that for all practical purposes the prediction results are almost exclusively related to the third column representing an angle of 48 degrees which is chosen by an overall majority vote. It should be noted, though, that the angle of illumination should be optimized for every powder in its specific implementation and application context.

Table 3 lists all possible combinations of the four alternative angles, with the two alternative surface treatments, as well as the internal data modeling alternatives of using either standardized scaling or not (see the section on PLS above).

It is critical to understand why we have chosen this practical and complete model prediction assessment as optimization criterion. This has to do with the fact that the indices employed in the statistical prediction strength assessment (slope and correlation coefficients) are manifestations of the total, accumulated prediction accuracies and precisions respectively. These multivariate calibrations encompass all the practically relevant sampling, illumination, modeling and prediction error components associated with presenting powder samples to an image analytical acquisition system. This type of empirical testing is much to be preferred *en lieu* of various internal statistical tests only.

Table 3.
Selected prediction model evaluation parameters (Slope = Sl and regression coefficient = r^2) from models for Amaranth, Couscous and Paprika. These two statistics constitute the most informative, immediately accessible prediction performance indices for lateral comparison between models, Esbensen (2001).

		Angle 1 (23º)		Angle 2 (35º)		Angle 3 (48º)		Angle 4 (63º)	
		Scaled	Non-sc.	Scaled	Non-sc.	Scaled	Non-sc.	Scaled	Non-sc.
Amaranth	Scrap.	Sl: 0.88 r^2: 0.86	Sl: 0.93 r^2: 0.98	Sl: 0.81 r^2: 0.83	Sl: 0.79 r^2: 0.93	Sl: 0.92 r^2: 0.92	***Sl: 0.95*** ***r^2: 0.96***	Sl: 0.78 r^2: 086	Sl: 0.80 r^2: 0.88
	Prist.	Sl: 0.78 r^2: 0.77	Sl: 0.78 r^2: 0.81	Sl: 0.88 r^2: 0.92	Sl: 0.96 r^2: 0.79	Sl: 0.87 r^2: 0.64	Sl: 0.92 r^2: 0.66	Sl: 0.91 r^2: 0.83	*Sl: 0.99* *r^2: 0.85*
Couscous	Scrap.	Sl: 0.34 r^2: 0.29	Sl: 0.32 r^2: 0.21	Sl: 0.75 r^2: 0.62	Sl: 1.10 r^2: 0.96	Sl: 0.94 r^2: 0.98	*Sl: 0.97* *r^2: 0.98*	Sl: 0.92 r^2: 0.94	Sl: 0.88 r^2: 0.94
	Prist.	Sl: 0.85 r^2: 0.83	Sl: 0.92 r^2: 0.92	Sl: 0.90 r^2: 0.98	Sl: 0.97 r^2: 0.96	***Sl: 1.00*** ***r^2: 0.98***	Sl: 0.95 r^2: 0.98	Sl: 0.85 r^2: 0.85	Sl: 0.84 r^2: 0.81
Paprika	Scrap.	Sl: 0.91 r^2: 0.92	Sl: 0.94 r^2: 0.94	Sl: 0.92 r^2: 0.92	Sl:0.93 r^2: 0.92	Sl: 0.95 r^2: 0.96	Sl: 0.98 r^2: 0.98	Sl: 0.99 r^2: 0.98	***Sl: 1.00*** ***r^2: 0.98***
	Prist.	Sl: 0.72 r^2: 0.77	Sl: 0.86 r^2: 0.90	Sl: 0.88 r^2: 0.90	Sl: 0.91 r^2: 0.94	Sl: 0.95 r^2: 0.96	*Sl: 0.98* *r^2: 0.96*	Sl: 0.79 r^2: 0.76	Sl: 0.78 r^2: 0.79

Bold & italics designates best model. Italics alone designates best model for *opposite* surface treatment.

3.2 First Models for Agglomeration Quantification

Our first foray into quantitative modeling of agglomeration made use of the following seven powders: amaranth, couscous, paprika, ground corn, red lenses, rice (brown, short) and sand (fine, clayey), which were analyzed in the exact same way as above (Table 3). Selected prediction parameters for the final models are presented in Table 4.

All final models are *outlier-screened*, 11-segment cross-validated prediction valuations, following the procedures in [9,16,17]. We here only present the two main statistics: the slope of the "predicted versus measured" model fit, Sl, and the squared correlation coef-

ficient, r^2, of which the former is a visually sharp assessor of the overall accuracy of prediction, the latter taking good care of the precision (there is a one-to-one relationship with the alternative precision statistic, RMSEP).

Table 4.
Prediction performance statistics for final models for all powders at an angle of 48º.

Initial experiments			
		Auto-scaled	Non-scaled
Ground corn	Scraped	*Sl: 0.93 - r^2: 0.92*	Sl: 0.92 - r^2: 0.92
	Pristine	***Sl: 0.96 - r^2: 0.98***	Sl: 0.93 - r^2: 0.98
Couscous	Scraped	Sl: 0.94 - r^2: 0.98	*Sl: 0.97 - r^2: 0.98*
	Pristine	***Sl: 1.00 - r^2: 0.98***	Sl: 0.95 - r^2: 0.98
Paprika	Scraped	Sl: 0.95 - r^2: 0.96	***Sl: 0.98 - r^2: 0.98***
	Pristine	Sl: 0.95 - r^2: 0.96	*Sl: 0.98 - r^2: 0.96*
Amaranth	Scraped	Sl: 0.92 - r^2: 0.92	***Sl: 0.95 - r^2: 0.96***
	Pristine	Sl: 0.88 - r^2: 0.64	*Sl: 0.92 - r^2: 0.66*
Red lenses	Scraped	*Sl: 0.56 - r^2: 0.59*	Sl: 0.50 - r^2: 0.44
	Pristine	***Sl: 0.97 - r^2: 1.00***	Sl: 0.95 - r^2: 0.96
Short brown rice	Scraped	Sl: 0.79 - r^2: 0.79	*Sl: 0.85 - r^2: 0.81*
	Pristine	Sl: 0.81 - r^2: 0.96	***Sl: 0.88 - r^2: 0.98***
Clayey sand	Scraped	*Sl: 0.96 - r^2: 0.98*	Sl: 0.93 - r^2: 0.94
	Pristine	***Sl: 0.98 - r^2: 0.85***	Sl: 0.95 - r^2: 0.96
Subsequent experiments			
Plastic pellets	Pristine	***Sl: 0.37 - r^2: 0.26***	Sl: 0.37 - r^2: 0.25
Clay	Pristine	Sl: 0.97 - r^2: 0.92	***Sl: 0.99 - r^2: 0.94***
Sand (0.18 mm)	Pristine	***Sl: 0.95 - r^2: 0.94***	Sl: 0.93 - r^2: 0.94
Cellulose	Pristine	***Sl: 0.96 - r^2: 0.96***	Sl: 0.94 - r^2: 0.94
Cement	Pristine	***Sl: 0.76 - r^2: 0.72***	Sl: 0.73 - r^2: 0.66
Alumina	Pristine	Sl: 0.93 - r^2: 0.92	***Sl: 0.94 - r^2: 0.92***
Talc	Pristine	***Sl: 0.98 - r^2: 0.96***	Sl: 0.98 - r^2: 0.94
Microdolomite	Pristine	Sl: 0.92 - r^2: 0.79	***Sl: 1.02 - r^2: 0.79***

Bold & italics denotes best model. Italics alone denotes best model for *opposite* surface treatment.

4 DISCUSSION

The above first foray is now augmented by an additional eight powders. All powders are presented in full in Table 4, in which the last eight powder type entries are only represented by the pristine surface treatment (reasons given below). The surface treatment optimization is to be one of the primary results from the present data analysis (reported on in both this and the companion paper II). At this stage we really do not know (yet) which is performing "best".

In this results overview we have also included the data analytical ambiguity of whether to use "auto-scaled" data or not. Typically, from overall chemometric experience, there is no telling which of the latter alternatives are best in any new data analysis setting [9,16,17]. It is often a very practical matter to be settled by appropriate experimentation. The laboratory pilot-study context should always be closely similar to the future industrial setting.

The results are revealing in their apparent first *ambiguity* that there is no universal optimal setting (illumination angle, surface treatment, data analytical scaling) for all powders. For each powder type in Table 4, we have highlighted the "best" prediction model (bold and italics) as well as the second-best (italics) - for the *opposite* surface treatment.

The first significant observation is that there are not very big differences between many model alternatives, so that many "best" choices, in reality are only marginally so in detailed comparison. This apparently ambiguous feature will actually allow us to make more general conclusions however.

1. There would appear to be approximately as many "best" prediction models for the auto-scaled (standardized) data models as for the non-scaled alternatives. As this internal data analytical choice is of absolutely no practical consequence whatsoever with regard to establishing PLS-R models (it takes exactly the same amount of computer time to do either), this can very well be decided upon on an individual powder type basis. Whether to scale or not when making calibrations for powders is problem-dependent.
2. There would clearly appear to be a higher proportion of pristine "best models". Detailed inspection of Table 1 reveal a 5 out of 7 best "non-manipulated" models (only gently shaken pre-treatments). Also, for the two "opposite" models, which are deemed "best" with regard to the scraped surface treatment, the alternative pristine models are only marginally worse off (amaranth) or insignificantly different (couscous), all in all strongly pointing to the pristine (non-manipulated surface) option as the final verdict. This is the reason we only include these models for the last eight powders in the full results reporting. Presumably it is also most desirable with the *least* mechanical industrial image analytical setup, i.e. no mechanical surface scraping on top of the conveyor belt load etc.
3. There are clearly many possible ways to achieve quite satisfactory prediction models for automated quantitative prediction of the degree of agglomeration by using an off-the-shelf, inexpensive R/G/NIR image analytical system, combined with simple, well-tested chemometric data analysis (AMT image pre-processing and PLS-R prediction modeling).
4. The present models have all been subjected to 11-segment so-called full cross-validations, which are quite satisfactory for the *comparative purpose* of finding one (more) optimal predictor model configuration amongst a set of alternatives.
5. Still, we have *also* performed the ultimately desirable, most relevant and most reliable practical test, based upon two independent new *test sets* [16]. Thus we have duplicated the entire experiment for two representative powders: sand and clay. We present these results, and compare them with the original models for these two powders, in Table 5. Here we show results for both scaled as well as non-scaled models. Detailed comparisons show quite convincingly that the results in Table 4 are not a consequence of statistical flukes, as both test set validations compare very closely with the original, cross-validated models. This is proof positive, from the practical context (new, independently drawn test sets), that we have indeed been able to develop – and fully validate using comparative cross-validation – the desired image analytical powder agglomeration predictor. Figures 11 through 14 show the model overview of the four test-set validated models (auto-scaled and non-scaled) for clay and sand (0.18 mm).
6. In the planned companion paper II we shall further explore a more full understanding of which powder characteristics lies behind this success of practical agglomeration description and quantitative prediction modeling as well as other objectives.

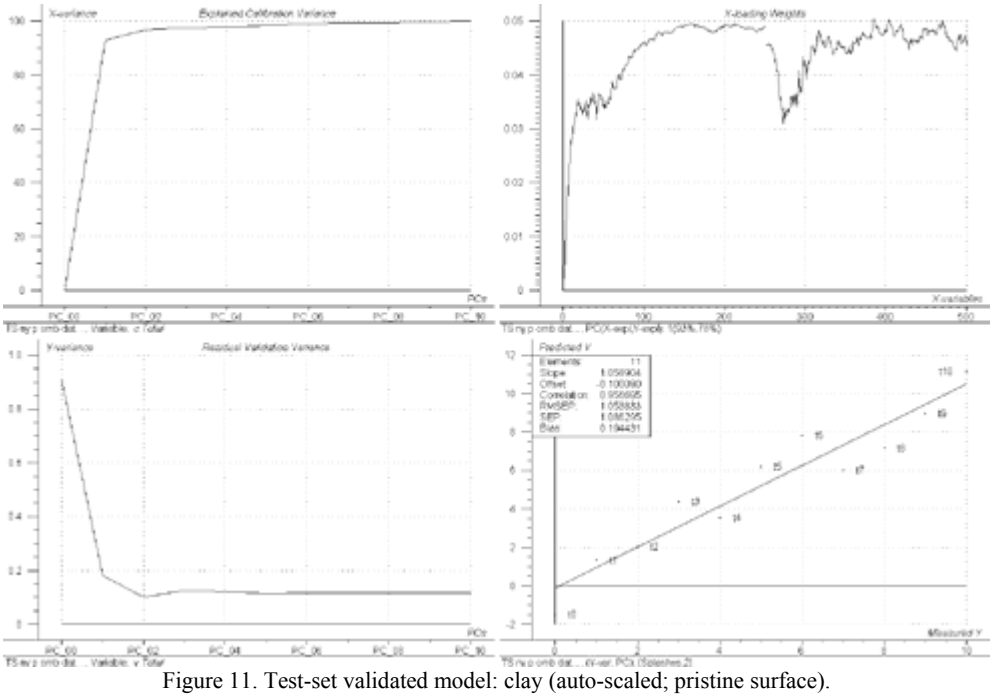


Figure 11. Test-set validated model: clay (auto-scaled; pristine surface).

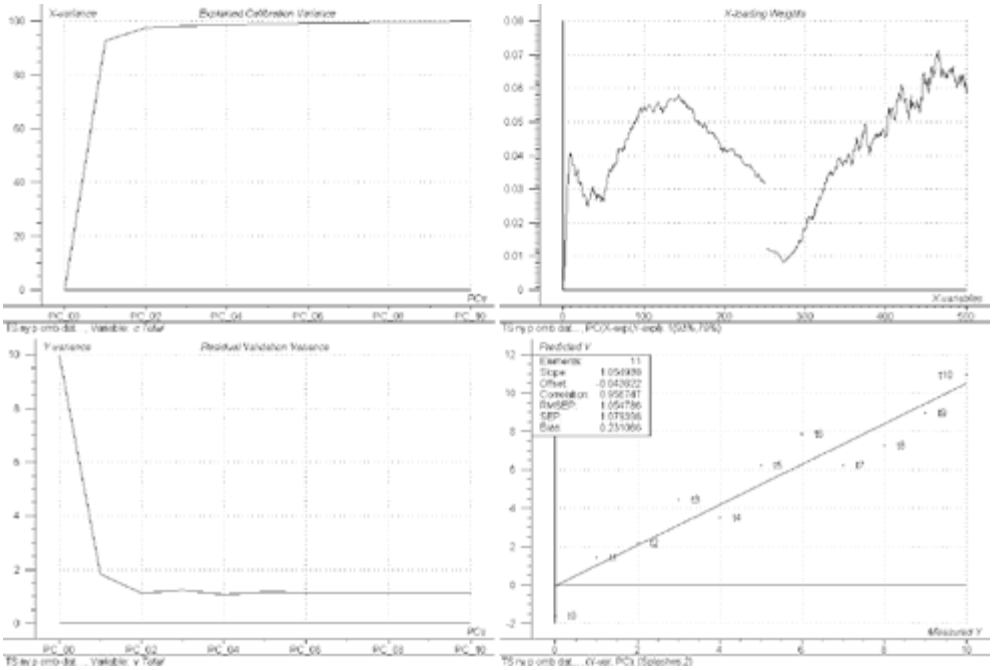


Figure 12. Test-set validated model: clay (non-scaled; pristine surface).

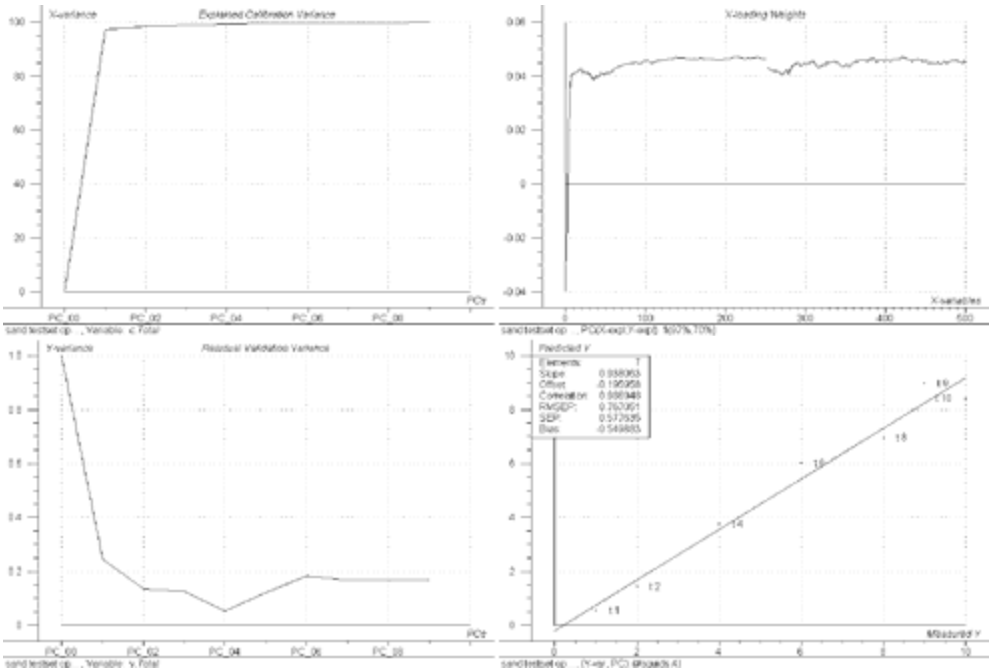


Figure 13. Test-set validated model: sand (auto-scaled; pristine surface).

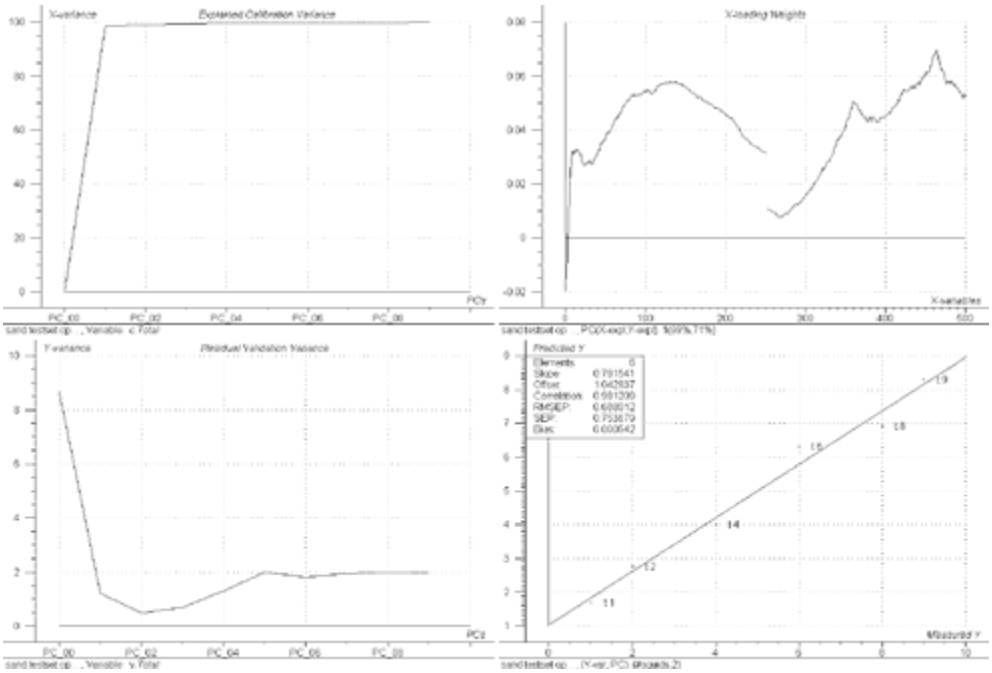


Figure 14. Test-set validated model: sand (non-scaled; pristine surface).

Table 5. Summary of slope and regression statistics for test set-validated models.

		Auto-scaled	Non-scaled
Clay	Non-m.	Sl: 0.97 - r^2: 0.92	Sl: 0.99 - r^2: 0.94
Sand (0.18 mm)	Non-m.	Sl: 0.95 - r^2: 0.94	Sl: 0.93 - r^2: 0.94
Clay test-set	Non-m.	Sl: 1.06 - r^2: 0.92	Sl: 1.05 - r^2: 0.92
Sand test-set	Non-m.	Sl: 0.94 - r^2: 0.98	Sl: 0.79 - r^2: 0.98

Interpretation of results: Individual Models

As a final illustration of the coverage of the 15 powders employed, we present all the relevant predicted vs. measured assessment plots. For all the powders we have results for the pristine surface, while only the first seven were modeled for the scraped (non-manipulated) surface treatment.

The graphical depictions below are intended to convey a direct graphic overview of the degree of difference with which it has been possible to construct useful models for predicting the degree of agglomerative behavior, and thus also the potential for indirect prediction of moisture content. There are some powders in this selected set which do not lend themselves to agglomeration modeling (at all); these will give rise to distinctly inferior prediction model statistics, while many other powders are subject to easy modeling. Only these latter will be amendable to the methodology developed in this work. Naturally not all powder types are significantly affected by moisture-induced agglomeration.

Ground Corn

Ground corn is of the intermediate class of powders with regard to particle size in the present setting. Ground corn is relatively well modeled, both surface scraped and pristine. The reason for this, we believe, is found in the fact that the powder is “sticky”. Ground corn absorbs water with relative ease.

Couscous

Couscous is extremely well modeled with the pristine surface, but a little worse off when scraped. The reason for this is found in that very large “snow-ball-effect” agglomerates are produced when the sample is scraped. Increments 7-10 for the scraped samples produce very nearly equal results, possibly because of saturation and thus had to be removed from the dataset.

Paprika

Paprika expresses no clear preference towards either surface treatment and is well modeled in both cases.

Amaranth

Amaranth produces “intermediately good models” due to the fact that it is mostly water-films on the surface of the seeds which are binding/holding the agglomerates together. The scraped models are best, possibly due to the fact that nicely distributed “crevices” appear when the knife passes by. The non-manipulated models are not that good, probably because the agglomerates are easily broken even during the gentlest transportation/ shaking.

Red Lenses
This powder was initially believed impossible to model, due to the very large grains. The scraped models were of very low quality because the knife removed almost all the sample material from the Petri dish due to the fact that all particles were effectively "glued" together into one giant nugget. The pristine models, however, turned out very well, even though the particles "glued" together very rapidly after addition of water, and the initial Petri dish "gentle shaking" had absolutely no effect.

Short Brown Rice
Another example of where pristine models exceed the scraped ones in prediction performance, mainly due to the same reasons as for Red Lenses above (though in a downscaled version). A few samples with high addition of water had to be removed from the models due to saturation effects.

Clayey Sand
The models for the pristine surface were slightly better. Sand is difficult to model, because the agglomerates are very instable and the heat from the (1000 W) lamp causes the water to evaporate relatively fast from the surface. Even the gentle rotation of the sample-holder resulted in destruction of some agglomerates. This calls for *extensive* effort to produce similar conditions if applied to industrial large scale conditions.

Plastic Pellets
Very low quality models were produced because of the large particle size, and the lack of ability of the particles to find "joint surfaces" where the water can bind the particles together. This is due to the fact that the shapes of the particles are rather irregular ("morphed" cylinders).

Clay
The latter eight powders were only modeled for the pristine surface, due to superior models for this option for the previous seven powders. Clay is well modeled due to the stickiness of the powder, which is certainly a reflection of its very high water holding capacity.

Sand (0.18 mm)
The same comments are valid as above for Clayey Sand.

Cellulose (micro-crystalline)
Very good models are established, hence opening for a wide range of potential application areas in the pharmaceutical industries where micro-crystalline cellulose is a major filler component. The high contrast between the white powder and the illumination-generated shadows is believed to influence positively on the modeling.

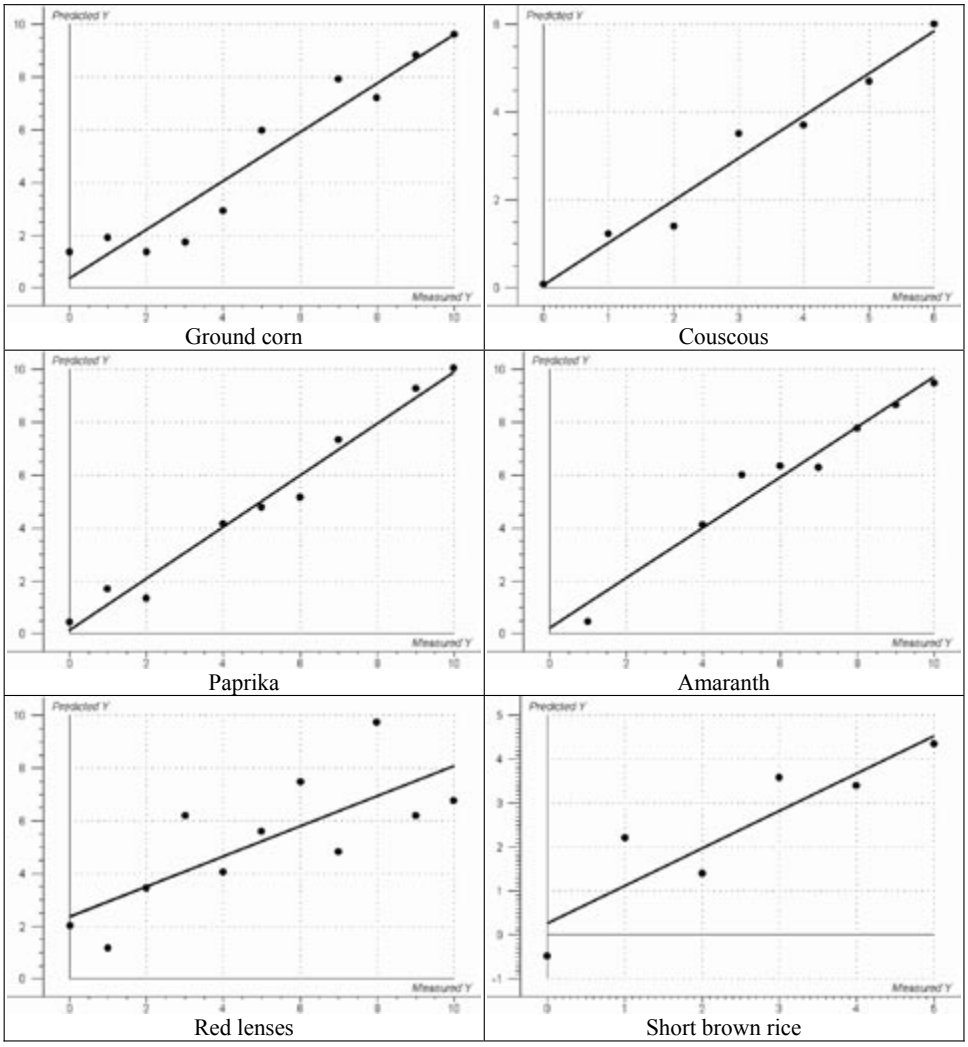
Cement
The models produced are only of intermediate quality, mainly due to the (obvious) ability of cement and water to create very hard agglomerates! The reason for the lower quality models is also believed to be associated with the low contrast between the cement and the shadows.

Alumina
Again good models could be established, mainly due to identical reasons as for cellulose.

Talc
Also very good models – explanation are identical to those for cellulose and alumina.

Microdolomite
This last, white, powder also produced very good models, due to the same explanations as with the other fine, white powders in the selected training set.

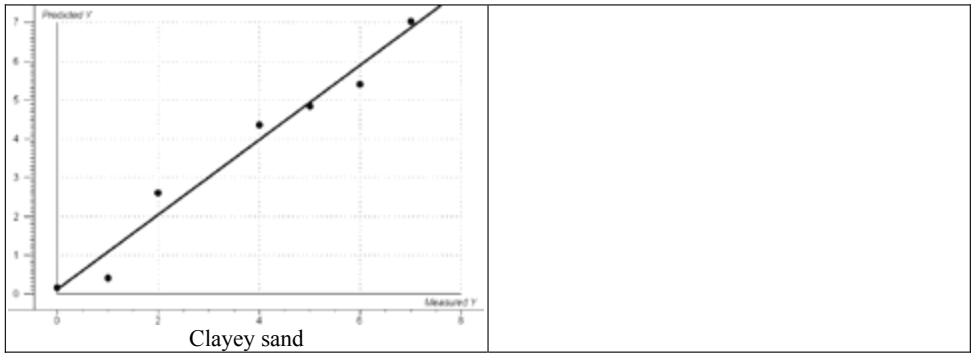
Figures 15 and 16 below provide an overview of the regressions (predicted vs. measured values) for all fifteen powders with both surface treatment alternatives, scraped or pristine, respectively. The large span in model quality is easily appreciated from comparison of the figures and relating to the above individual comments.



Ground corn | Couscous

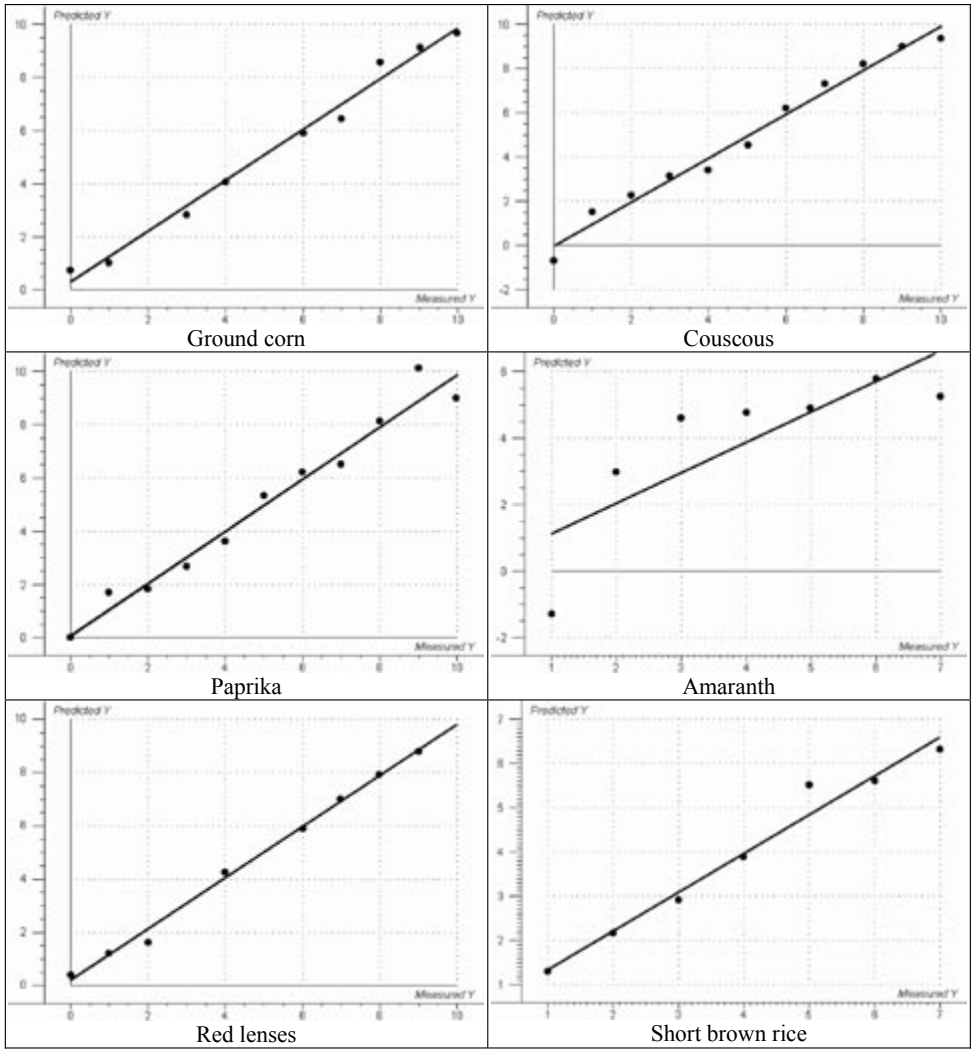
Paprika | Amaranth

Red lenses | Short brown rice



Clayey sand

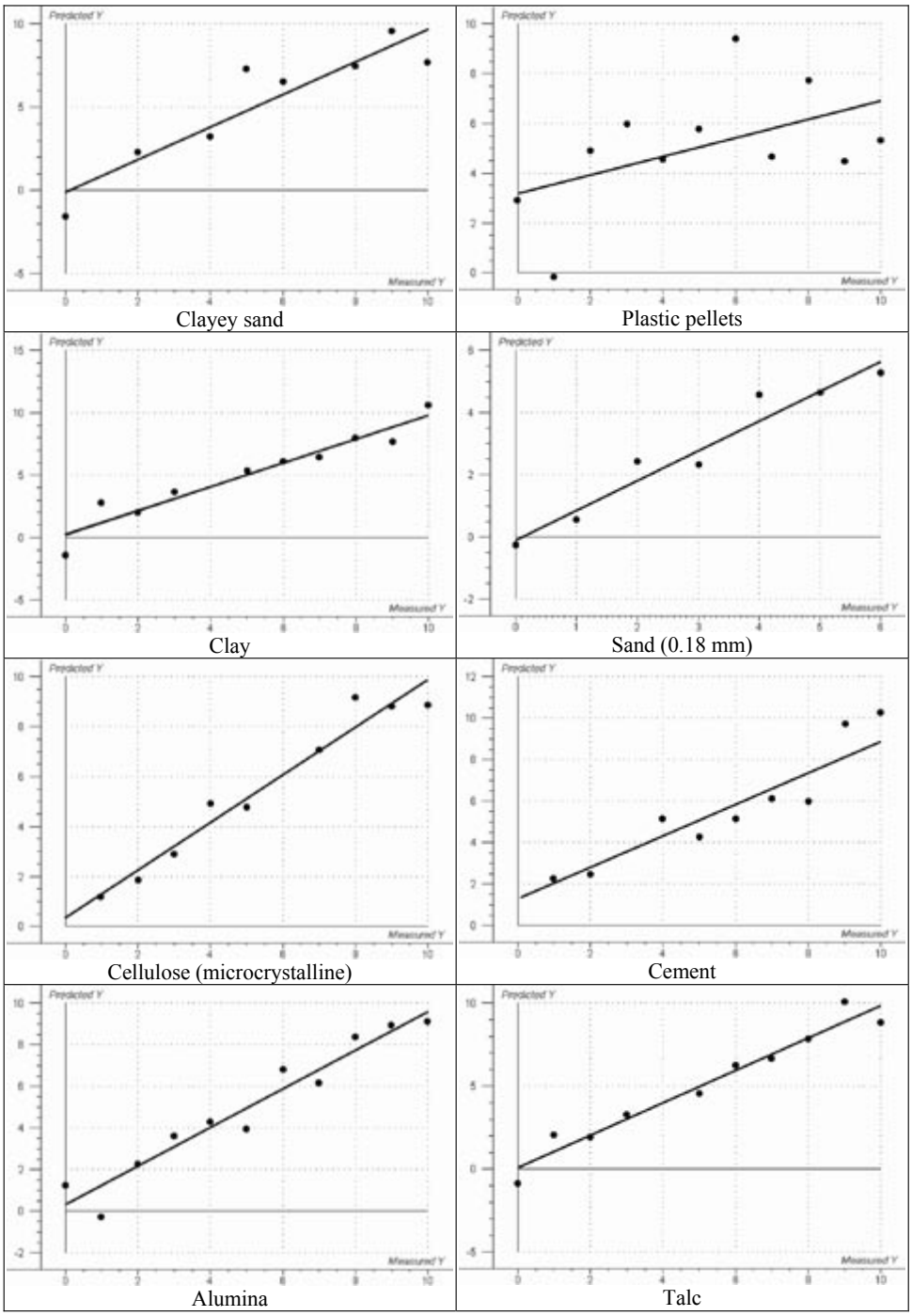
Figure 15. Predicted vs. measured plots for the best *scraped* models (see Table 4).



Ground corn | Couscous

Paprika | Amaranth

Red lenses | Short brown rice

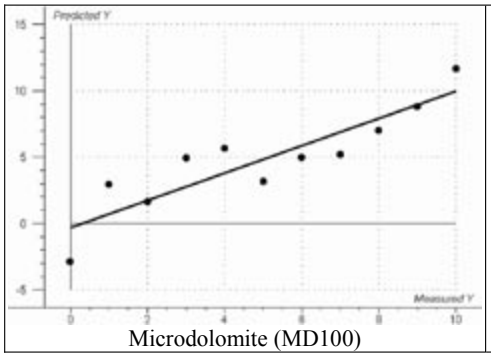


Clayey sand | Plastic pellets

Clay | Sand (0.18 mm)

Cellulose (microcrystalline) | Cement

Alumina | Talc



Microdolomite (MD100)

Figure 16. Predicted vs. measured plots for the best pristine models (see Table 4).

5 CONCLUSIONS

A series of 15 representative industry-related powders have been subjected to a new image analytical quantitative agglomeration characterization, tantamount to indirect moisture determination. All details of the new methodology (fully developed and described here), were optimized *on average* for this set of powders.

An optimal illumination angle of 48° (of the four tested) was used throughout all experiments. It must be emphasized that the angle of illumination must be optimized for every specific industrial application context.

The results clearly show that the pristine surface treatment option gave the best models for the powders tested, which has desirable consequences for industrial applications (no conveyor belt scraping necessary, or similar). Thus this feasibility study finds the methods described in the present work as successful candidates for future *automated* quantification of powder agglomeration. The degree to which this will hold up also in relatively rough industrial settings will depend on good process automation engineering; judging from today's many other types of successful image analytical implementations this would appear eminently doable.

Only very little effort has yet gone into optimizations of *individual* powder-specific image acquisition and the corresponding data analytical modeling, which for all particular industrial setups will result in higher accuracy and precision than those for the present global models in future applications. The successful global results presented here for 15 decidedly very different powders bode very well for this issue.

Data analysis of the extensive results did not resolve whether to use auto-scaling or not, so this has to be decided upon on an individual application basis, as indeed all aspects of which data and model pre-treatment to use. A global answer to this issue was sought, but none appeared. However this is of no practical consequence, since either alternative will not result in any discernable difference in computation time. Different (spectral) types of cameras, different focal distances and especially a higher number of sample replicates should of course be tried out in the specific implementation contexts.

The present results are not necessarily restricted only to water-induced agglomeration. We can recommend further application of this type of chemometric image analysis/AMT/PLS methods also for other types of particulate matter e.g. powder mixtures, slurries as well as other industrial materials of similar surface texture manifestations, even certain types of solid surfaces amongst other.

6 REFERENCES

[1] N. Zeuch, Understanding and Applying Machine Vision, 2.nd Ed. Marcel Dekker, Inc. ISBN 0-8247-8929-6
[2] M. Rhodes, Principles of Powder Technology, John Wiley and Sons Ltd., ISBN: 0-471-92422-9 (1990)
[3] M. Rhodes, Introduction to Particle Technology, John Wiley and Sons Ltd. ISBN: 0-471-98483-3 (1998)
[4] Y. Wanibe, T. Itoh, New Quantitative Approach to Powder Technology, John Wiley and Sons Ltd., ISBN: 0-471-98154-0 (1998)
[5] K.H. Esbensen, K.H. Hjelmen, K. Kvaal, The AMT Approach in Chemometrics – First Forays, Journal of Chemometrics, Vol. 10, pp. 569-590 (1996)
[6] J. Huang, K.H. Esbensen, Applications of Angle Measure Technique (AMT) in Image Analysis, Part I: A New Methodology for In Situ Powder Characterization, Chemometrics and Intelligent Laboratory Systems 54, pp. 1-19 (2000)
[7] J. Huang, K.H. Esbensen, Applications of AMT (Angle Measure Technique) in Image Analysis, Part II: Predictions of Powder Functional Properties and Mixing Components Using Multivariate AMT Regression (MAR), Chemometrics and Intelligent Laboratory Systems 57, pp. 37-56 (2001)
[8] T.T. Lied, Multivariate Image Regression (MIR) for Quantitative Predictions - Prototype Software Implementation and Selected Industrial-Technological Pilot Studies, Ph.D. thesis, the Norwegian University of Science and Technology (NTNU), ISBN: 82-7984-126-1 (2000)
[9] J. Huang, Developments in Applied Chemometrics – AMT, Acoustic Chemometrics and N-way Image Analysis, Ph.D. thesis, the Norwegian University of Science and Technology (NTNU), ISBN: 82-7984-187-3 (2001)
[10] K. Kvaal,J. P. Wold, U. G. Indahl, P. Baardseth, T.Næs, Multivariate Feature Extraction from Textural Images of Bread, Chemometrics and Intelligent Laboratory Systems 42, pp. 141-158 (1998)
[11] U. G. Indahl, T. Næs, Evaluation of Alternative Spectral Feature Extraction Methods of Textural Images for Multivariate Modelling, Journal of Chemometrics 12, pp. 261-278 (1998)
[12] B. Egelandsdal, et.al, Evaluation of Scanning Electron Microscopy Images of a Model Dressing Using Image Feature Extraction Techniques and Principal Component Analysis, Scanning 21, pp. 316-325 (1999)
[13] P. P. Mortensen and K. H. Esbensen, Optimization of the Angle Measure Technique for Image Analytical Sampling of Particulate Matter, Chemometrics and Intelligent Laboratory Systems (in print 2004)
[14] R. Andrle, The Angle Measure Technique: A New Method for Characterizing the Complexity of Geomorphic Lines, Mathematical Geology 26, pp. 83-97 (1994)
[15] J. Huang, S. Ose, S. de Silva, K.H. Esbensen, Non-invasive monitoring of powder breakage during pneumatic transportation using acoustic chemometrics, Powder Technology 129, pp. 130-138 (2003)
[16] K.H. Esbensen, Multivariate Data Analysis – in Practice, 5th ed., CAMO ASA, ISBN: 82-993330-2-4 (2001)
[17] H. Martens, T. Næs, Multivariate Calibration, John Wiley and Sons Ltd., ISBN: 0-471-93047-4 (1989)

Paper II

Jens Bo Holm-Nielsen*, Casper K. Dahl, and Kim H. Esbensen
"Representative Sampling for Process Analytical Characterization of Heterogeneous Bioslurry Systems — a Reference Study of Sampling Issues in PAT"

Chemometrics and Intelligent Laboratory Systems
Volume 83, pp. 114–126, 2006.

Available online at www.sciencedirect.com
ScienceDirect
Chemometrics and Intelligent Laboratory Systems 83 (2006) 114–126
Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

Representative sampling for process analytical characterization of heterogeneous bioslurry systems—a reference study of sampling issues in PAT

Jens Bo Holm-Nielsen *, Casper K. Dahl, Kim H. Esbensen

ACABS Research Group, Aalborg University Esbjerg, Niels Bohrs Vej 8, DK-6700, Esbjerg, Denmark

Received 29 March 2005; received in revised form 28 December 2005; accepted 13 February 2006
Available online 6 March 2006

Abstract

Extensive experimentation and evaluation of alternative sampling methods of highly heterogenous bioslurries are carried out in the context of the theory of sampling (TOS) in order to delineate optimal, *representative* sampling procedures in PAT (process analytical technologies). The analytical methods investigated are at-line NIR and image analysis (IA) for monitoring of an industrial bioenergy anaerobic digestion processes (AD), subject to stringent economic bracketing: PAT solutions have to be both *practical* and *inexpensive* in the bioenergy sector where cost efficiency of instrumentation and monitoring systems is an absolute must. Focus is on development of minimum expenditure methods for sufficient characterization of the very heterogeneous types of biomass feedstock as used in industrial scale, continuously stirred tank reactors (CSTR). Product and process characterization necessarily involves a chemometric multivariate calibration predictor (PLS regression). The general goal is development of appropriate sampling/PAT facilities for at-line/on-line process monitoring in typically low-tech bioenergy, agro-industrial sectors. Experimental laboratory reactor evaluations, based on biomass feedstock and digested products from a full-scale biogas plant, were initially run in batch mode for a week, followed by fed-batch addition of maize silage, introducing a systematic increase in total solids allowing properly spanning multivariate calibration models. Measurements on 55 laboratory reactor samples taken during a complete 14-day fermentation cycle included three key process parameters: total solids (TS), volatile solids (VS) and chemical oxygen demand (COD), representing difficult-to-sample analytes. NIR spectroscopy and image analysis, including the angle measure technique transform (AMT), were evaluated for characterization of different feedstocks as well as continuously extracted process samples with respect to selected chemistry and dry matter characteristics. Optimized sampling on four different scale-levels allowed acceptable PLS prediction models for TS and VS for both NIR and image analysis compared to chemical reference analysis, while it was not possible to predict the COD levels satisfactorily due to large uncertainties in mandatory reference measurement protocols. This feasibility study is promising for NIR as at-line prediction of TS and VS content as well as other AD parameters, which can be measured on the same sample types, while image analysis is currently too complex and expensive for these industry sectors. The findings in the present bioslurries study have a considerable generalization potential: all PAT approaches are critically dependent on *representative* reference calibration sampling, which has to be fully compliant with the theory of sampling (TOS).

Keywords: PAT (process analytical technologies); Representative sampling; TOS (theory of sampling); Chemometrics; AMT (angle measure technique); Image analysis; PLS regression; Anaerobic digestion; Bioslurry; Biogas production

1. Introduction

Fermentation processes at large-scale biogas plants are sensitive to sudden changes in feedstock composition, which cause significant variability in the process conditions. Today, fermentation process control is achieved through manual sample extraction (shown below to be highly problematic due to the very heterogenous nature of the fermentation media) with off-line (sometimes even off-site) analysis of a few key process parameters such as total solids (TS), volatile solids (VS) and chemical oxygen demand (COD), as well as volatile fatty acids (VFA), total-N and ammonia content. However, other parameters, e.g., process temperature, pH, volumetric biogas yield as well as methane, carbon dioxide and hydrogen sulfur content

* Corresponding author.
E-mail address: jhn@aaue.dk (J.B. Holm-Nielsen).
URL: http://www.acabs.dk (J.B. Holm-Nielsen).

0169-7439/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2006.02.002

in the gaseous headspace of the reactor, are also monitored more-or-less on-line as control parameters in more advanced plants [1,2].

Even though time-constants involved in anaerobic digestion (AD) are not critically short (average hydraulic retention time of feedstock in a semi-continuously fed-batch digestion system is of the order of 10 to 25 days), there is nevertheless a desire for implementation of on-line technologies, of a distinctly inexpensive, robust types. The perspective is to develop real-time process monitoring data at full-scale commercial biogas plants above a certain minimum size to be used for improved routine managing tools, i.e., for identification of critical process state parameters, upsets and for monitoring when unavoidable deviations in feedstock composition occurs. By combining low-tech, but reliable sensors with critical at-line facilities if/where/when physical samples must be examined, the goal is to contribute towards a new generation of fully integrated process management systems for the daily operating plant staff [3,4]. This endeavor fits well with current EU renewable energy initiatives for considerably increased utilization of biofuels including biogas, integrated in the overall energy as electricity, heat and in the transportation fuel sectors.

Improved control of this type of fermentation processes especially calls for more controllable raw materials characterization throughout, which today are often based on slow off-line measurements. Improved optimization of anaerobic digestion processes can only be achieved by incorporating new, relevant sensor technologies for semi-continuous at-line/on-line measurements of the composition of both raw materials and processed biomass at important process stages. The most relevant sensor deployment locations with this scope include:

- *Feedstock* arrival depositories (raw materials)
- *Inlets* (feed-points) for fermentation reactors
- *In-process deployments* (in reactors and transportation pipelines)
- *Outlets* (exit-points) for the fermentation reactors

Any process analytical technology (PAT) is critically dependent on the representativity of the signals obtained from on-line or in-line probes or sensors in pipelines or reactor tanks as well as representativity of the samples analysed for calibration. For example, what is the relationship between a sensor or probe's field-of-view (a few mm^2) in relation to the entire cross-section of a transportation pipeline or the entire reactor volume in question? What with TOS is termed "incorrect sampling procedures", which are legion in the AD regimen, will always result in biased, non-representative and hence unreliable analytical results [5–9]. All traditional sampling procedures involved in the general AD process flow are therefore carefully evaluated in the present study in the light of TOS (theory of sampling) and the specific sampling errors are estimated where needed in order to put these hitherto largely neglected sampling issues on a firm quantitative footing.

Sampling is a critical component in any analytical procedure and must always ensure representativity of the primary, secondary and tertiary sampling steps. Total sampling errors typically can be in the order of 100+ times the specific analytical error. "Incorrect sampling" always results in non-representative samples because of an uncontrollable bias even though all subsequent steps have been performed in a correct way [5–8].

Animal manure forms a highly complex and heterogeneous suspension because of a significant proportion of solids, "dry matter" (3–12%): straw, grass, undigested lignocellulosic fiber particles, besides not yet digested macromolecule aggregates such as starch, sugars and proteins. Such suspensions are therefore always prone to segregation on several scales, which sets in immediately after agitation has been terminated. Sampling from such systems consequently will introduce a significant sampling bias if not properly counteracted, wherefore this study will deal exhaustively with all relevant sampling issues using the complete TOS toolbox [8]. In this context, focus will be on a general sampling hierarchy, covering all aspects from primary field sampling to the ultimate, apparently insignificant, mass reduction involved in securing the often minute analytical volume.

This study also evaluates the potential of an imaging technique in characterizing important physical and chemical parameters. Since NIR reflectance/transmission has shown good potential in mixed liquids in linked fermentations sectors, a NIR reflectance system was employed for reference comparison and/or as the main PAT technique. NIR has been tested extensively with varying degrees of success as a spectroscopy tool for on-line bioslurry monitoring of continuous AD systems and similar [2,10–13].

We here combine image analysis (IA) and the angle measure technique (AMT), a powerful texture extraction technique [14–19], with chemometric multivariate modeling (PLS-R) in a pilot study with a potential for on- or at-line monitoring and prediction of visible features such as the surface manifestation of bioslurries and the TS content. By virtue of stoichiometric closed array relationships, there is also a possibility for *indirect* multivariate calibration modeling for correlated parameters, VS and perhaps even COD a.o. Such methods constitute a potential for monitoring and optimization of full-scale production plants.

2. Sampling in AD systems for multivariate calibration

Procurement of feedstock as well as digested biomass in preparation for AD trials and testing took place at a full-scale industrial biogas plant, the Ribe Biogas plant, Denmark. Primary and secondary sampling and preparation will be described in detail below. Experimental fermentation trials–including a tertiary sampling step–is a commonly used procedure both in academic studies as well as in commercial laboratories to study biomass feedstock compositions and to characterize their biogas potentials [1,2]. This is followed by a quaternary sampling procedure for securing representative analytical samples for the final NIR and IA analysis. Fig. 1 is a schematic illustration of the entire AD process and the attendant sampling hierarchy as examined in this study.

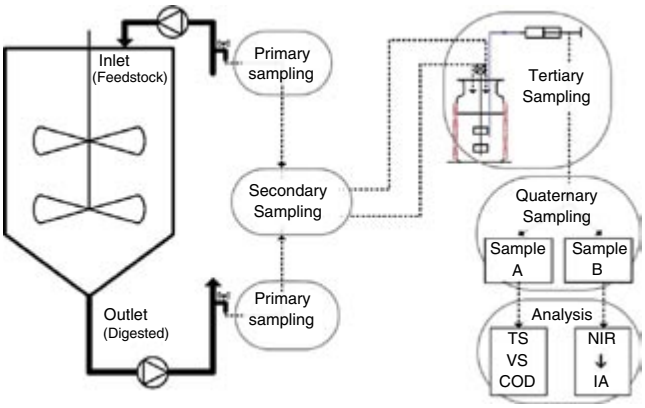


Fig. 1. Schematic illustration of the necessary sampling hierarchy for the AD process, including primary sampling at the full-scale biogas plant, secondary sampling preparing samples for fermentation trials, tertiary sampling from laboratory fermentation reactor and quaternary sampling associated with analysis.

2.1. Primary sampling: full-scale biogas plant

Appropriate field sampling procedures based on TOS were employed in order to obtain proper extraction of primary samples from a full-scale biogas plant, the Ribe Biogas plant. The feedstock in terms of total solid contents is here composed of approximately 80% pig and cattle manure to which is added ~20% industrial organic waste from Danish food processing plants (primarily from slaughteries). All feedstocks of this origin utilized for biogas production are very heterogeneous by nature and thus constitute a relevant vehicle for displaying all issues involved when sampling heterogenous materials in general.

Suitable primary extraction sites were located for both feedstock and final digested biomass sampling. Significant gravitational segregation always occurs in horizontal pipelines with higher concentration of solid material along the bottom. Thus, it was decided to sample from a vertical pipeline through side-valves: upward flow greatly facilitates mixing due to turbulence, gravity now acting to boost mixing. The pre-mixed feedstocks was extracted *before* addition to the full-scale fermentation reactor (Fig. 2, left), while digested biomass was sampled directly at the outlet pipeline immediately *after* the fermentation reactor (Fig. 2, right).

It is not only important to point-sample heterogeneous material in well-mixed upflow pipelines—it is equally important to employ extensive time-averaging composite sampling schemes, tuned to the specific situation(s). Thus, primary sampling took place by spanning eight buckets each of 10 L of biomass with 3-min intervals in the continuous pumping.

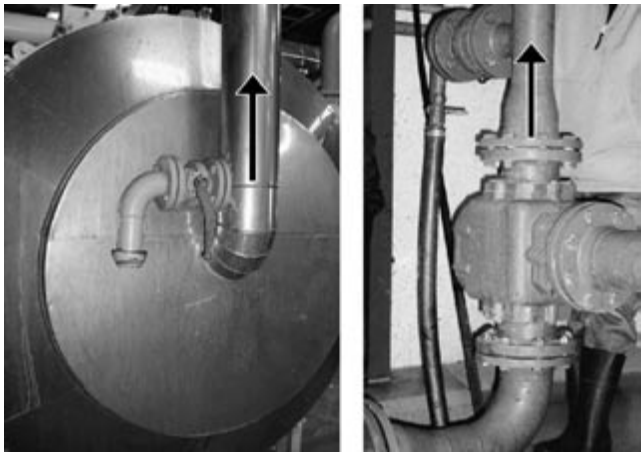


Fig. 2. Primary sampling locations for feedstock and final digestate. Left: outflow from heat-exchanger immediately before inlet to CSTR reactor. Right: sampling of digested biomass at the CSTR reactor outlet. Note sampling only in vertical upward flow.

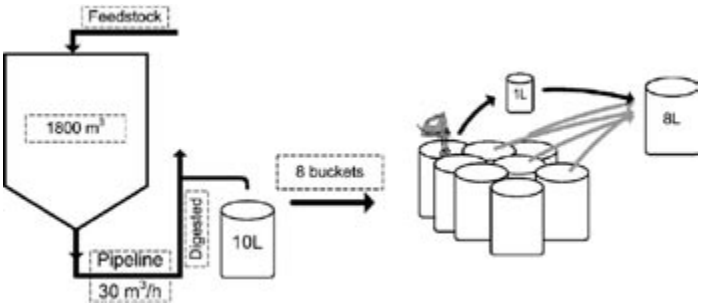


Fig. 3. Schematic illustration of primary sampling at the full-scale biogas plant in which a two-step composite sampling approach was used. Eight 10-L primary *increments* were individually mass-reduced to 1 L, before being compounded. Mechanical agitation is essential to keep the bioslurries in a state of maximum homogenization while being subsampled. This compound sampling scheme is in full accordance with the principles of TOS [5–9].

Feeding and outlet pumping sequences both consist of 30 m^3 biomass/h from an 1800-m^3 biogas reactor. Although the composite sample itself only corresponds to $(8\times10/30{,}000)\times100\sim0.3\%$, the composite sampling rate covered approximately $(15/1800)\times100\sim0.8\%$ of the total reactor volume in the 30-min pumping interval. Sampling rates of the order of 0.1–1.0% are typical of reactor volumes of the current magnitude (1800+ m^3), for very hardheaded practical reasons (Fig. 3).

The bioslurries were first collected from the outlet pipeline in a series of eight 10-L buckets representing time sampling over 3 min. From each of these 10-L buckets, a 1-L increment was extracted while the content was thoroughly mixed by use of a hand-held mixer to ensure maximum homogenization. The resulting eight 1-L increments were subsequently combined into an intermediate 8-L composite sample, for further secondary sampling.

2.2. Secondary sampling: sample size reduction for the fermentation trials and analysis

From this 8-L composite bucket, secondary sampling increments were extracted using a fractional *ladling* technique: each 750-mL secondary sample consists of 15 such increments, which were extracted at various depths in the continuously stirred material. Taking many small increments under continuous agitation leads to a minimization of the dominant sampling error involved, the grouping and segregation error (GSE). This secondary sampling scheme is also in accordance with all appropriate principles in TOS [5–9].

This contrasts drastically with, e.g., one 750-mL grab sample *taken* directly from the pipeline, which would be thoroughly non-representative, ibid. Use of such a scheme for direct grab sampling, i.e., going directly for the final sample volume without TOS-correct mass reduction of the a larger composite alternative delineated above completely destroys any hope of representativity. Such direct grab sampling unfortunately more-or-less rules the day in very many routine, ill-reflected sampling protocols.

The final part of the secondary sampling followed a scheme of randomly choosing 750-mL bottles, which were further subdivided into 120-mL volumes. These were even further subdivided into a number of 30-mL volume vials. At each level, mass reduction was carefully performed by similar agitated fractional ladling, although using with appropriately scaled-down tools, etc. The entire TOS-correct mass reduction of biomass feedstocks (as well as final digested biomass) resulted in thoroughly representative material utilized for all

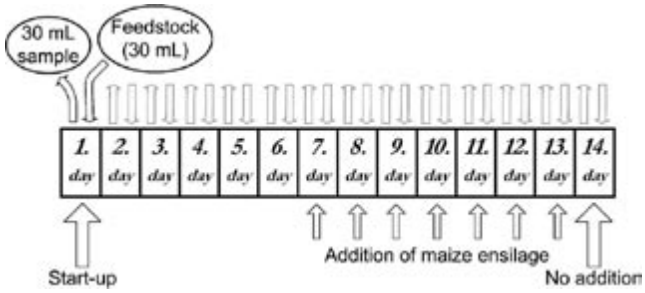


Fig. 4. Fed-batch laboratory reactor operations flowsheet outlining balanced sampling and feeding schemes. Addition of maize ensilage was used to boost biogas production. All 30-mL sampling/30-mL feedstock replacement takes place every sixth hour (four times each day).

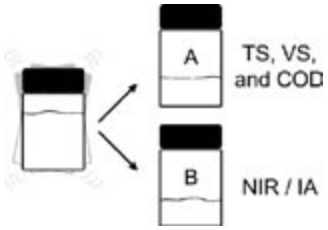


Fig. 5. Schematic illustration of typical laboratory subsampling procedure, here used for chemical reference analysis (subsample A) and NIR/IA (subsample B). *Fractional pouring* (in a composite sampling mode) was used to split the initial sample into two subsamples. Two other alternative subsampling methods in the laboratory were also evaluated, see text.

experimental digestion trials, for chemical reference analysis and for testing the selected PAT tools.

2.3. Tertiary sampling from the laboratory reactor

Mixtures for laboratory-scale fermentation trials were *formulated* as 20% raw feedstock added to 80% the outgoing digested biomass in order to be as identical as possible to the average complex material found in the full-scale fed-batch operated reactors of the biogas plant. In a broad sense, this mixture can be viewed as typical for a great many contemporary AD biogas reactor contents, especially as regards the two-phase bioslurry characteristics.

All anaerobic digestion trials in the laboratory were run for 14 days. After the sixth day, 1% finely macerated maize silage was added daily to increase the dry matter content in the feedstock (Fig. 4). Use of this TS-augmented medium is motivated by new feedstock strategies being contemplated in the European biogas sector focusing more on well-defined feedstocks originating exclusively from farms (in combination with the ubiquitous manure). This study conveniently also uses this total solids (TS) increase to produce a significant calibration span for TS in the multivariate calibration context.

Over the entire 14-day biogas cycle, a 30-mL digested biomass sample was extracted every sixth hour from the laboratory reactor by a syringe device (Fig. 1). This provided all samples used for reference analyses (Y) for both NIR and image analysis X-data. Each 30-mL volume was promptly replaced by an equal amount of fresh feedstock, making the laboratory reactor run in a similar fed-batch mode as its full-scale origin. These 30-mL samples reflect the typical practical difficulties and attendant sampling uncertainties caused in batch-fed or continuously fed reactors—problems typically caused by segregation or sedimentation, inefficient/insufficient steering, inhomogenous fibre distribution a.o.

2.4. Quaternary sampling for analysis

Each 30-mL sample was finally split in two by a fractional pouring technique. From each sample, 15 mL was subsampled for reference analysis (subsample A) by vigorous shaking of the vial followed by immediate pouring of a small fraction into another container. This procedure was repeated several times until approximately 15 mL were obtained as illustrated in Fig. 5; absolutely identical weight was not a critical issue for the subsequent chemical analysis, as different weights were corrected for. The remaining volume constituted the parallel subsample B. This procedure is also meant to represent very many contemporary laboratory techniques as closely as possible.

2.5. Analytical procedures

TS, VS and COD were chosen as experimental analytes in this comparative study, since they are typically used to characterize both feedstock and digested biomass in full-scale biogas production contexts. Fig. 6 shows the uncertainties of the reference measurements—calculated as the relative standard deviation (RSD), defined as [20]:

$$\mathrm{RSD} = \frac{s}{x} \cdot 100\%$$

RSD for reference measurement duplicates show twice as high an analytical uncertainty for COD compared to TS and VS. For routine AD monitoring in the industrial regimen, it is customary not to invest in overly accurate nor precise analytical equipment (including sampling equipment and procedures); one is usually content with analytical reproducibilities of the order of, say, 5+%, which is not acceptable for pilot study laboratory runs however. For COD, the current total reproducibility standard deviation almost reaches 11%, which is way too high for serious data analysis and interpretation. Nevertheless, for the integrity and practical relevance of the present sampling feasibility study, it was necessary to use this standard industrial COD analytical procedure.

2.5.1. Ultimate subsampling of bioslurries (analytical volume)

A special study was deemed necessary, to determine if–or to what extent–the intrinsic material heterogeneity affects splitting into two supposedly identical subsamples (cf. Fig. 5), because later multivariate calibrations are based on these split sample series (parallel samples producing the X- and Y-data, respectively). Subsample A was poured from the predecessor vial, while subsample B is made up of the remaining material. Thus, if segregation occurs at this ultimate, minute scale-level,

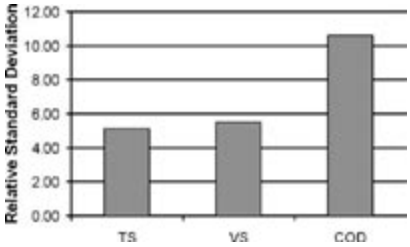


Fig. 6. Relative standard deviation (RSD) for *duplicate* reference analyses for three selected analytes.

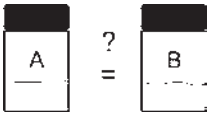


Fig. 7. Analytical volume sample splitting test. Will (50/50) splitting into parallel A and B aliquots by pouring result in two *equal* subsamples (mass, composition), i.e. to which degree will they be representative of the predecessor sample? Sample A is poured from the predecessor vial, while subsample B contains the remaining material.

subsample B should show a higher concentration level of solid matter, but to what extent—significant or negligible (Fig. 7)?

No less than four different types of relevant heterogeneous materials/mixtures were examined for these splitting experiments. The level of heterogeneity, and hence the magnitude of the resulting sampling error, will depend on the material and its specific characteristics. The following *formulations* are based on extensive practical experience with AD systems and materials:

- Feedstock (raw material from the Ribe Biogas plant)
- Digested biomass (also from the Ribe Biogas plant)
- FD (20% feedstock, 80% digested biomass)
- FDM (10% feedstock, 88% digested biomass, 2% maize silage)

The first two substrates were chosen for obvious reasons, since they constitute the relevant bracketing (raw material→end-product) of the industrial biogas production process studied. FD has a composition similar to that under initial laboratory bioprocessing conditions, while FDM has strong similarities with important industrial mixtures, which are routinely "improved" by adding maize silage.

These four material substances were each produced in a series of 30 duplicate split vials, each of 30 mL, by fractional ladling, following identical procedures as outlined above for secondary sampling (Section 2.1). In the results section, we evaluate the effectiveness and representativity of this replication experiment. Analytical results from the comprehensive splitting tests are visualised in Fig. 8 for TS and VS.

In general, VS shows no difference between A and B samples, while there is a clear systematic difference between A and B samples as concerns TS for all four materials. The total solids content of course reflect the dry matter content, the

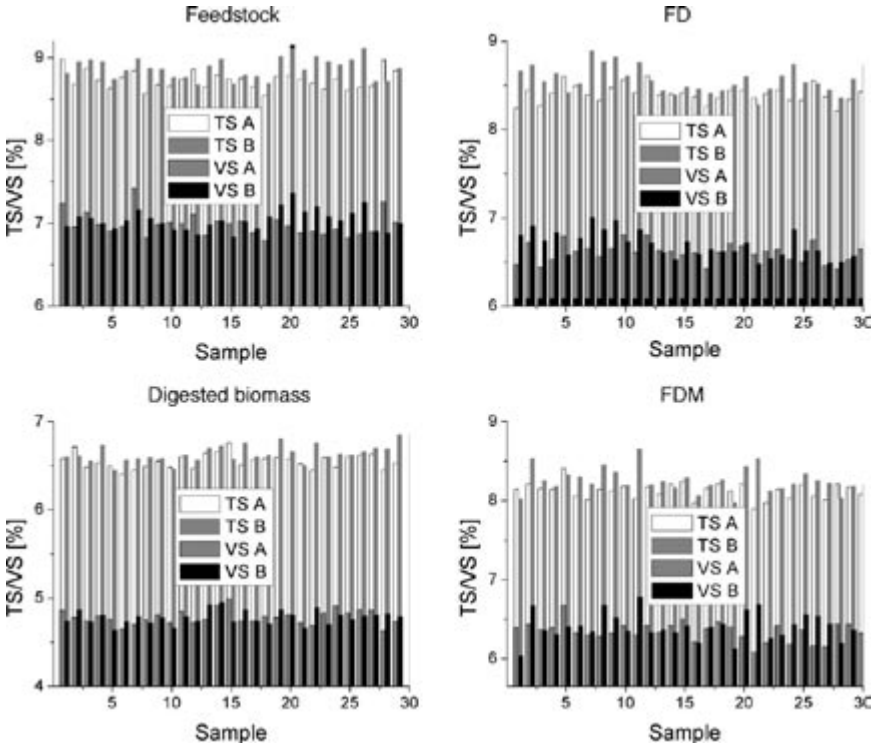


Fig. 8. Compositional differences between A and B vial splits (TS and VS) for four types of materials, typical in biogas production: feedstock, digested biomass, FD and FDM. Note *truncated* concentration axes, visually augmenting A/B differences. Note a clear systematic difference in TS contents for B samples for all four compositions, while this tendency is missing for VS.

Fig. 9. Equipment used in evaluation of vial-to-Petri dish extraction. Left: "pseudo COLIWASA" (operated manually by thumb-pipetting). Right: cut, standard automated pipette. Note that both these alternatives allow for sampling of "sticky" material following [6].

fraction prone to segregation. Splitting into A and B subsamples by fractional pouring in the current study will consequently be slightly biased. Even the extreme homogenization efforts employed in this study have not been able to do away entirely with this tendency.

2.6. Multivariate calibration implications

Since the consecutive 30-mL samples from the fermentation process resulted in relatively minor (but not negligible) differences between the A and B subsamples, some mismatch between reference analysis and NIR/IA was unavoidable. The net effect of this is to broaden the imprecision associated with the final prediction models; everything else equal, a mismatch between the X- and Y-samples will inflate RMSEP levels. But because this feasibility study is aimed at showing the prospects of PAT methodologies on heterogenous systems, including inherent sampling problems the bias is accepted here (it is far smaller than what is usual from sampling from this type of heterogeneous system). This bias can play a didactic role as to the pervasiveness of *incorrect sampling errors* at even the smallest scales in the total sampling-and-analysis process chain.

2.6.1. Vial extraction method

Another subsampling test was carried out to discriminate between three alternative methods for sampling (extraction) from a vial to a Petri dish. Two of the above mixtures were used again:

- DM (90% digested biomass and 10% maize silage).
- FDM (10% feedstock, 88% digested biomass and 2% maize silage).

The following three extraction techniques were investigated:

- "Pseudo Coliwasa", a plastic tube made from the cut end of an automatic pipette, which allowed large(r) particles to be sampled manually.
- Standard automatic pipette with cut end, also allowing large particles to pass.
- Incremental fractional pouring directly from the vial, with continuous shaking to minimize segregation effects (as described above).

The COLIWASA (COlumn LIquid WAter SAmpler), described in detail in [7], is supposed to allow extraction of a virtually undisturbed column through the entire thickness of a possibly stratified medium. This is an especially appealing approach for segregated materials, since such a column by definition is the only which can be fully representative, ibid. Pitard [6] also describes weakly conical inlet tube configurations, which allows effective sampling of "sticky materials" (within "reasonable" limits), as is illustrated in Fig. 9.

All methods were tested with continued agitation of the super-nascent suspensions. These alternative extraction techniques from vial-to-Petri dish are compared in Fig. 10 for the two prepared mixtures DM and FDM, again using the relative standard deviation. RSD is here a measure of the reproducibility of the specific extraction method, based on 10 replicate extractions.

From Fig. 10, all three methods would appear close to equal in extraction performance, with a grand average RSD of some 4.5%. None of the methods are very much different from one another over both materials. This result was slightly surprising, as the COLIWASA approach has otherwise shown great potential in many other macroscopic application scales. The present application is distinctly on the mini-scale however. Fractional pouring, showing a minimum RSD for DM and intermediate for FDM, was chosen as the extraction technique used for analysis by NIR and image analysis below.

3. Analysis

3.1. NIR measurements—PAT

The specific reflectivity and absorptivity of the complex AD organic materials makes it ideal for analysis by NIR spectroscopy. NIR is also advantageous for on-line characterization due to its often *supposed* minimal or no sample

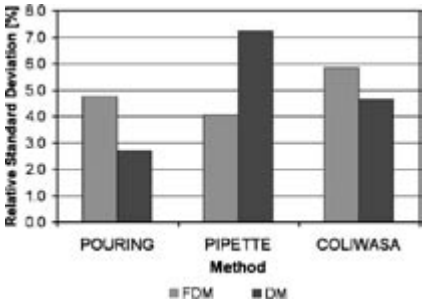


Fig. 10. RSD comparison of reproducibility of alternative vial extraction techniques, based on two different mixtures (FDM, DM). Average results, based on 10 replicate extractions for each technique and biomass material.

Table 1
NIR prediction evaluations for TS, VS and COD; identical 11-segment X-val

NIR	TS	VS	COD
Slope	0.87	0.90	0.77
R^2	0.81	0.91	0.77
Outliers	9	9	7
# Comp	4	3	3

pretreatment and rapid analysis/sample prediction in standard PAT operations [21–26]—which very often pays no attention to the gamut of sampling representativity issues however.

NIR measurements were carried out on the B subsamples (see Fig. 5) using a NIR Systems 5000 spectrophotometer equipped with a mirror transflexive probe head on an optical fiber (scanning range between 1100 and 2500 nm). The samples were homogenized prior to measurements by vigorous shaking. Having completed the NIR measurements, the same samples were subsequently subjected to image analysis.

3.1.1. Results—NIR

NIR spectra were modeled using PLS-1 with segmented cross validation for all three Y-variables: TS, VS and COD; Table 1 summarizes these results. The final 44 measurements (objects) were divided into 11 segments, selected from a sorted listing of ascending TS/VS/COD values. For the purpose of comparison *between* alternative prediction models, it is acceptable, indeed optimal, to use an appropriate segmented cross-validation scheme—contrary to the demands for absolute prediction performance assessment, in which the use of a independent test set is mandatory [27].

Fig. 11 (left) presents the loading weights for the first and the second PLS components from a NIR model of the volatile solids, VS. Stochastic noise was dominant for the highest wavelengths, wherefore variable selection was performed, cutting away variables >2300 nm without loss of information —indeed resulting in improved models; there were a relatively high fraction of outliers.

Final assessment of any PLS model rests with the "predicted vs. measured" plot, or rather with the attendant statistics, pertaining to a fitted regression model. A slope of 0.90 or higher is highly acceptable for the present strongly heterogeneous system. A RMSEP of 0.23 signifies a quite satisfactory prediction precision. Based on the mean VS (4.7%), this RMSEP corresponds to a prediction error ~10% (±2 RMSEP). Another way to express prediction precision is by the squared correlation coefficient. For the NIR VS model, r^2 scores as high as 0.91. These statistics attest a satisfactory potential for predicting VS directly from (*properly sampled*) at-line samples.

Table 1 shows similar statistics for all three AD process parameters, TS, VS and COD. While TS does not achieve as quite good prediction assessment statistics as do VS, they are still acceptable, but COD cannot be modeled with any degree of satisfaction. COD is a measure of the chemical oxygen demand, which includes all compounds in the sample, including water-soluble molecules a.o. It is not overly surprising that no correlated signals appear in the NIR spectra. This can partly also be explained by the particularly high analytical uncertainties for COD compared to TS and VS (see Fig. 6).

3.2. Image analysis

An at-line image of the surface of a specially prepared version of the "B" samples is acquired by an appropriate imaging system in the laboratory, which has been described extensively in [18,19]. This type of imagery is isotropic for which reason the AMT transform can be put into use [14–19].

The angle measure technique (AMT) is a preprocessing technique used for characterizing the complexity of one- and two-dimensional technological data series [14–18]. Image analysis (IA), combined with AMT and PLS-R, is a powerful technique for qualitative and quantitative description of visual features, especially surface morphology and texture [15]. This combined imaging technique (IA/AMT) has mostly been used for characterizing solid materials (surface textural features) and bulk materials (particle sizes, shapes a.o.) [14–19]. The current study takes this technique into a new domain, into characterization of the visual appearances of complex slurry textures as presented by Petri dish appearances in a backlighted setting [18,19]. This may have potential interest in pulp-based industries of various kinds (dairy, paper, mineral extraction,

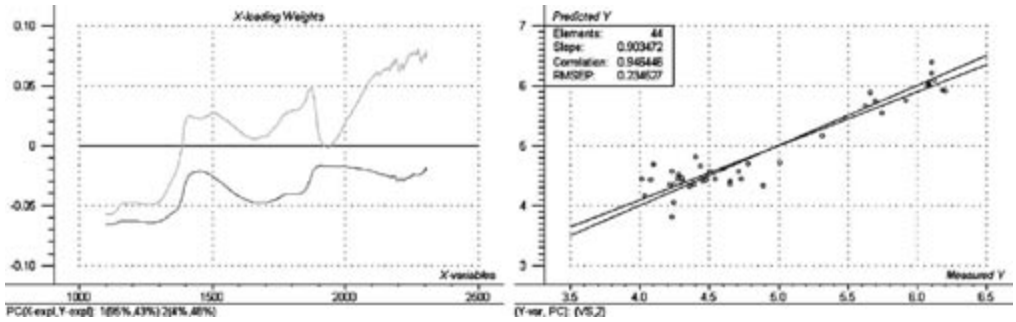


Fig. 11. Left: loading weights for the first and second PLS components for a NIR model of the volatile solids, VS. An indication of the >2300-nm stochastic noise can be appreciated, especially for the second (topmost) component. Right: predicted vs. measured plot for this model, based on 11-segment cross-validation.

chemical process industries), which was the reason to include image analysis/AMT as a possible PAT modality here.

From the outset, it was expected that IA/AMT should be able to model, e.g., TS effectively, while the attendant VS (and COD) might, or might not, also be amendable to PLS modeling and prediction as well. This will depend on the closed array stoichiometrics of the anaerobic digestion system: In a compositionally closed system, it may sometimes be possible to arrive at satisfactory prediction facilities based on *indirect* characterization, necessarily heavily dependent upon realistic test-set validations, etc. [27].

The AMT algorithm was originally developed for one-dimensional data and cannot process a color image as such (in this study R/G/B imagery is used). Each color channel must hence be tested separately, which–in turn–results in three AMT complexity spectra; one for each color channel. The AMT method has been described in full detail in the chemometric literature [15–19].

The imaging system was a QImaging filter-wheel CCD camera equipped with a Nikkor 120 mm F2.8 macro lens, supporting 12-bit grayscale images with a resolution of 1280 × 1024 pixels. A 4″ × 4.88″ backlight platform from Schott Fostec was used to illuminate the material being imaged. Backlighting can be used to create a crisp edge around the present solid material particles in transparent liquids (water). This was thought to be especially useful for image complexity characterization with AMT, because the algorithm would benefit from a clear definition of the transition between matrix (water) and object (dry matter particles).

3.2.1. Replication scheme in image analysis

Each final 30-mL B subsample was distributed into two Petri dishes by fractional pouring as related in detail above. Two non-overlapping images from each Petri dish subsample were acquired. Subsequently, the Petri dish was gently shaken (*not stirred*) and another set of images was acquired. This procedure was repeated four times, resulting in eight replicates for each split subsample—altogether 16 replicate images for each B subsample (Fig. 12). The rationale for this extensive replication relates to the need for evaluating the camera-presentation sampling variance, essentially the "pouring variance", which will always be present for an at-line imaging system. This pouring sampling error cannot be ignored a priori; it has never been adequately assessed [5–7]. In addition, the material is still exceedingly heterogeneous at this stage; the more image replicates, presumably the more representative their average manifestation. This procedure was repeated for each of the 55 samples drawn from the entire 14-day fermentation cycle. The complete imaging procedure is illustrated in Fig. 12.

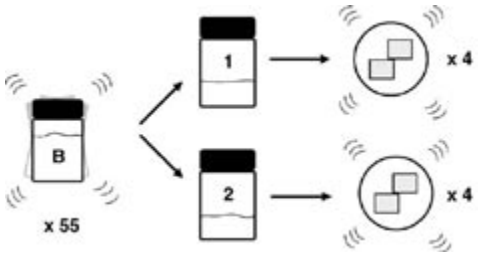


Fig. 12. Replication protocol for imaging procedure. Subsamples (B), also used for NIR analysis, are mixed thoroughly before being distributed into two subsamples (termed 1 and 2) by fractional pouring. Each subsample is poured into a Petri dish and subjected to duplicate imaging. This is followed by gentle shaking of the Petri dish four times with duplicate images after every redistribution of the sample. Altogether this results in an extensive 2 × 8 replicate image series for each B sample.

3.2.2. Results—IA

All images were processed using AMT using both MA as well as MDY complexity spectra [15–19]. Subsequently, all images were subjected to a correction procedure (gamma correction and grey-level enhancement) to optimize image contrast in order to obtain the best possible AMT definition. All images were treated identically. These images were all transformed into AMT complexity spectra and were subsequently modeled using PLS-1 with *identical* segmented cross-validation procedures as for the validation of the NIR models.

The loading weights from this PLS model reflect the wavelengths, which are most correlated with VS. The MA spectra (left half of MA + MDY) reveal that neighbourhoods up to a distance of some 35–85 pixels carry the main correlated information. This corresponds well with the pixel size/resolution of typical dry matter particles in the original images. The main contributions to the prediction model come from the first and second PLS components (92/58 and 6/15% of the total X/Y-variances modeled, respectively).

The prediction assessment in Fig. 13 shows only a slightly lower prediction precision ($r^2 = 0.88$) for TS than for VS. The accuracy statistic slope is equal to 0.86, which again is fully acceptable for the complex system involved. Because of essential interconnections, the relative RMSEP-level is also closely comparable.

Fig. 14 shows the same predicted vs. measured relationship for the VS model as in Fig. 13, also showing models based on all 16 individual replicates, a model based on averaging of the image duplicates and a model based on averaging both image duplicates and shaking replicates. Fig. 14 very clearly attests to the enormous advantage of the averaging operator in the presence of the significant variability associated with the imaging sampling process.

The most important feature in Fig. 14 is the enormous sampling error associated with the original pouring into a Petri dish, viz. the extremely large individual differences in the 16 replicates for each split (upper left panel). This is a poignant reminder that sampling errors (sampling variability) can, and will, crop out at any scale if not recognized; from the primary field situation to the last instance of pouring some 15, apparently innocent, milliliters into a Petri dish (very often considered totally "insignificant" in the greater order of things). The lesson is clear though. If the attending sampling variability is not known, there is every chance that unrecognized errors of

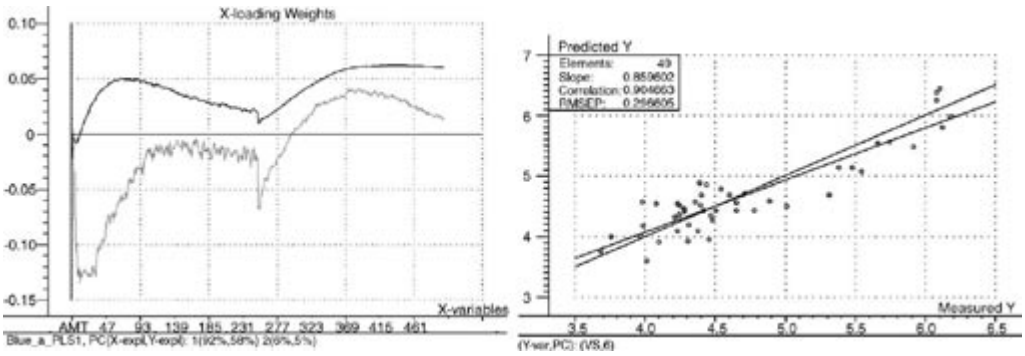


Fig. 13. Evaluation of image analysis PLS-1 model for VS. Left: loading weights for the first and second components. No variable selection was needed for AMT. The model is based on corrected images (described in text) from the blue channel. Right: predicted vs. measured plot for the AMT model of VS, based on a grand average of all 16 image-replicates (Fig. 14 shows the same results *without* averaging).

highly detrimental magnitudes may easily corrupt the reliability of the analytical results. It is only through the diligent, extremely careful, replication/averaging scheme employed in the present study that any image analysis/AMT prediction model could be obtained at all (lower right panel).

Table 2 summarizes averaged results from the whole set of image analysis calibrations, based on the red, green or blue image channel, respectively.

From Table 2, it can be seen that acceptable prediction models are possible for both TS and VS for all three channels. In

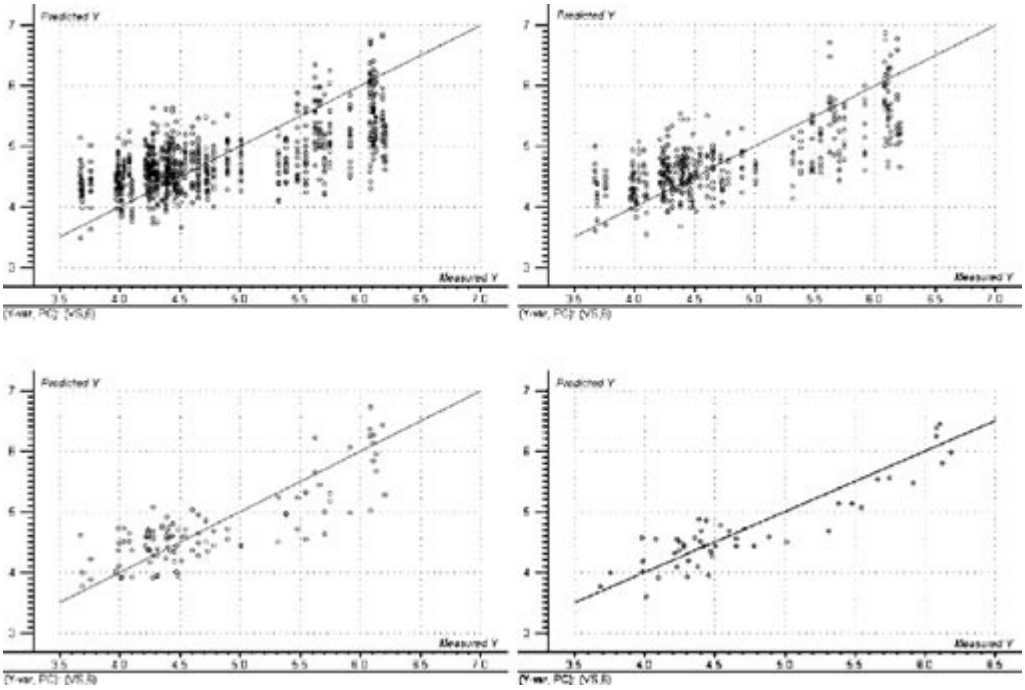


Fig. 14. Averaging effect on the "predicted vs. measured plot" for the PLS-1 model of IA/AMT for VS. Top left: model without averaging (880 images). Top right: model with averaging over image duplicates (440 images). Bottom left: averaging of image duplicates and shaking replicates (110 images). Bottom right: final model based on all three levels of averaging.

Table 2
PLS prediction results for TS, VS and COD from IA/AMT on three alternative color channels

IA results		TS			VS			COD		
		Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
Original images	Slope	0.77	0.72	0.84	0.84	0.79	0.80	0.67	0.63	0.71
	r^2	0.72	0.66	0.76	0.81	0.76	0.79	0.64	0.64	0.67
	Outliers	6	5	6	6	6	7	7	6	8
	# Comp	5	5	6	4	3	1	3	3	5
Corrected image	Slope	0.70	0.73	0.78	0.73	0.71	0.86	0.53	0.53	0.51
	r^2	0.59	0.71	0.72	0.66	0.67	0.81	0.50	0.50	0.49
	Outliers	6	8	4	6	4	6	5	6	6
	# Comp	5	5	5	4	5	6	2	2	3

The images were modeled both with and without a contrast enhancement (original vs. corrected images).

general, the models were not improved from the contrast correction however. The blue color channel would appear to give the best prediction models on an overall basis, though the red channel works well also for VS. It is not possible to model COD by any image approach. The computationally intensive image correction should be omitted, since this does not improve the predictability.

4. Discussion and perspectives

4.1. Representative sampling—a must

This study illustrates the importance of a fundamental quantitative heterogeneity characterization for all new materials to be sampled for the first time [8]. Considerable efforts were spent on zooming in on appropriate sampling procedures, from the initial field sampling to the minute sampling for the ultimate analytical volume. A considerable informed diligence with relation to TOS is necessary in order not to oversee any link in the sampling vs. scale representativity hierarchy. The innate heterogeneity of the material systems were highly significant, which made for an exceptionally clear case of the difficulties involved in representative sampling and mass reduction. There exist a plethora of identical complex, heterogeneous materials in process industrial sectors to which the chemometrics community currently introduces the PAT concept. Reliable, representative sampling must play a central role in this context, in the critical reference calibration context as well as concerning appropriate sensor localizations and representativity in accordance with the entire volume or cross-section of reactors or pipelines.

4.2. NIR and image analysis perspectives

This feasibility study tested a parallel NIR and AMT methodology for characterizing specific complex, heterogeneous bioslurries (sampling based on proper TOS principles). Optimization of the individual experimental conditions will necessarily be closely linked to the particular industrial implementations. Many parameters, such as illumination and camera options, sample preparation (especially dilution and sample amount/thickness) and bioslurry sampling must always be optimized when dealing with image analytical applications, but this will necessarily be related to potential, specific industrial implementations. With problem-dependent optimization, both the NIR and IA prediction models can most likely be somewhat to significantly improved. Below, we discuss general features only.

From related research on newly developed circulation flow-through cells (carried out at Christian-Albrecht's-University, Kiel), termed "Transflexive Embedded Near InfraRed Sensor" (TENIRS), joint work in progress has shown very promising results for the same types of difficult medias. Here, manure slurries and digestate slurries not only gave good models for TS and VS, but also for total-N, ammonium and phosphorous. In this type of circulation configuration, much larger analytical volumes of slurry can be measured in an at-line, or even on-line NIR context. This system is currently for use distinctly as an at-line facility, but by-pass options from primary biogas plant pipelines or reactors can easily be envisaged and are not technically difficult. By-pass solutions must of course also oblige all the necessary TOS criteria.

The challenge is to develop, and especially to maintain, robust calibration models based on NIR measurements, due to the extreme variability of the feedstocks encountered in the industrial biogas sectors originating from a bewildering array of manure types and organic by-products as well as waste-streams from food and pharmaceutical industries. There is an ultimate challenge due to the fact that systems have to be developed, and maintained, at price levels far below what is the case for, e.g., the food manufacturing and pharmaceutical sectors. This is because the bioenergy sector has to produce, indeed make a profit, far below these price levels per units of product [3,4]. This puts very stringent demands as to the level of high-tech equipment that can be contemplated.

4.3. PAT for anaerobic bioconversion processes—perspectives

The present method developments in laboratory-scale studies are currently being followed up by extensive testing of the TENIRS flow-through cell system both for at-line and on-line NIR, augmented by acoustic chemometrics characterizations [28,29], aimed at overcoming the immense difficulties for TOS-correct reactor sampling, which is virtually impossible. This will be done by introducing a *recurrent loop* sampling concept, essentially transforming the three-

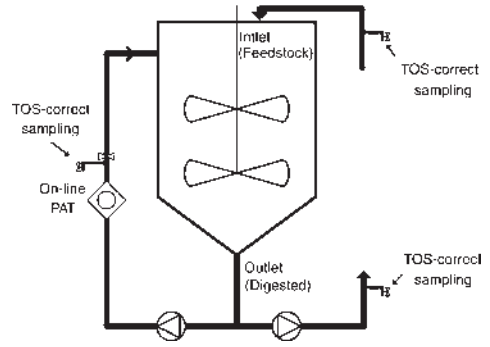


Fig. 15. On-line PAT measurement and TOS-sampling concept in a full-scale AD plant context.

dimensional bioreactor sampling issue into a one-dimensional pipeline sampling situation, which in comparison is fairly easy [8,9]. The prospects here also lie in reducing and minimizing the time consumption and cost expenditures for the necessary analysis.

PAT, process analytical technologies is here, as in pharmaceutical and food production contexts, directly aimed at improved process control by applying suitable sensor technologies for on-line measurement of biomass feedstock composition and of important intermediary and end products of the anaerobic digestion processes. Appropriate sensors and physical sampling modalities shall be deployed at critical control points: (1) at raw materials and feeding inlet/outlet localizations, (2) in recurrent loops positioned strategically in the process and (3) at post-treatment localizations. These developments will contribute towards the goal of an integrated process and product monitoring management and regulation system, orders of magnitude more reliable than the current state of art.

By introducing the type of *reliable* on-line PAT sensor systems delineated in Fig. 15 in addition to the at-line analysis approaches surveyed here, it will be possible to deliver fully comprehensive monitoring data for improved control throughout the AD processes. This will make it possible to adjust process parameters on a real-time basis based only on characterization of feedstock variations coupled with a library of relevant PLS predictor models. This concept transgresses the narrow AD scope and will address the much more general field of anaerobic bioconversion (ABC) process regimen without loss of significant carry-over potential because of the consciously selected adverse general dry-matter/suspension characteristics in the present study.

5. Conclusions

A down-scaled laboratory bioreactor was used to survey appropriate sampling for two selected potential process analytical technologies, NIR and image analysis/AMT, for highly heterogeneous anaerobic digestion processes in large-scale biogas plants.

Table 3
NIR modeling/prediction results

NIR	NIR		
	TS	VS	COD
Slope	0.87	0.90	0.77
R^2	0.81	0.91	0.77

All critical sampling issues for at-line analysis across the relevant scale domains, primary sampling, reactor and pipeline sampling, secondary laboratory sample mass reduction, tertiary mass reduction/splitting in the laboratory trials and quaternary NIR/IA analytical sample preparation, were analyzed in their proper TOS contexts. Appropriate representative sampling procedures and equipment were developed and implemented, commensurate with the intrinsic heterogeneity characteristics of the materials involved, as exemplified by multicomponent organic feedstocks and digested bioslurries. An attempt was made to illuminate the critical role of proper quantitative heterogeneity characterization before design of representative sampling protocols, without which chemical analysis and data analysis will always be in potential jeopardy.

The quantitative sampling_cum_analysis error levels involved were empirically estimated in order to demonstrate the effectiveness of the sampling measures adopted. Each of 55 laboratory reactor samples following the normal and enhanced (maize silage spiking) biogas process development was split into two subsamples by use of optimised fractional pouring. One of these subsample series was subjected to NIR followed by image analysis/AMT, while the other was used for chemical reference analysis of total solids (TS), volatile solids (VS) and chemical oxygen demand (COD), followed by multivariate calibration (PLS regression).

Acceptable prediction models could be obtained for TS and VS for both the NIR and IA/AMT techniques, as summarized in Tables 3 and 4. These two important routine AD parameters show promising prospects for at-line/on-line prediction of complex bioslurries, while it was not possible to model COD satisfactory due to very large analytical uncertainties in the routine COD analysis standards. This latter aspect both can and will have to be improved.

There would appear to be a significant advantage in using a larger primary camera footprint—in the form of many image replicates in order to suppress the specific image sampling errors encountered, dramatically displayed by Fig. 14. While we here were able to suppress this error satisfactory (by extensive replication), if image analysis indeed is to be a viable PAT candidate for this type of material system, it must be amendable

Table 4
IA/AMT prediction results (blue channel)

IA	Original image—blue color channel		
	TS	VS	COD
Slope	0.84	0.80	0.71
R^2	0.76	0.79	0.67

to extensive automated sample preparation (presentation to the camera)—which is tantamount to a far too laborious, and even more too expensive, task for full-scale facilities at present.

Thus, the NIR approach comes out as a clear winner in this context, especially as the prospects for full-scale implementation already has been amply proven for many other media in very many other process analytical contexts. While the critical (although often too easily dismissed) sampling issues have been treated in full detail here, the attending specific probe-localization sampling error issues have been only hinted at so far however; they will be addressed in full elsewhere.

Acknowledgements

Baseline sampling and NIR studies were carried out by an enthusiastic complement of our Romanian and Polish 8th semester MSc (Eng.) students at Aalborg University Esbjerg: Monika Antczak, Liliana Jensen and Adam Stachowski, to whom we extend our sincere thanks for their hard work. They and their fellow students achieved a good learning process about TOS and NIR.

References

[1] Å. Nordberg, M. Hansson, I. Sundh, E. Nordkvist, H. Carlsson, B. Mathisen, Monitoring of biogas using electronic gas sensors and near-infrared spectroscopy (NIR), Water Science and Technology 41 (2000) 1–8.
[2] B.K. Ahring, M. Sandberg, I. Angelidaki, Volatile fatty acids as indicators of process imbalance in anaerobic digestors, Applied Microbiology and Biotechnology 43 (1995) 559–565.
[3] K.H. Gregersen, Centralised Biogas Plants—Integrated Energy Production, Waste Treatment and Nutrient Redistribution Facilities, 2005 (11 Feb.). http://websrv5.sdu.dk/bio/pdf/centra.pdf.
[4] T. Al Seadi, Danish Centralised Biogas Plants—Plant Descriptions, 2005 (11 Feb.). http://websrv5.sdu.dk/bio/pdf/rap2.pdf.
[5] P. Gy, Sampling for Analytical Purposes, John Wiley and Sons, 1996.
[6] F.F. Pitard, Pierre Gy's Sampling Theory and Sampling Practice, 2 edn. CRC Press Ltd., 1993.
[7] P.L. Smith, A Primer for Sampling Solids, Liquids, and Gases, Siam, 2001.
[8] L. Petersen, P. Minkkinen, K.H. Esbensen, Representative sampling for reliable data analysis: theory of sampling, Chemometrics and Intelligent Laboratory Systems 77 (2005) 261–277.
[9] L. Petersen, C.K. Dahl, K.H. Esbensen, Representative mass reduction in sampling—a critical survey of techniques and hardware, Chemometrics and Intelligent Laboratory Systems 74 (2004) 95–114.
[10] M. Hansson, Å. Nordberg, B. Mathisen, On-line monitoring during anaerobic treatment of municipal solid waste, Water Science and Technology 48 (2003) 9–13.
[11] P.F. Pind, I. Angelidaki, B.K. Ahring, A new VFA sensor technique for anaerobic reactor system, Biotechnology and Bioengineering 82 (2003) 54–61.
[12] A. Legin, Y. Rudnitskaya, C. Vlasov, E. Di Natale, A. Mazzone, Application of electronic tongue for quality and quantitative analysis of complex liquid media, Sensors and Actuators. B, Chemical 65 (2000) 232–234.
[13] A. Legin, D. Kirsanov, B. Seleznev, A. Rudnitskaya, J.B. Holm-Nielsen, K.H. Esbensen, Y. Vlasov, Analysis of feedstuff from biogas production plant by the electronic tongue. The 10th International Symposium on Olfaction and Electronic Nose, ISBN 9984-18-222-3.
[14] R. Andrle, The angle measure technique: a new method for characterizing the complexity of geomorphic lines, Mathematical Geology 26 (1994) 83–97.
[15] K.H. Esbensen, K.H. Hjelmen, K. Kvaal, The AMT approach in chemometrics—first forays, Journal of Chemometrics 10 (1996) 569–590.
[16] J. Huang, K.H. Esbensen, Applications of angle measure technique (AMT) in image analysis: Part I. A new methodology for in situ powder characterization, Chemometrics and Intelligent Laboratory Systems 54 (2000) 1–19.
[17] J. Huang, K.H. Esbensen, Applications of AMT (angle measure technique) in image analysis: Part II. Prediction of powder functional properties and mixing components using multivariate AMT regression (MAR), Chemometrics and Intelligent Laboratory Systems 57 (2001) 37–56.
[18] C.K., Dahl, L., Petersen, K.H., Esbensen, Automatic image analytical characterization of powder agglomeration by AMT-regression (angle measure technique): Part 1. Development and validation of method, Journal of Powder Technology (submitted for publication).
[19] C.K. Dahl, K.H. Esbensen, Image analytical determination of particle size distribution for bulk solids, Chemometrics and Intelligent Laboratory Systems (submitted for publication).
[20] D.A. Skoog, J.F. Holler, D.M. West, Fundamentals of Analytical Chemistry, 7 edn.Harcourt Brace College Publishers, 2005.
[21] D.F. Malley, L. Yesmin, R.G. Eilers, Rapid analysis of Hog manure and manure amended soils using near-infrared spectroscopy, Soil Science Society of America Journal 66 (2002) 1677–1686.
[22] A. Millmier, J. Lorimor, C.H. Fulhage, J. Hattey, H. Zhang, Near-infrared sensing of manure nutrients, American Society of Agricultural Engineers 43 (2000) 903–908.
[23] Y. Zhang, Z. Zhang, N. Sugiura, T. Maekawa, Monitoring of methanogen density using near-infrared spectroscopy, Biomass & Bioenergy 22 (2002) 489–495.
[24] P. Jørgensen, J.G. Pedersen, E.P. Jensen, K.H. Esbensen, On-line batch fermentation process monitoring (NIR)—introducing 'biological process time', Journal of Chemometrics 18 (2004) 81–91.
[25] Y. Li, C.W. Brown, F.-M. Sun, J.W. McCrady, R.W. Traxler, Non-invasive fermentation analysis using an artificial neural network algorithm for processing near infrared spectra, Journal of Near Infrared Spectroscopy 7 (1999) 101–108.
[26] S. Vaidyanathan, S.A. Arnold, L. Matheson, P. Mohan, B. McNeil, L.M. Harvey, Assessment of near-infrared spectral information for rapid monitoring of bioprocess quality, Biotechnology and Bioengineering 74 (2001) 376–388.
[27] K.H. Esbensen, Multivariate Data Analysis—In Practice, 5th edn., CAMO ASA, 2002.
[28] K.H. Esbensen, M. Halstensen, T.T. Lied, A. Saudland, J. Svalestuen, S. de Silva, B. Hope, Acoustic chemometrics—from noise to information, Chemometrics and Intelligent Laboratory Systems 44 (1998) 61–76.
[29] K.H. Esbensen, B. Hope, T.T. Lied, M. Halstensen, T. Gravermoen, K. Sundberg, Acoustic chemometrics for fluid flow quantifications: II. A small constriction will go a long way, Journal of Chemometrics 13 (1999) 209–236.

Paper III

Casper K. Dahl*, Pentti Minkkinen, and Kim H. Esbensen

"Image Analytical Monitoring of Paper Quality - a Feasibility Study"

TAPPI JOURNAL
Volume 5, Issue 11, pp. 18-24, 2006.

Image analytical monitoring of paper quality – a feasibility study

CASPER K. DAHL, PENTTI MINKKINEN, AND KIM H. ESBENSEN

ABSTRACT: We evaluated 14 selected paper quality parameters using new image analytical techniques for characterization and monitoring of paper quality (multivariate AMT regression). We tested the technique on six major paper types. Of the parameters tested, 13 could be modeled with only minor optimization of the initial experimental image recording parameters. Although this analysis was performed in the laboratory with static sheets of paper, this new "remote sensing" image technique shows promising application potential for on-line product monitoring of the specified physical qualities.

Application: The research presented here should lead to an on-line method for evaluating paper quality during production.

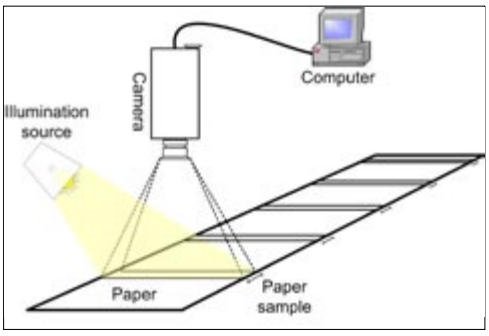
Off-line testing (lab testing) of finished sheet characteristics is still the most common practice to ensure that process and quality specifications are being met. These methods are often too slow for on-line process control strategies, and also have an inherent sampling problem: physical samples are usually cut only from the end part of a master roll.

Commercial technologies for on-line measurement of finished paper characteristics have proven to be limited in both speed and range of tests. Image analytical techniques could become a very useful tool for on-line measurement of paper properties. Image processing technology is used in the industry today, but it cannot be used to control all quality parameters alone and must often be combined with other methods. The complexity of integrating these other methods into the image analysis puts high demands on the reliability and standardization of information, both of which are necessary to develop robust process and product control systems.

An alternative approach, combining all necessary characterization methods into one, may some day come into existence, presumably based on optical (visual [VIS] and near-infrared [NIR]) and physical-radiometric (neutron activation analysis [NAA]) techniques, amalgamated into one.

This paper investigates how far optical paper quality characterization can reach on its own merits. The technique used is a combination of digital image acquisition, followed by the angle measure technique (AMT) [1-3,6-11,13]. We designed an image analytical system to simulate automatic monitoring of as large a set of paper quality parameters as possible. The system relies on problem-dependent oblique unilateral illumination together with multivariate calibration of transformed images.

In this study, isotropic images of low-angle oblique illumination enhanced paper surface texture are transformed into so-called AMT complexity spectra [3,7,8], which are subjected to multivariate calibration using partial least squares (PLS) regression [4,12].



1. Typical image analytical system setup.

IMAGING

Figure 1 shows a proposed on-line (on-machine) image system setup. Recorded images of the paper surface provide information about its texture and other physical properties. The critical vision system parameters include the camera performance specifications, which should be adequate for the task required (though not detailed in this study), and the illumination conditions. It is often also important to carry out problem-dependent image processing, which in the present case included enhancement of selected surface features (i.e., texture). By using low- to very low-angle (oblique) unilateral illumination, complex shadows are created, which in turn enhance the feature extraction function of the AMT. With the paper lane suitably illuminated, images are acquired automatically with a camera connected directly to a computer by FireWire or similar fast computer interface connection. A fast connection is needed to capture high resolution, high quality, real-time images of the moving paper. Although the paper speed may exceed 1000 meters/min., only a small fraction of the total lot (length) of

paper needs to be sampled (strip-imaged) and analyzed, typically using a sampling rate of 1:100 or less.

Image analysis can be used for both quantitative and qualitative measurement. The images of the virtual paper strips "cut out" by the camera system should be representative of the entire paper master roll. Two key issues must be addressed for this to be valid:

1. The sampling rate must be adequate along the longitudinal length-dimension of the master roll (the total width of the paper lane is supposedly captured).
2. The degree to which the upper surface of the paper lane will function in relation to the properties present throughout the total thickness of the paper.

The Theory of Sampling (TOS) [6,17,18] is a comprehensive theory for obtaining structurally correct samples (samples that represent the average composition of the lot material). TOS (1-dimensional sampling) uses variogram analysis [6,15,16] methodology to ensure that the fractions (strip areas) sampled by the camera are representative of the whole 1-dimensional production lot or paper master roll, with a specified sampling error. This technique can be used for sampling on-line measurements and is well described by Petersen et. al. [17]. Variographic techniques and PLS-regression have recently been used to study pulp bleaching processes [7,8].

The sample strip longitudinal dimension will be determined on a problem-dependent application basis. Because nonphysical, remote sensing (imaging) is being used, we can analyze as many and as large samples as the variogram analysis demands. Only in the calibration process of the prediction model is it necessary to extract the physical samples from which the reference values for the paper quality parameters are determined. The sampled strip-images are then subjected to feature extraction by the AMT transform. This results in a "complexity-spectra," x, which can be subjected to multivariate calibration using PLS-regression with paper quality parameters, y.

PLS-REGRESSION

PLS-regression is a bilinear projection method, where samples expressed in the x-space (defined by the original x-variables) are projected onto new axes (termed PLS-components), which define a new space constituted by a few latent variables only. The projected locations of the samples in this PLS-component space are termed scores. These PLS-projections are dynamically inter-correlated with the y-space (one or more y-variables, see further below), in such a fashion that the y-space data structure is used actively to find only the relevant information in x (i.e., the most correlated). This optimizes the covariance between the two spaces and often leads to very simple models consisting only of a few latent PLS-components in each space (often as low as 2, 3, or 4).

Two different versions of PLS exist: PLS1 and PLS2. In PLS1, one y-variable is modelled only. PLS2 models on all y-variables simultaneously. PLS1 is optimal for predicting a single y-variable, while PLS2 is used for overview modelling only, because of the advantage of modelling all correlations between all x- and y-variables simultaneously. Predictions are always based on PLS1 exclusively. PLS-regression can generally be schematized in the following way:

$$X = \sum_{A} T \cdot P^T + E$$

$$Y = \sum_{A} U \cdot Q^T + F$$

T: *x*-scores
U: *y*-scores
P: *x*-loadings; W: *x*-loading weights (*"effective" loadings*, see [4,12])
Q: *y*-loadings
E: Error matrix for the *x*-space
F: Error matrix for the *y*-space
A: The number of PLS components
T: Transpose operator

Both the *x*- and *y*-space are modelled by summarized product of scores (T/U) and loadings (P/Q) and error matrixes (E/F). The product of the scores and loadings (loadings are the coordinates of the new axes in the original space, which is a reflection of how well each variable is taken into account by the PLS-components) can then be seen as a compact model of the covariance data-structure, while the error matrices contain the remaining, noncorrelated information.

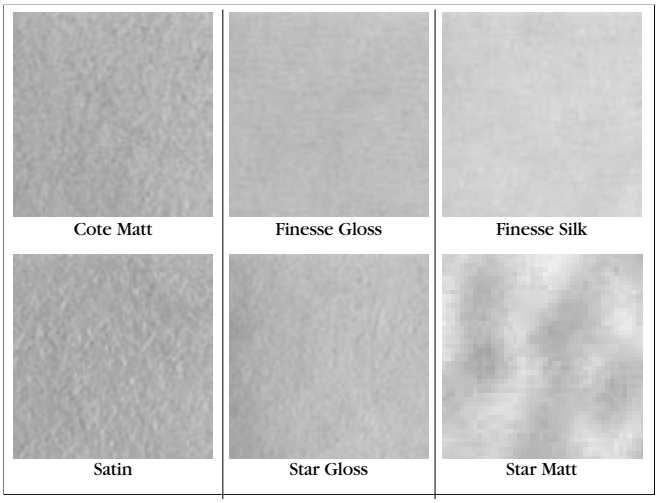
MATERIALS

Six different paper types from UPM-Kymmene Corporation (Finland) were studied (**Fig. 2**). Each represents a printing grade with different degrees of finish. The Cote matte grade represents a coated sheet, which must have characteristics of functionality, consistency, and good print results on both gloss and matte surfaces (Cote matte paper is imaged in this study). The Finesse grade is a wood-free, high brightness and opacity sheet, with differing finishes, from gloss to matte (silk). The different finishes provide a good basis for image and text reproduction. The Satin sheet allows print to stand out on both matte and glossy surfaces (only matte was imaged). Star paper is double-coated and used for high-quality print reproduction. It exhibits high brightness, gloss, and smoothness. Both matte and gloss finish were sampled.

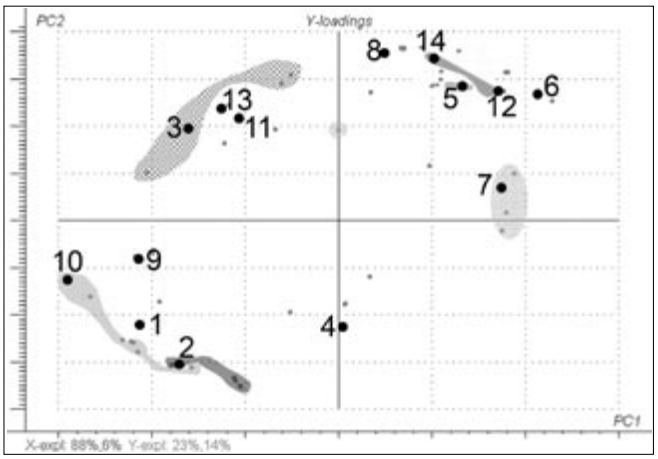
We measured 72 paper quality parameters for each paper type. A PLS2 model [4,12] was developed between these parameters and the AMT-spectra to observe relationships (**Fig. 3**).

The *y*-loadings (Q) for the first PLS component are plotted against the y-loadings for the second PLS component. Similarly behaving variables are grouped tightly, which indicates a strong internal correlation [4,12]. For example, six different measures for brightness have been measured and modeled and these are grouped as contiguous gray-shaded area number 5 in Fig. 3.

The 14 *y*-variables numbered in Fig. 3 and named in **Table I**) that span the depicted loading space were selected for further individual analysis with PLS1. This reduced set represents key variables important to paper quality. Also, the vari-



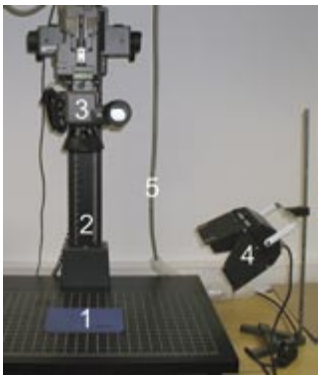
2. Visual appearance of the six imaged paper types.



3. PLS analysis graph (PLS2 loading weights) showing groupings of correlated variables (see Table I for variable groups). The two PLS-components (plotted as the x- and y-axis, respectively) use 88 + 6 = 94% of the x-variance to explain 23 + 14 = 37% of the y-variance.

1. Ash	5. Brightness	10. Roughness
2. b*C2	6. Density	11. Tear strength
3. Bending resistance	7. Gloss	12. Tensile strength
4. Bending resistance ratio	8. Grammage	13. Thickness
	9. Opacity	14. Whiteness

I. Paper quality parameters in this study.



4. Image analytical setup: 1. paper sample; 2. camera stand (elevator); 3. camera; 4. low angle illumination source; 5. cable to computer.

ables loading space spans a summary of their individual PLS1 modeling and prediction abilities, and as such will provide a condensed, yet reliable view on the current feasibility study objectives.

EXPERIMENTAL SETUP

The optical system was composed of a QImaging black-and-white CCD camera with a Nikkor 120 mm F2.8 macro lens, supporting 12-bit grayscale images with a resolution of 1280 x 1024 pixels, all mounted directly to a Kaiser rePro photo stand (**Fig. 4**). The entire camera stand was placed in a dark room to eliminate stray external light. This prevented inaccurate color reproduction and extensive natural light variations, which among other variables would negatively affect modeling. Maintaining as constant a light-setting as possible minimizes modeling error in an on-line industrial application.

The paper was placed directly underneath the camera and illuminated with a FalconEyes LHS-500 lamp holder fitted with a 150W bulb. The holder was vertically adjustable to provide low to very-low oblique illumination. We used this combination to enhance the appearance of the sheets' surface features. The recorded images were transferred to a personal computer via a FireWire connection.

We used standard experimental design approaches to optimize the imaging conditions. For the six paper types, surface feature (texture) images were sampled by changing both illumination angle and image size (resolution). Four different experimental setups were selected (series 1-4, **Table II**).

	Paper footprint size: 1.7 x 1.35 cm (low resolution)	Paper footprint size: 1.0 x 0.8 cm (high resolution)
Illumination angle: 21°	1	4
Illumination angle: 17.5°	2	3

II. Experimental overview. The four experimental series are numbered as indicated.

We used illumination angles of 17.5° and 21° to image the same sample of Star matte paper. Significant improvements in the surface roughness and compositional texture definitions were observed in the images using the lower angle (**Fig. 5**).

5. Visual appearance with 17.5° (left) and 21° (right) illumination angles on Star matte paper.

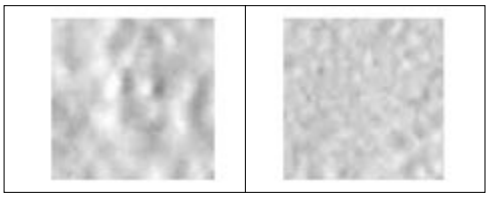
Two footprint sizes (absolute size of the camera field-of-view) were used to determine whether a few large image cutouts or many small(er) samplings would be optimal. Theoretically, and in industrial applications, the entire paper lane cross-section should be imaged to obtain the most representative sample characteristics. This paper investigates only a first estimate of the usable sampling rate relationships needed (see the study by Mortensen and Esbensen [13]).

We investigated footprint sizes of 1.7 x 1.35 cm (low resolution) and 1.0 x 0.8 cm (high resolution). These footprint sizes roughly correspond to a realistic industrial sampling rate. **Figure 6** shows that this difference in size is markedly visible-the texture definition is of course changed as a function of the pixel-size scale resolved.

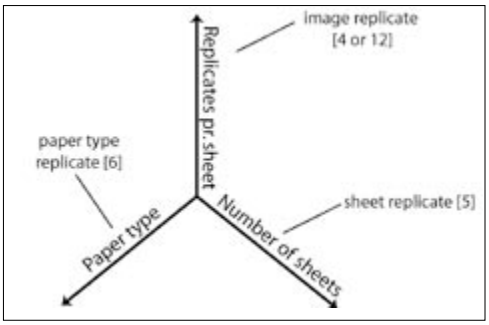
To minimize measurement errors, we used a number of replicates between and within the paper type sheets (**Fig. 7**).

Within each paper sheet, two sets of replicates were used. For the low-resolution settings, four image replicates for each sheet were acquired, while 12 were acquired for the high-resolution settings (**Fig. 8**).

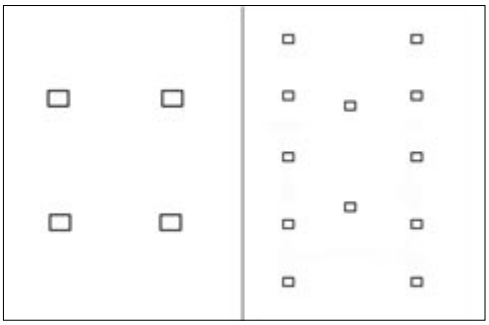
Five sheet replicates of each of the six paper type replicates were subjected to imaging, resulting in 120 images for the larger, low-resolution image samples and 360 images for the smaller, higher resolution image samples.



6. Resolution alternatives studied. Star paper matte finish, 17.5° illumination, 100 x 100 pixels. Pixel size is 8 x 6 µm (high resolution) (left) and 13 x 10 µm (low resolution) (right).



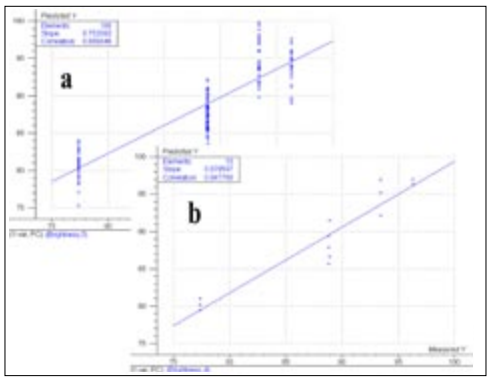
7. Layout of replicate regimen studied.



8. Sheet replicate scheme for low-resolution images (experimental series 1 and 2) (left) and high-resolution images (experimental series 3 and 4) (right).

Feature extraction of the images using the AMT transform yielded 500 x-variables for each image, of 5000 initial points (pixels), A, which corresponded to approximately 0.5% of the total image size.

The overall sampling error and modeling quality were greatly improved by averaging the individual AMT spectra



9. Predicted vs. measured plot for brightness, no averaging (a) and averaging within individual sheets (b). (Cote paper was not included in brightness models due to lack of reference values)

within each sheet before performing the multivariate data analysis. This averaging resulted in 30 AMT spectra for multivariate data analysis (five replicates for each of the six paper types) (**Fig. 9**).

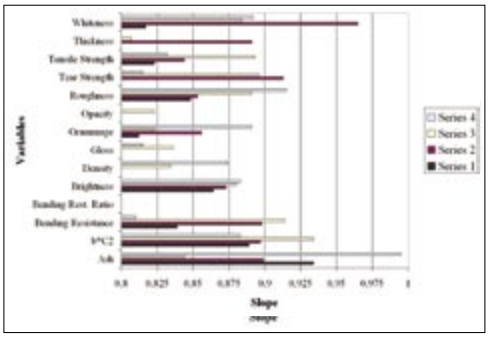
The *x*- and *y*-data sets were normalized to zero mean and each of the 14 selected variables was modeled using PLS1 with test set validation.

PREDICTION OF PAPER QUALITY FUNCTIONAL PROPERTIES

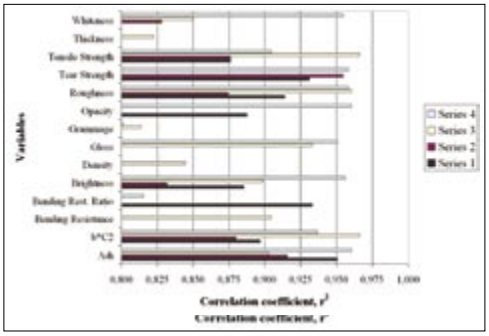
The 14 parameters were modeled using PLS. Either two or three of the five paper sheet replicates (for each paper type) were randomly chosen for testing the prediction models (test set definition). The fitted slope (accuracy) and the r^2 correlation coefficient (precision) were used to describe the quality of the established multivariate prediction models for the paper quality parameters (**Figs. 10** and **11**).

Results show that series 3 and 4 (smallest footprint areas, highest resolution) yielded the best prediction quality statistic (r^2), indicating that the physical area being imaged (or the pixel size in μm) is important in determining the precision of this image analytical method. However, the optimal experimental settings do vary, as evident in the parameters thickness, tensile strength, and density.

All 14 paper quality parameters cannot accurately be predicted with only one of the four experimental settings. However, 13 of them could be modeled reasonably well using the third experimental settings. The bending resistance ratio failed to model well with any of the four series. We do not yet fully understand why this physical material characteristic was not predicted and other physical tests were. Even though series 3 was able to predict most parameters, series 2 gave the best models for whiteness and tear strength. With



10. Prediction model slopes, reflecting overall accuracy of prediction.



11. Prediction model correlation coefficient indicating overall precision of prediction.

all 14 slopes (but one) and all correlation coefficients higher than 0.80, the potential for image analytical prediction in these grades of paper are quite promising.

DISCUSSION

In this study we investigated whether paper quality can be predicted from image analysis. The results show that this image analytical technique is capable of predicting many key paper quality parameters. To apply the off-line techniques discussed here into an on-line application, a faster imaging system with sufficient processing speed is needed.

To process the large amounts of data that are collected on-line, modifications must be made to the current algorithms for AMT, which are the limiting factor today. Optimized algorithms are in development [11]. The computational system requirements (and algorithms) are highly dependent upon the image size and image sampling frequency. The computational requirements are even greater if any image preprocessing is to be performed.

To obtain acceptable prediction models, it might be nec-

essary to optimize the system settings both before and after imaging. In this study, a set of illumination angles and camera resolution settings were varied to achieve optimum correlation to sheet properties. Additional experimentation with different illumination angles and sample sizes, or the use of a different problem-dependent illumination source may improve results even further. Existing knowledge on light sources versus AMT performance may also assist in making the optimal equipment, imaging, and data analysis selections [14]. Finally, the set of ideal system parameters may need to be established by trying them under a series of realistic on-line conditions and comparing the results with the use of the techniques discussed in this paper.

Prediction performance may be further improved by preprocessing the raw images with various image analytical filters before data analysis. Small changes in contrast or brightness can greatly affect image features. A previous study [2] showed that some variables could not be modeled at all without preprocessing, while much more satisfactory models were obtained with preprocessing.

All paper quality parameters can roughly be divided into two groups: optical (surface relevant properties) and physical (mechanical). It should be possible to describe almost all optical parameters using the present imaging technique. Some physical parameters, such as bending resistance ratio, appear more difficult to predict, while those correlating with paper texture such as thickness and density, are more accurately modeled. However, correlations might be reduced for paper grades that do not focus on surface properties (such as tissue, linerboard, corrugated medias, etc.), especially if the relevant parameters are not correlated with surface properties.

The current system only investigated a small fraction of the image area, but it provided quite satisfactory first-order models. Larger images (composite images from well-sampled increments over the entire paper lane width) will undoubtedly improve the models [5]. This strategy for improvement in real-world implementations corresponds to increasing the number of sheet replicates (five in the present investigation).

This study investigated 14 different quality variables, although 72 different parameters are often used in the quality control of the paper grades investigated. In Fig. 3, all 72 parameters are modeled simultaneously using PLS2. Most functional groups were modeled and a large part of the variable space was covered by the 14 paper parameters selected here. Most of the remaining variables can most likely be modeled with individual PLS1 models similar to those used here.

CONCLUSION

Two different illumination angles and two different resolutions (image sizes) were tested within a relevant parameter space. The combination of 17.5° angle illumination and high resolution images could predict most parameters. The only variable that was not modeled with acceptable accuracy was the bending resistance ratio.

The following paper quality parameters were modeled successfully with only minor optimization of the initial experimental parameters: ash, b*C2, bending resistance, brightness, density, it can definitely be beneficial to the paper industry.

The next step in developing this technique would be to perform a mill trial.

INSIGHTS FROM THE AUTHORS

A former research project was the forerunner for this project and hence woke my interest to dig deeper into the possibilities of combining image analytical sampling with monitoring of paper and bulk quality.

This work supports previous work done in cooperation with a Finnish paper mill, but it also differs from previous research projects regarding the application of this new imaging technique.

The most difficult aspect of this work was to find a set of relevant parameters, which potentially could characterize 72 process parameters. However, a large PLS2 model did successfully find 13 parameters, which spanned the whole measurement space.

It was expected that the surface images could be correlated to the surface properties such as brightness. However, it was surprising that it was possibly to correlate the surface images with the mechanical bulk properties, such as paper thickness.

Because this cheap technique can be applied to on-line monitoring of paper and bulk quality parameters

Dahl

Esbensen

Minkkinen

Dahl and Esbensen are with ACABS Research Group, Aalborg University Esbjerg, Esbjerg, Denmark; Minkkinen is with the Department of Chemical Technology, Lappeenranta University of Technology, Lappeenranta, Finland; email Dahl at ckd@aue.auc.dk or Esbensen at kes@aue.auc.dk.

gloss, grammage, opacity, roughness, tear strength, tensile strength, thickness, and whiteness.

The present image analytical AMT/PLS technique has shown a promising application potential for monitoring paper quality by being able to model those 13 representative key quality parameters by remote optical sensing.

ACKNOWLEDGEMENTS

We thank Teuvo Leppänen, UPM-Kymmene Corporation, and Marja Salo, UPM Research Center's Paper Laboratory, for kindly supplying the paper samples and the reference analysis necessary for this study. We thank Jim Burger for his valuable instrument, programming, and interfacing competences, adding significantly to ACABS' capabilities when the camera used in this study was procured and implemented.

Submitted: June 15, 2005
Revised: May 23, 2006
Accepted: July 23, 2006

LITERATURE CITED

1. Dahl, C.K., Petersen, L., and Esbensen, K.H., "Automated image analytical characterization of powder agglomeration by AMT-regression," Powder Tech., submitted.
2. Dahl, C.K. and Esbensen, K.H., "Image analytical determination of particle size distribution for natural and industrial bulk solids aggregates," Chemomem. Intell. Lab Syst., submitted.
3. Esbensen, K.H., Hjelmen, K.H., and Kvaal, K., J. Chemom., 10(5-6): 569(1996).
4. Esbensen, K.H., The Introducer Package: Multivariate Data Analysis In Practice, 5th edn., CAMO AS, Oslo, Norway, 2002.
5. Gy, P., Sampling for analytical purposes, John Wiley and Sons, New York, New York, USA, 1998.
6. Huang, J. and Esbensen, K.H., Chemom. Intell. Lab. Syst., 54(1): 1(2000).
7. Huang, J., "Developments in applied chemometrics-AMT, acoustic chemometrics and N-way analysis," Ph.D. thesis, Høgskolen i Telemark, Norway, 2001.
8. Huang, J. and Esbensen, K.H., Chemom. Intell. Lab. Syst., 57(1): 37(2001).
9. Huang, J., Ose, S., de Silva, S., and Esbensen, K.H., Powder Tech., 129(1-3): 1(2003).
10. Huang, J., Wium, H., Qvist, K.B., and Esbensen, K.H., Chemom. Intell. Lab. Syst., 66(2): 141(2003).
11. Johansen, S.B., Laugesen, J., Frøst, M.B., et al., Angle measure technique-proposal of new algorithm and systematic examination of method performances, in prep.
12. Martens, H. and Næs, T., Multivariate Calibration, John Wiley and Sons, New York, 1989.
13. Mortensen, P.P. and Esbensen, K.H., Chemom. Intell. Lab. Syst., 75 (2): 219 (2005).
14. Zuech, N. and Miller, R.K., Machine Vision, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1989.

Paper IV

Casper K. Dahl* and Kim H. Esbensen

"Image Analytical Determination of Particle Size Distribution Characteristics of Natural and Bulk Aggregates"

Chemometrics and Intelligent Laboratory System
Volume 89, issue 1, pp. 9-25, 2007.

IMAGE ANALYTICAL DETERMINATION OF PARTICLE SIZE DISTRIBUTION CHARACTERISTICS OF NATURAL AND INDUSTRIAL BULK AGGREGATES

Casper K. Dahl* & Kim H. Esbensen
ACABS Research Group
www.acabs.dk
Aalborg University Esbjerg, Niels Bohrs Vej 8, DK-6700 Esbjerg, Denmark

ABSTRACT
Image Analysis combined with multivariate regression on Angle Measure Technique (AMT) transformed imagery and the Theory of Sampling (TOS) is here presented as a comprehensive for *Image Analysis Sampling* (IAS), which takes all aspects of sampling representativity into account - especially 2-dimensional image versus 3-dimensioanl bulk compositions issues. Every IAS application has to be based on optimized image acquisition parameters: camera and illumination type, illumination angle, sample thickness as well as image post-processing, which are all examined here in order to delineate the general requirements for optimal prediction models for particle size distribution of natural and industrial bulk solid aggregates. We present a complete optimization study in order to show its intrinsic problem-dependent nature. This optimization allowed an original 60-sample data set to be compressed to an essential 22 natural coastal sands array with equally varying composition ranges – which was subjected to IAS in order to characterize the specific particle size distribution curves. In addition to D_{50} (50 %-tile), six other size classes were successfully predicted, while extreme size classes (extreme low or high particle sizes) showed a too narrow training data set span, illustrating a critical *grain size contrast* which will always bracket successful models of particulate matter being imaged for grain size characterization. All classes with a satisfactory (representative) calibration interval span can be quantitatively predicted due to the powerful scale-dependency of the central AMT feature extraction combined with PLS multivariate calibration. The present application to natural sand aggregate size distributions forms a vehicle to illustrate the full potential of image analysis in general, IAS in particular, also for technological/industrial manufacturing on-line product and process monitoring applications, or quality control purposes, with similar grain-size prediction objectives. There is a significant carrying-over potential to parallel industrial scenarios.

Keywords: Image Analysis, Image Analysis Sampling (IAS), Theory of Sampling (TOS), Angle Measure Technique (AMT), multivariate calibration, Grain-size Prediction

INTRODUCTION

In many aggregate materials and powder handling/processing operations particle size and particle size distribution play a key role. It is an essential quality control characteristic of many industrial product specifications often affecting important bulk characteristics, e.g. particle size distribution is a major determinant regarding flow properties. But most methods used for determining the particle size of powders and aggregates are designed to be used off-line, such as sieving, image analysis of prepared samples, laser diffraction. However advantageous with precise off-line methods, today it is of increasing importance to many industries to carry out this analysis on-, at or in-line. An on-line technique should both be fast and robust in delivering the results – typically for a large number of measurements – and reliable in having a high level of accuracy and precision.

Traditional low-level image processing techniques have been used for on-line measurements during the last ten years; a good overview of many successful implementations and applications can be

found in Zeuch [1]. The disadvantage of these approaches is that calculations are heavy, thus a limited number of measurements per time-interval can only be made [2,3] – and many of the dedicated analyzing instrument will necessarily represent a significant capital outlet. This paper presents an alternative extended image analysis technique, which potentially can be used for on-line particle size prediction as well as many other physical characteristics – based on standard, off-the-shelf, inexpensive camera technologies. In fact, it was originally thought of as a spin-off from other machine vision or image inspection operations. The new method is based on feature extraction from standard digital imagery using the Angle Measure Technique (AMT) with subsequent multivariate data analysis, Partial Least Squares regression (PLS-R). This combined approach was termed MAR (Multivariate AMT Regression) by Huang [4,5].

In the event of significant material segregation *before* imaging, this necessitates use of the Theory of Sampling (TOS) to ensure valid correspondence between the sample surface characteristics (the 2-dimensional image) and the true bulk (3-dimensional, volume) sample characteristics. Combining MAR with representative TOS-sampling introduces a new way to deal with this issue, which could be viewed as *Image Analytical Sampling* (IAS). IAS deals with all aspects related to this basic 2D (surface characterization) vs. 3D (true bulk characteristics) match, or conflict.

In the present work we develop a method for particle size distribution determination for natural beach sands excavated from the Western Coast of Jutland, Denmark. The Danish Coastal Inspectorate (*"Kystdirektoratet"*) makes extensive use of this type of sample as part of monitoring campaigns of coastal protection activities of and on the Jutland Coast, which is regularly subjected to significant erosion from the North Sea with its prevailing western wind in an intense and frequent storm regimen. Coastal monitoring is carried out in a variety of ways, including satellite surveillance, morphological analysis, mathematical modeling. As an important part of the monitoring, on-site sand samples are also routinely extracted from selected beach locations, which are characterized by their particle size distribution. The results are a.o. used for modeling of movement of sand as a function of various experimental erosion abatement parameters in several different projects. This material is here used to illustrate the potential of the proposed image analytical methodology for characterization of particle size class mixtures. It was decided to use beach sands samples because these furnish a realistic vehicle for modeling grain size in a very relevant particle diameter range. Indeed, if anything, natural sand mixtures are significantly more complex – distribution-wise – than the gamut of industrial powder and particle aggregates also considered as candidates for image analysis, which are usually confined by much narrower grain size distribution brackets. Thus the present approach is of a much wider application potential within science, technology and industry.

All surface morphology imaging approaches to aggregate materials will necessarily break down at some material-dependent limit-of-application because of the simultaneous presence of both large-diameter particles, θ_{max}, together with fines, θ_{min}. There is a problem-dependent limit to the fidelity with which an optical method can characterize both large and small particle diameters simultaneously – and their mixtures – due to transgression beyond a useful regimen of a material-specific grain-size contrast ratio $\Phi := \theta_{max} / \theta_{min}$. The overlying theme for IAS in the present context is then: By using sand aggregates as the medium, what can be learned about the general limit at which grain-size prediction models break down?

METHODOLOGY

IA for grain-size prediction is comprised of five elements:

1. Image Acquisition/Image Analysis (IA)
2. Image Sampling (IS)
3. Image pre-processing
4. Angle Measure Technique (AMT)

5. Multivariate calibration (PLS-regression)

1.1 IMAGE ACQUISITION/IMAGE ANALYSIS

There are two critical issues in image acquisition, which have to do with the quality of the images obtained as well as their representativity with regard to the characteristics of the lot material.

Image quality is a direct function of the imaging equipment as well as the relevant system settings. Figure 1 shows a sketch of a typical setup for digital image acquisition.

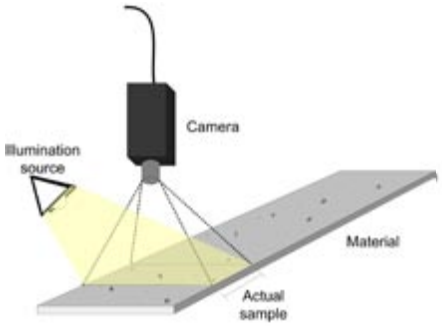


Figure 1. Principal image analytical system setup for characterization and monitoring of selected material characteristics. The camera only acquires surface images of a suitable illuminated "footprint" of a sample area; the images are transferred to a computer for subsequent processing and analysis.

Images only reflect the surface of the material; a certain area of the material is imaged only (the camera "footprint"). This area is determined from the geometry and setup of the camera/lens, which mainly is influenced by the optical distance to the sample surface. The camera quality is evident. The images are transferred to a computer for subsequent processing and Image Analysis (IA).

Selecting appropriate equipment and securing optimal image acquisition conditions is dependent on the experimental conditions. Also, camera and illumination source(s), and the optimal setting of these devices, must of course be selected based on the characteristics of the material to be examined, for instance particle size, shape, and color. However, such practical implementation problems are only addressed briefly in the present study, since the main focus is to show the prospects for monitoring bulk solid properties by the IAS approach only, as exemplified by particle size and particle size distribution for particle aggregates. We refer to the excellent textbook by Zeuch [1] for a complete review of these essential basic issues.

For every specific application context, optimal illumination conditions must be found in order to accentuate the key features and characteristics of interest, i.e. bring out colors, produce textural contrasts, create shadows (essential for the present AMT purpose) a.o. Illumination sources might e.g. be a UV, fluorescence or an ordinary tungsten source (VIS). The number of illumination sources, the illumination angle(s) and the distance from the sample surface must also receive serious consideration. For most image analytical techniques it is of course important to keep illumination as constant as possible. To keep the illumination constant in harsh industrial environments, it may at times even be necessary to use e.g. a dark chamber (or similar) around the equipment.

Selecting a camera appropriate for the tasks at hand is not trivial either. Efforts should for each application context be focused on the following principal issues:

- Camera type (digital/analogue)
- Chip type (CMOS/CCD)

- Lens
- Filter(s)
- Camera resolution
- Spectral characteristics (BW/RGB/multi- & hyper-spectral)

IMAGE ANALYSIS

Several different IA techniques are used extensively today, also within the powder industry. Traditional IA is typically based on filtering techniques in which images are subjected to various thresholding in the image and frequency domains, combined with object segmentation - for example for digital particle separation. This would then typically be followed by a series of relatively simple size, form and texture measurements of the individual particles. Standard IA software suites can be used here, which are based on a wealth of different image-processing techniques; being able to find the optimal method requires at least some insight and experience though.

Other techniques, such as MIA (Multivariate Image Analysis) [6,7] and Multivariate AMT Regression (MAR) [4,5], can be used as well. MIA is based on decomposition of the spectral domain into a principal components domain (scores and loadings), based on PCA. MIA has been used for both qualitative as well as quantitative analysis. MAR is based on the Angle Measure Technique (AMT) which extracts features of interest from the image domain into a scale domain. This transformation results in the so-called AMT complexity spectra which represent a compressed description of the morphological texture relationships of the original image simultaneous at *all* scales simultaneously (see further below). These spectra can be used as 1-dim object vectors in further multivariate data modeling for finding possible correlations between image features and material – or material functional characteristics.

Acquiring an image suitable for image analysis can be a task in itself, but it is always *doable* at the current technical development stage of IA. Ensuring that the image actually depicts the material below the camera in a *representative* fashion is quite a different, non-trivial matter, however. This study focuses primarily of image analytical representativity issues.

1.2 IMAGE ANALYTICAL SAMPLING

Acquiring representative images includes an understanding of the Theory of Sampling (TOS), which is a complete theory detailing how to obtain representative samples from all given materials and lot types [8-11]. The present work deals with material lots that essentially occupies only one, dominating dimension, such as conveyor belts and pipelines (to a first approximation the width and depth is negligible compared to the length dimension) – subject to the necessary condition that both images (X) as well as corresponding reference samples (Y) must *cover* the transverse dimension(s) of the lot completely, always consisting of an entire cross-section in width and depth respectively, as outlined in Figure 1.

Image Analytical Sampling (IAS) deals within the realm of “proxy-sampling”. It is often tacitly assumed that the 2-dim surface making up the camera’s Field-Of-View (FOW) also represents the average composition of the bulk (3-dim) sample without further qualifications, i.e. the material being imaged is *assumed* to be identical throughout the vertical dimension underlying the FOW. However, this assumption is valid if, and only if, the material to be analyzed can be demonstrated to be homogeneous, to wit:

> "Homogeneity is an abstract mathematical concept that does not exist in the real, material world."
> [P. Gy, p.24, Sampling for Analytical Purposes]

Since all materials are indeed inhomogeneous by nature (dependent on the scale of observation) representative reference samples can only be obtained through TOS-correct physical sampling. This is manifestly the case for samples used in MAR, since any calibration should be based on reference measurements (Y) determined from representative samples in order to produce a valid prediction model. Prior to physical extraction of samples, the surface of the lot is subjected to imaging (X). It is vital to estimate the degree of mismatch between the image analytical 2-dim rendition of the surface of a product relative to its 3-dim bulk characteristics. Only materials which can be/have been demonstrated not to be significantly segregated in their natural state can be directly imaged; image analysis will in general be representative for such situations (examples: potato chips on a conveyor belt, metal surfaces etc.). However, most particulate materials are manifestly prone to segregation to some significant degree and must therefore, as an absolute minimum, be subjected to thorough mixing before presentation to the camera. Thorough mixing will not make samples homogeneous however, but makes them only less heterogeneous. It will always be necessary to document the effective residual heterogeneity obtained by any particular mixing operation.

Combining mixing with a large number of *replicate samples* will also help effectively towards higher representativity.

Two types of *replicates* can be used in the imaging context:

- Physical replicates (pouring, loading) – reconstituted presentation of the sample (or samples) to the camera (simulates a number of replicate images *along* the conveyor belt etc.)
- Imaging replicates (sample rotation) – the sample is illuminated from different angles (simulates more than one illumination source or more than one camera angle)

Physical replication: The sampling error can be significantly minimized by reconstituting the sample (mixing it thoroughly) and re-subjecting it to imaging. This way, the bulk of the material will be introduced to the camera, but differently: another mixing surface will be imaged. Different views of the material will be recorded and interpreted by the image analytical facilities with an objective of *averaging*. On a conveyor belt this can be done for example by implementing a plow-like mixer just before the sample is presented to the camera hereby minimizing segregation effects; there are of course many other possibilities for physical mixing – all delightful engineering challenges.

Imaging replication: In addition, the samples might also be illuminated from different angles – one by one for unilateral illumination. This will produce "imaging replicates" which are not physical replicates, but each manifestly provides another *view* of the same surface of the material – it will be shown that this type of replicates also has great merit.

This study distinguishes clearly between these two types of replicates. The easiest way to minimize image sampling errors is to average over a multiple of imaging replicates, which is almost without extra costs in practice since the effort of taking one or more images (everything else being equal: illumination etc.) is virtually nil – but this endeavor will ultimately be limited by computation power especially in real-time contexts.

Sampling for reference analysis (calibration) is very much another matter, since these samples must be extracted physically (3-D bulk samples). Reference sampling is often the distinctly most costly

element in any replication context, for which reason it is economically "the first" item on all money-saving horizons. Whether this also is a scientifically sound approach is a totally uncorrelated matter however, especially when TOS sampling errors are involved. Image replicates (X) can of course only be regarded as true new samples *if* they are complemented by complementary new-physically extracted reference samples (Y).

1.3 IMAGE PREPROCESSING

UNWANTED EFFECTS

Images can contain many types of unwanted effects or artifacts, e.g. shading, unwanted objects, blurring (or other optical deficiencies), noise etc. Most of these effects can be reduced or eliminated by low-level image-processing techniques in the spatial domain [2,3] or, if need be, by careful application of special, i.e. problem-dependent, corrections. In the present work a significant shading effect in the shape of a clear illumination gradient was observed, which greatly affected the multivariate models. This gradient originated because of the necessary unilateral low-angle illumination used to optimize the AMT feature extraction [2]. Figure 2 shows an example of how this gradient appears to the camera as well as after removal by a simple gradient correction.

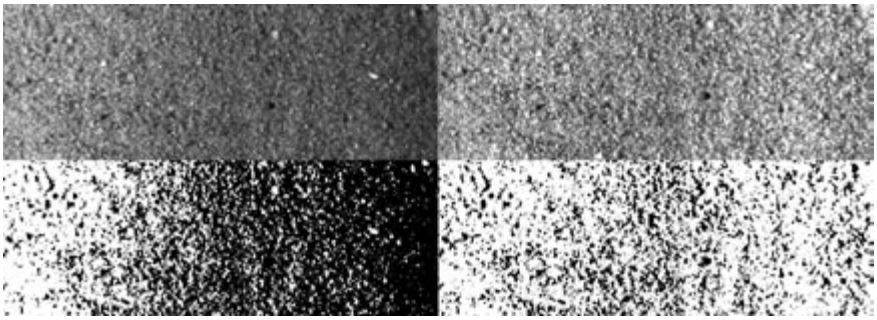


Figure 2. Sand sample with extensive illumination gradient (left) and in corrected form (right). Both images are also shown (lower panels) in a contrast-enhanced version for improved appreciation of the correction.

Since the gradient was found to be nearly linear in the horizontal direction (and non-existent in the vertical direction) the average pixel intensity per column can be used for a simple correction, which takes the form of a linear regression. If each pixel is corrected by the inverse of this function the result will be an image without this detrimental illumination gradient.

Other filtering techniques could have been used for removing shadings, e.g. morphological, homomorphic or lowpass filtering. Zamperoni [12] has pointed out that morphological filtering can be used for removal of uneven illumination effects, while the Quantitative Imaging Group at Delft University of Technology has shown [13] that homomorphic or lowpass filtering can be used for estimating any shading effects generally. Nevertheless, there is a serious drawback: More or less all filtering techniques, including the linear gradient removal function developed specifically for this paper, add heavily to the real-time computational load – especially for large number of images (replicates). This should of course be taken into account when setting up potential on-line IA applications.

1.4 THE ANGLE MEASURE TECHNIQUE (AMT)

The Angle Measure Technique (AMT) is a signal characterization method for describing the complexity of 1-dim data series. AMT was originally developed by physical geographer Robert Andrle expressly for characterizing 2-dim "geomorphic lines" (map contours, coastal lines, river trajecto-

ries) as an alternative to fractal analysis [14]. This technique has been further developed within chemometrics where it today can be applied to *all* generic 1-D data series [5]. For the present study AMT was used to facilitate extraction of useful surface morphology features from unfolded images, essential in the context of aggregate grain size descriptions and - prediction.

AMT focuses on extracting features, both global and local, as a domain transform operating analogously to the Fourier transform, but AMT transforms data series (unfolded images in this study) into a new scale domain; this results in so-called "complexity spectra". Following this domain transformation a standard multivariate data modeling approach, such as partial least squares (PLS), can be used for establishing relationships between the extracted signal features (AMT's complexity spectra) and the reference data in question – which in the present study will be grain-size frequencies. For a detailed description of AMT we refer to the extensive literature, e.g. [14,16,23-25].

Figure 3 gives a schematic overview of the complete AMT conversion process from images to prediction model (PLS-R), with Y being the independent data to be modeled and predicted; X represents the spectral data *after* the AMT domain transformation, T is the PLS score-vector, and W the loading weight. In the regression model, the two standard AMT spectra MA and MDY are concatenated in a fixed succession [MA,MDY]. The full workings of AMT + PLS has also been described extensively in the chemometric literature [17-19].

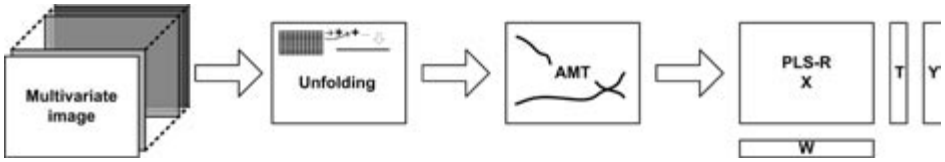


Figure 3. Schematic overview of AMT as a PLS-regression pre-processing transform. One color channel has to be selected, i.e. the channel with the highest contrast. This is followed by AMT feature extraction on the unfolded imagery [MA,MDY], from which a regression model can be established.

Before any multivariate regression is carried out, averaging of the AMT spectra for replicates is often performed. The average for each AMT scale is found and used as a more robust expression of the image complexity. Using averaging dramatically improves both the precision (RMSEP) and accuracy (slope) of properly validated prediction models. Figure 4 is an example of the averaging procedures as used in this study, expressed by two complementary prediction model validations ("Predicted vs. measured" plots with associated statistics), with – and without averaging.

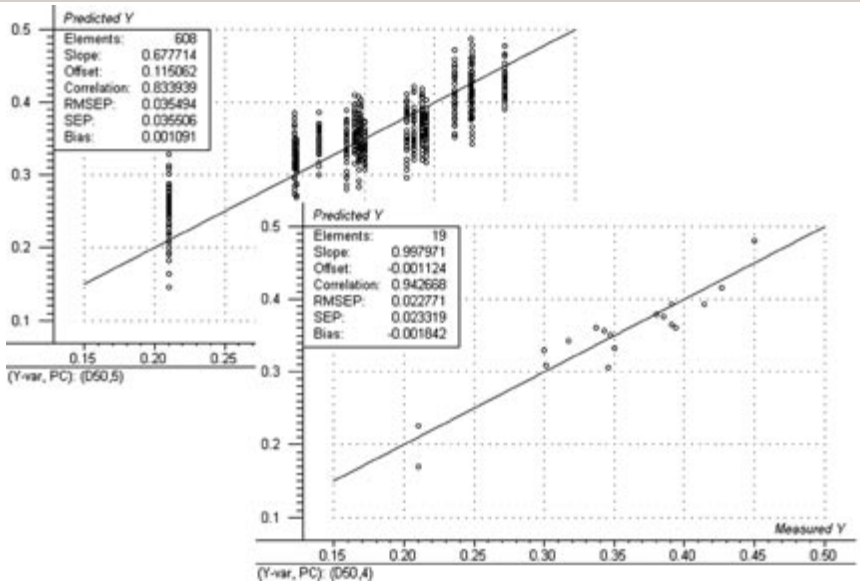


Figure 4. PLS-regression validation ("Predicted vs. Measured" evaluation plot) for a particular grain-size feature (D_{50}) with and without averaging of both physical and imaging replicates (described in text). Note the dramatic improvement in prediction performance for the averaged data (lower right panel).

1.5 MULTIVARIATE REGRESSION

For the present work PLS-R is used for establishing regression models between the AMT spectra (X) and the reference variables (Y).

VALIDATION

When a regression model has been established, its predictive performance must be evaluated through a suitable validation procedure. The only way to know if a model will truly give realistic predictions of future observations is to test its performance on a completely new and independent data set (called the test set). All data sets (both training and test sets) need to be obtained in a way so as *to be representative of the future situation in which the prediction model is to perform*. Needless to say both types of data sets must be based on samples which have been extracted using proper sampling techniques for the bulk material in question: representative sampling [8-11,16, 22]. Test set validation with samples extracted according to TOS is the only way to get a *realistic* assessment of the full prediction error [17]. On the other hand, all cross-validation schemes underestimate the physical sampling error seriously, as there is only one data set - the calibration data set. One needs the additional test data set in order at least to have sampled the material ("the population" if one so prefers) twice, so as to incorporate two different TOS sampling-error manifestations in the validation procedure; if the material is severely heterogeneous, using more than one test set would appear a prudent approach, *ibid.*, always subjected to the economics and logistics of the situation at hand.

The one and only case, where it is statistically sound to use other validation approaches, is for comparative evaluations regarding alternative model specifications. In this special context it can be argued that using a well-reflected version of cross-validation is sensible, since no additional errors will be introduced through different application on the same test set. However, any final prediction model should *always* be tested on a completely independent data set to get a fully realistic estimation of the prediction strength.

MATERIALS

For the present study we used a master sample set consisting of 60 pristine coastal sands with significantly different particle size distributions. These sands have arrived at their present size distribution status through a set of complex coastal and beach processes (erosion – transportation – deposition – reworking). These sands come from the entire length of the Danish west coast (approx. 300 km in length), making up 60 widely spanning sands, each with its own specific history. It is certain that this calibration data set is considerably more complex than very many industrial aggregates; many industrial and manufacturing aggregate products will display a much more constrained grain size distribution, while some raw materials could be as complex so as to be comparable to the present set. From experience with IA/AMT characterization of solid aggregates, the present data set can be viewed as representative both in overall grain size distribution span as well as w.r.t. sorting.

The sixty samples were selected from the vast coastal sand archives by senior research personnel at the Danish Coastal Inspectorate (*Søren Bjerre-Knudsen*). Each sample originally consisted of three increments excavated from three different depths from one pit-hole, combined and mixed in the field, this being is the traditional sedimentological characteristic employed for beach sands. After transport to the inspectorate storages these samples are mass-reduced to a prescribed archival mass using riffle splitting in accordance with optimal TOS principles [11], and the particle size distribution of each sample is finally determined by standard laboratory sieving using 15 trays. The particle sizes are hereby classified in sixteen different size classes, which in these experiments are used as the independent data to be modeled (variables Y_1 to Y_{16}). These reference data were supplied for the present study together with the selected archival samples themselves. These samples and their reference grain-size distribution data are considered highly trustworthy as they all originate from the same laboratory (they were in fact produced by the same, very experienced laboratory technician).

The overall variations in composition, particle size and sorting can be appreciated from Figure 5, in which the entire spanning ranges for these features have been captured by a selected series.

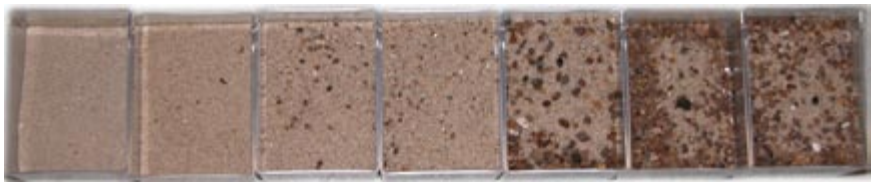


Figure 5. Illustration of the wide material composition, grain size and sorting span of the natural coastal sand calibration set. Note significant differences especially regarding particle size, the objective of the present study.

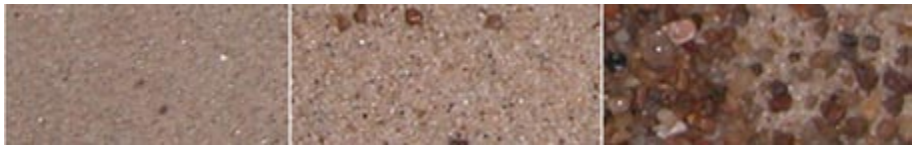


Figure 6. Close-ups showing extreme span in average particle size as realized in the master data set, as well as large differences w.r.t. particle size contrast (sorting). Images are to identical scale.

Figure 7, 8Figure 8 and 14 describe the grain-size distribution characteristics of the training set. In terms of absolute grain sizes, many samples display a relative small average particle size ranging between 0.2 and 0.8 mm, with a positively skewed tail covering larger average particle size; only 3 out of 60 samples (5 %) were found to have an average particle size (D_{50}) above 1.0 mm. The sort-

ing of these natural sands is well characterized by their cumulative size distribution curves as illustrated in Figure 7.

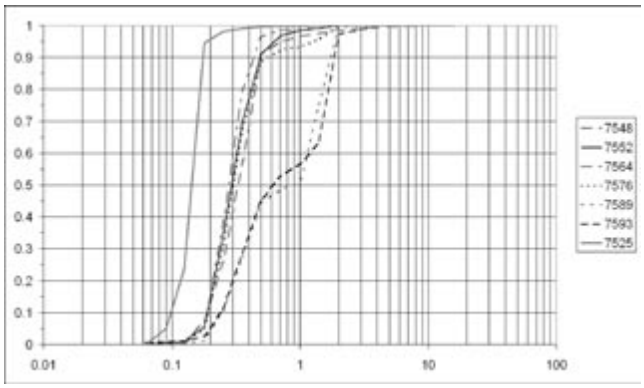


Figure 7. Selected cumulative distribution curves for seven sand samples with D_{50}-values ranging from 0.200 to 1.259, covering the full span of the present prediction models.

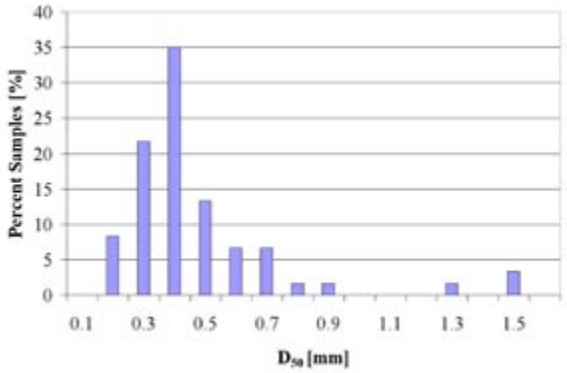


Figure 8. D_{50} particle size distribution for the DCI 60-sample data set. D_{50} values were derived from the 15-bin sieving data. Only three samples display D_{50} above 1.0.

The particle size distribution skewness (inherited from the DCI data sets) will cause some inherent calibration difficulties. This is a minor issue only however as this study was not meant to focus on applicability of image analysis on all types of beach sands specifically (neither necessarily covering the gamut of 16 size classes). This study instead focus on illustrating the *general* applicability potential for determining complex bulk aggregate/powder parameters – especially aggregate particle size distributions to the number of size classes realizable.

EXPERIMENTAL

Obtaining images of non-moving materials is relatively easy, in comparison with moving objects; exemplified by a conveyor belt, perhaps suffering adverse from illumination hindrances etc. Under all circumstances the image acquisition conditions are always of critical importance. Especially illumination source, camera type and illumination angle when images are to be preprocessed using AMT.

To zoom in on optimal imaging conditions for the sand samples acquired, a number of pre-experiments were performed, which can be divided into three series with the following objectives:

1. Selection of optimal illumination source conditions
2. Optimization of camera type and illumination angle
3. Final modeling of 16 different size classes based on these optimized conditions

All images from the second and third experimental series were subjected to AMT. Prior to this the images were unfolded by row-wise concatenation. 5000 points were selected randomly from each unfolded image as seed points for AMT and each of these points was analyzed for the first 250 AMT-scales (for both MA and MDY).

1.6 SERIES ONE – ILLUMINATION SOURCE

Illumination is alpha and omega in image analysis. In the first half of the present study three different illumination sources were investigated for optimizing the image contrast on a particle-to-particle scale:

- LED Darkfield Ringlight, which is often used to enhance the contrast of surface features.
- LED Ringlight is used in pattern recognition due to its shadow-free illumination.
- A standard unilateral, oblique illumination source: a LHS-500 lamp (FalconEyes) was used with a 150 W bulb. This type of unilateral illumination has been used extensively in earlier AMT-applications giving rise to shadows which enhances differences in particle sizes/shapes very efficiently. In the initial study an illumination angle of 45° was used, based on said earlier experiences.

Eight different sand samples were used in series one (a widely size spanning end-member series). Figure 9 serves to demonstrate the manifestations of the three different illumination sources.

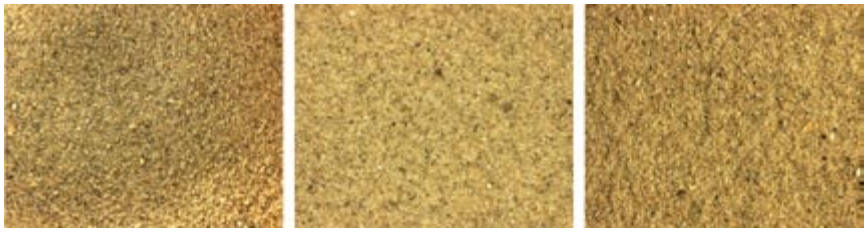


Figure 9. Photographs of the same sand sample, illustrating contrast differences resulting from the three alternative illumination sources. Photos (left to right) were obtained using the darkfield ringlight, ringlight and the LHS-500 lamp, respectively.

Evaluating these alternatives over eight very different sand samples, it was decided to use the LHS-500 lamp, because of its optimal contrast, as evidenced in the right panel in Figure 9. The unavoidable illumination gradient could easily be removed by the corrections already described. The choice of a unilateral illumination source corresponded well with experience from many previous AMT experiments in which an identical setup has indeed also proven superior in many cases [5]. In the end illumination optimization will always depend exclusively on the specific application context at hand.

1.7 SERIES TWO – CAMERA AND ILLUMINATION ANGLE

For unilateral AMT illumination, the illumination angle is of critical importance. A very high illumination angle close to 90° (w.r.t. the horizontal) will remove shadows, while lower angles will add shadow effects to the image – the lower the angle, the larger the shadows. These shadows are aiding the AMT-algorithm in distinguishing between, for instance, particles with different size, texture and/or shape. Unilateral illumination is a unique strength to AMT.

In addition to the illumination angle, two different cameras were also tested in series two:

- Sony 3-CCD analogue camera type DXC-390P with a 16 mm F1.4 C 2/3" objective. Digital images were obtained using a 3-channel Matrix Vision frame grabber, type PCI-mage-SRGB, with a resolution of 752x572 with 8 bit color in each channel.

- QImaging digital b/w camera, type QICam with a 105 mm F2.8 AF Micro-Nikkor objective. To obtain color images a RGB Liquid Crystal Color Filter is attached. QI images display a resolution of 1280x1024 also with 8 bit color in each channel.

The image resolution and the pixel size are both important, as they limit the ultimate small-scale texture characteristics possible in the AMT complexity spectra. For experimental series two, the images were obtained with identical distance from camera to sample, but with different resolutions due to the inherent different camera specifications. This difference is shown with great fidelity in Figure 10, which shows two alternative renditions of the same sample. Both images are 200x200 pixels in size and both are at their maximum resolution.

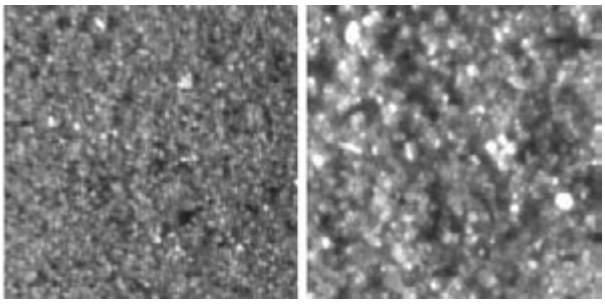


Figure 10. Comparison of maximum image resolution for the Sony (left) and the QImaging camera (right). Images were acquired at identical camera distance and both images represent 200x200 pixels.

From Figure 10 it can be seen that the QImaging camera (right panel) has more detailed information at smaller scales than the Sony camera. In the final modeling this could make a difference, unless the higher resolution acquired by the QImaging camera does not add useful information.

Both cameras were tested with alternative illumination angles 28° and 45°. For this optimization trial, 16 other samples were selected from the DCI master sample set based on their visual appearance (covering the whole distributional curve for D_{50} for all 60 samples); the selection spanned almost the full grain-size range covered, except for the absolutely smallest grain-sizes, Figure 11.

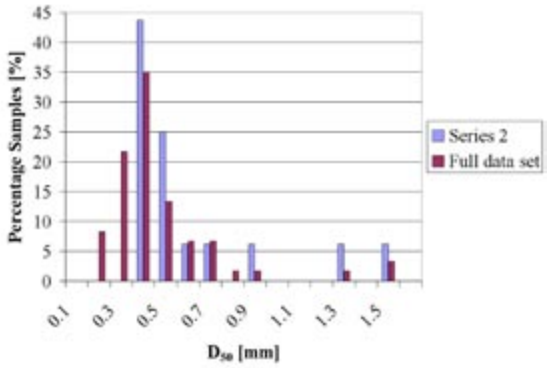


Figure 11. D_{50} size distribution for the total DCI data set (60 samples) and the series two selection termed "16 samples" in these studies.

Series two - replication scheme

To be fully realistic it was decided to use two true replicates (re-poured samples for renewed presentation to the camera) and eight imaging replicates (rotations) for each sample, giving a total of 16 images for each individual sand sample; each rotation was 45°. This setup is illustrated in Figure 12. Altogether using these settings on two illumination angles and for two cameras, 256 x 2 x 2 = 1024 images were obtained for the series two evaluation.

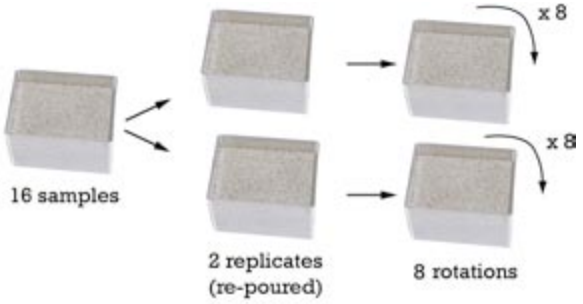


Figure 12. Sample replication setup for experimental series two. Sixteen samples were selected on basis of their visual appearance, covering the widest possible D_{50} span. Each sample was poured twice, and rotated eight times. This procedure resulted in 16 x 2 x 8 = 256 images per sand sample (for each camera and illumination setup).

In addition to illumination angle and camera type, two other parameters will be important for the present type of material (sands): choice of color channel and whether to use scaling or not in the ensuing chemometric data analysis. With a color image composed of three grayscale images (R/G/B), 3 x 1024 images were obtained. These images were all subjected to AMT domain transformation, which was followed by averaging for the 2 x 8 replicates for each sample.

Altogether 24 different multivariate models were established for optimizing the settings with regards to camera type (2), illumination angle (2), color channel (3), and whether to use auto-scaling or not (2). The images were all subjected to the gradient correction algorithm (or not), which resulted in 24 additional models. All models in series two were established for mean centered data using full cross-validation, here considered acceptable as a vehicle for internal comparison purpose for which it was desired to not compound the issues with physical re-sampling variability.

1.8 SERIES THREE – PREDICTION OF SIZE DISTRIBUTION

Based on the optimized set of parameters from series 1 and 2, it was found necessary to focus on sand models without the extreme highest grain sizes (D_{50} below 1.0 mm) as the upper part (D_{50} above 1.0) could not be adequately covered by three samples only. Without loss of generality, focus in the rest of the studies presented here is on $D_{50} < 1.0$ mm, see Figure 8 and Figure 11.

Spanning this main interval of the distributional curve regimen, 11 samples were selected to serve as the calibration set and 11 separate samples for the test set, the characteristics of which are both illustrated in Figure 13. These samples were selected on basis of a visual inspection of all original samples. Both sets were selected to be as widely spanning (while only broadly comparable) as possible, i.e. certainly not identical, in order to comply with the most stringent test set validation criteria [17]. The primary comparison feature was the D_{50}-values, while the grain size variances were left to random choice. Since experimental series three focuses on establishing the final prediction models, this deliberate calibration set/test set definition was essential. Test set validation is used throughout for all series 3 prediction models.

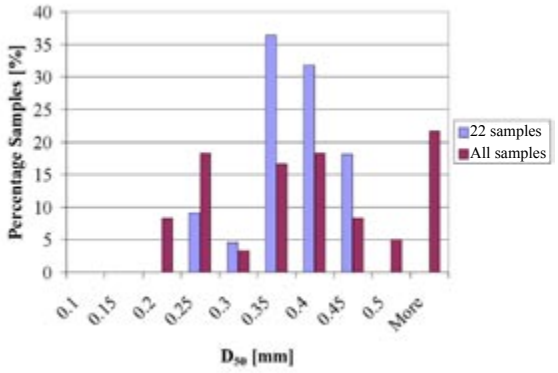


Figure 13. Comparison of D_{50} size distributions for the master data set and the 22 samples in the two series 3 data sets.

Figure 14. The combined (calibration, test) data set for the final third series (2 x 11 samples). Top left: highest D_{50} through lower right: lowest D_{50}. Every second sample in this lineup was selected alternatively for the training and the test set respectively.

Since the present sand samples are certainly not homogeneous (even though some are rather well sorted), sampling errors – often manifested as grain-size segregation – will necessarily occur to a significant degree, the point being that all necessary counteracting precautions must be applied (TOS): 1) mixing reduces sampling errors, 2) sampling errors will be lowered through averaging based on an increased number of replicates, 3) sampling errors will also be minimized by illuminating the sample surfaces from different angles.

Obtaining a higher number of images can be effectuated while at the same time reducing the potential 2-dim / 3-dim mismatch tendency through an increased surface area by a minimization of the sample thickness – a flattening of the sample layer as it were – ultimately (but rarely fully realizable) into a very thin, quasi mono-particle layer. With image analysis of stationary objects this is an easy task, whereas flattening of samples can be difficult to implement in applications with non-stationary, moving material (conveyor belts, pipelines). Irrespective of the practical difficulties e.g. in the industrial setting, sample thickness should *always* be optimized for each individual application. To illustrate the cardinal importance of sample thickness (or its reciprocal: surface area), a thin and a thick sample layer for each sand sample were produced during all series three experiments (based on an equal sample volume). In practice this was done with the use of the two alternative sample trays shown in Figure 15.

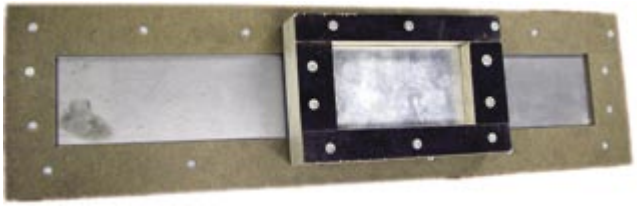


Figure 15. Alternative sample trays for the series three final experiments. The trays hold an identical sample volume; the large tray (bottom) is four times as long as the small (top).

The large tray holds the exact same volume as the small one and they both have the same width, 7.2 cm; the large tray is four times as long, 50 cm, while the small one is 12.5 cm. Compared to the sample masses used, the long tray comes very near to establishing a mono-particle layer of material (0.25 mm), with practically all sample material completely spread out, while the small tray has a rather small surface area and consequently a correspondingly greater thickness. Imaging a larger surface area with unchanged effective pixel size should indeed result in significantly better prediction models, in more-or-less direct proportion to the degree of segregation propensity displayed by the lot material.

For experimental series three it was decided to use four physical pouring replicates (re-introducing the samples four times completely anew to the camera). To simulate the effect of illuminating from different angles, imaging replicates were used as well. Due to the experimental setup it was not possible to rotate the large tray more than +/- 180 degrees, i.e. to simulate illumination from two opposite sides only, Figure 16.

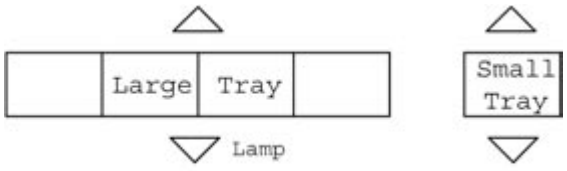


Figure 16. The large sample tray was virtually divided into four regions with the tray being illuminated from two sides (opposite). The same illumination setup was used for the small tray, but see Figure 17.

The large tray was virtually divided into four regions, each with the same surface area as the small tray. All regions were subjected to identical imaging; with illumination from each side eight images were obtained for the large tray. Each sample in the small tray was illuminated from eight different illumination angles in the horizontal plane, based on the results from series 2. Thus eight imaging replicates could be realized for each tray option.

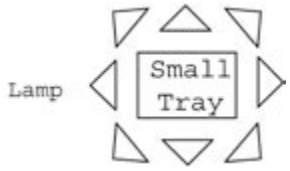


Figure 17. Eight different illumination sources were simulated by rotating the small tray. The sample is seen from the camera's perspective.

Replication rationale

- Using the above procedure for the large and the small tray respectively will make it possible to look into the effects of segregation - if detectable. Equally good prediction performance of models from both tray types would indicate insignificant levels of segregation.

- Comparing models based on the average AMT-spectrum for these eight replicate images (eight illumination angles) with models based on only one illumination angle will make it possible to investigate if better prediction models can be obtained by use of illumination replicates. If not, this would indicate that it will be sufficient to use only one illumination source making it easier to implement the technique in practice.

- With this compound experimental setup it will also be possible to investigate whether segregation can be detected or not; this should be seen from better overall prediction models for the experiments using the larger surface tray compared to the one with the smaller surface. Naturally, the eight imaging replicates for the small tray should represent the surface better than the two used for the large tray, but this is carefully counterweighted by the four-time increase in surface area. *If* segregation is minimal, the prediction models for both trays should be equally good.

- Series three thus probes directly at the core issue of this study, the potential 2-dim / 3-dim correspondence, or mismatch, between image data (X) and reference results (Y); the latter are provided by the external DCI grain-size distribution results.

RESULTS

1.9 INITIAL RESULTS – OPTIMIZATION OF EXPERIMENTAL PARAMETERS

Series two and - three experimentation was carried out based on one unilateral illumination source, following the salient pre-experiment conclusions reached above. The results from the second experimental series are based solely on the predictive ability of D_{50} and are summarized in Figure 18 and Figure 19.

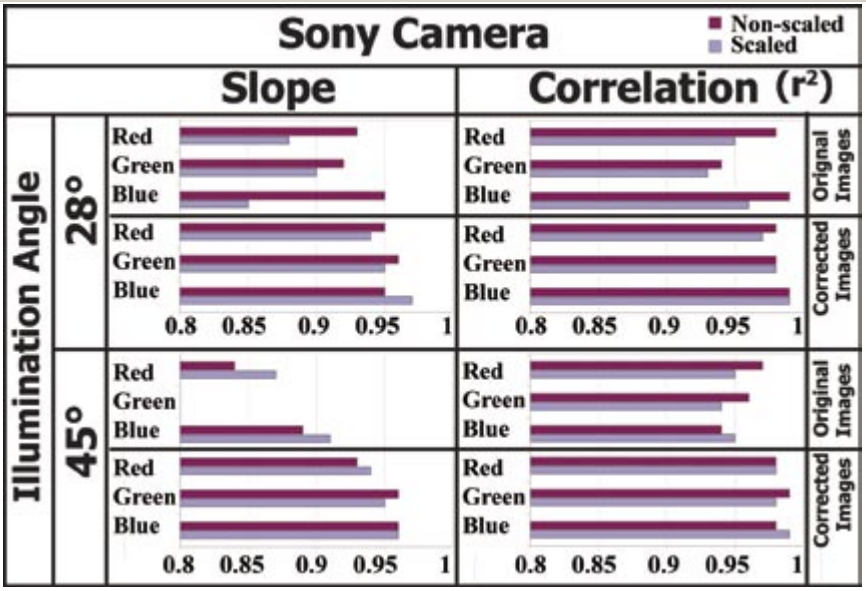


Figure 18. Results from the initial experiments, series 2: Sony camera. Results for scaled and non-scaled models for both original images (with illumination gradient) and the corrected images are displayed - for both illumination angles.

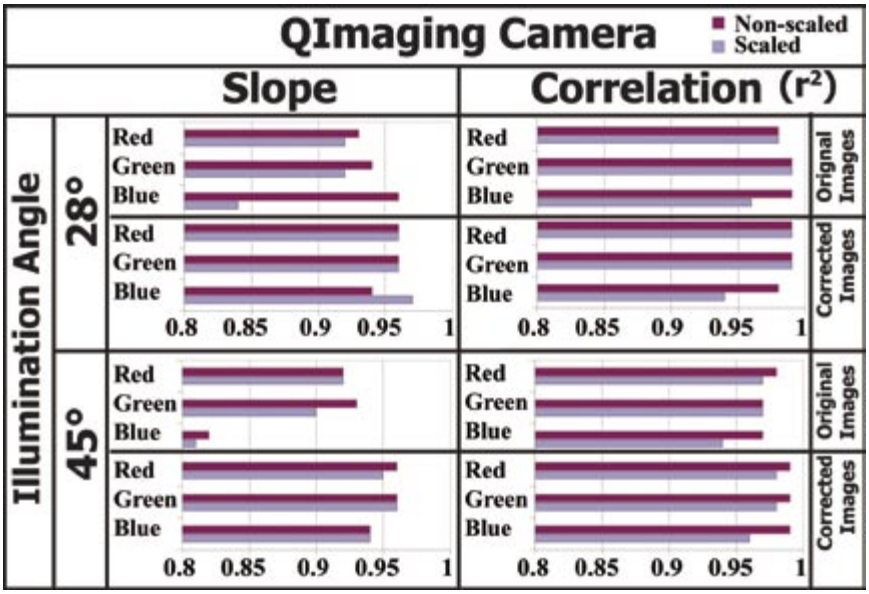


Figure 19. Results from the initial experiments, series 2: QImaging camera. Results for scaled and non-scaled models for both original images (with illumination gradient) and the corrected images are displayed – for both illumination angles.

For the comparisons below between the many alternative PLS-R prediction models, we make use of a two-parameter evaluation based on performance statistics pertaining to a fitted linear regression line for the "Predicted vs. Measured" relationships. The slope of this fitted line is a reflection of the overall accuracy of the predicted results, the closer the slope is to 1.0 the better the statistical accuracy of the prediction model. The squared correlation coefficient (Pearson second moment correlation coefficients), r^2, is the complementary measure of the overall precision of the predictions from the test sets; the closer this measure is to 1.0 the better the prediction model.

The primary objective was to investigate if any correlations between the AMT-spectra and the whole range of reference sieving values exist. Figure 18 and Figure 19 show that very good, indeed excellent models of D_{50} can be easily established. In general, the QImaging camera gives better models in comparison to the Sony camera due to the higher resolution (3x larger). Both cameras are affected by the illumination gradient artifact. Correcting this gradient on average gives slightly better models. Scaling the AMT-spectra does in general not seem to improve the models. Based on an overall assessment no clear conclusion can be drawn with regards to the selection of color channel, since the differences in general seem to be either marginal (non-significant) or to show no clear trends. These irrelevant minor nuances notwithstanding, the overall result from these experiments show a very good modeling potential with highly promising predictive abilities.

Based on the synoptic series 2 optimization results presented in Fig.s 18 and 19, it was decided to use the QImaging camera with a unilateral illumination source at a 28° angle. It was of course also decided to remove the illumination gradient. The red and the green color channel are equally good for describing the differences in particle size, while blue seemed to be slightly worse or not giving consistent results. It was therefore decided to use the red channel for the final modeling phase, also because the red channel previously showed equally good ability towards describing surface features of similar aggregates (sand, clay, cement …) [4,5,16,24]. Again, channel selection should always be a problem-dependent issue.

In order to determine if the number of illumination sources has a significant impact on model predictability, four different settings were tested. The results are based on the red color channel with the corrected images from the QImaging camera with 28° illumination angle.

Table 1. Influence of # of illumination sources. Four different settings were tested: Eight sources, four sources, two sources, and only one illumination source.

	Slope	Correlation
8 illumination sources	0.96	0.98
4 illumination sources	0.94	0.97
2 illumination sources	0.93	0.96
1 illumination source	0.87	0.95

As is evident from Table 1 the number of illumination sources *is* important. It was therefore decided to use eight sources (eight sample rotations in practice) for the third experimental series.

1.10 OPTIMIZED PREDICTION OF PARTICLE SIZE DISTRIBUTION

The calibration-prediction results in Figure 20 and Figure 22 aim directly at deciding on the core issue: is there a sample depth-influence on the realizable prediction performance (given the set of optimal experimental settings)? The results are singled out as large surface (area) vs. small surface (area) – and 2 x 4 imaging replicates vs. grab-sampling (2 x 1 image).

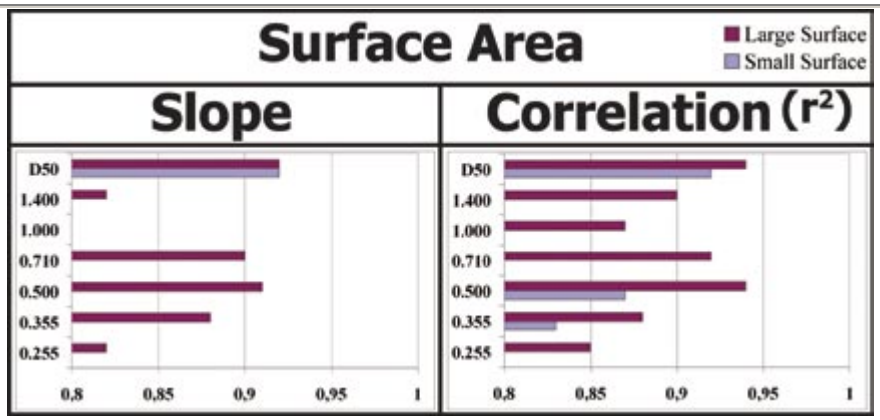


Figure 20. Final prediction models for sand based on eight replicate images for the large tray vs. two for the small tray. The models are for size classes 1.400, 1.000, 0.710, 0.500, 0.355, 0.255 mm as well as D_{50}.

Figure 20 shows very clearly that better models can be obtained by expanding the surface area. Slopes generally > 0.85-0.90 and correlations generally > 0.85-0.90 are *excellent* results for materials with the given levels of grain-size heterogeneity.

To determine the effect of eight illumination sources compared to only one, the models for these experiments are compared in a similar fashion in Figure 21.

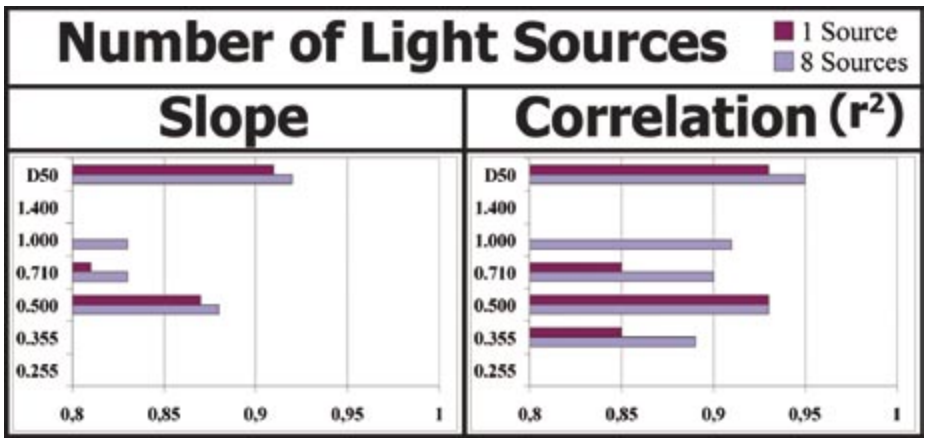


Figure 21. Comparison of final prediction models for the small tray using one and eight light sources respectively.

It can be seen that prediction performance are indeed improved by using several illumination sources. The present sand particles are quite rounded (reworked beach sands) and the sample surfaces were produced so as to be quite flat when being poured into the trays.

These complementary series 3 results show conclusively that using a larger surface area in combination with a larger number of illumination sources will result in superior prediction capabilities - but which of these factors has the largest comparative effect? Figure 22 delineates the salient differences between models with an increased surface versus an increased number of illumination sources.

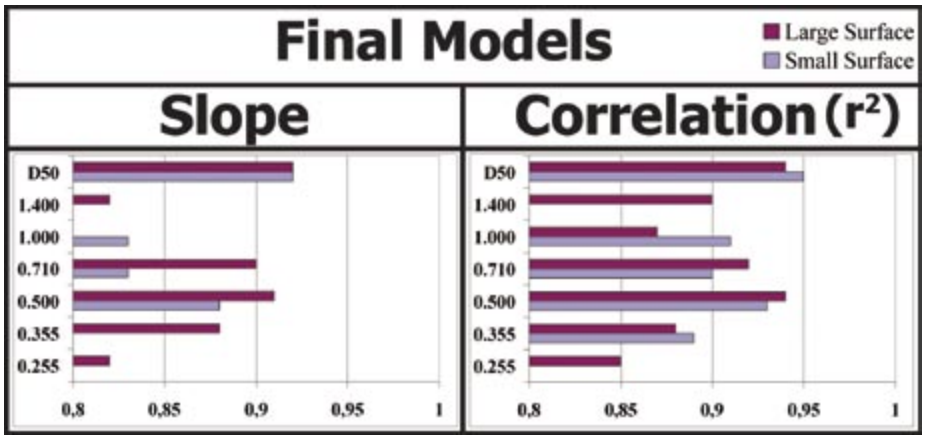


Figure 22. Final series 3 prediction models for sand based on eight images for the large tray (4 x 2 illuminations) vs. eight (rotations) for the small tray.

From Fig. 22 it is clearly seen that an increased surface area overall result in more accurate predictions (slope), while the situation w.r.t. precision (correlation) is rather equally balanced w.r.t. improvement; notice however that all models are of a highly satisfactory nature. From Figure 22 it can be concluded that very good prediction models are obtained for D_{50} and six size classes as well: 1.400, 1.000, 0.710, 0.500, 0.355, and 0.255 mm.

It bears noting, however that four minor deviations from these results (out of altogether 22), very likely attest to a combined effect from the fundamental sampling error (FSE) and the grouping and segregation error (GSE), well-known from TOS. Even at these small sample volumes, the effects of material manipulation (pouring, loading) as well as segregation still play a role in spite of the most strenuous efforts to control them. FSE and GSE are universal adverse agents. One would do well to be conversant about TOS to a sufficient, self-dependent degree.

DISCUSSION

The AMT-augmented image analytical methodology used in this study has shown its applicability for determining size distributions for sand mixtures of varying composition. Successful prediction models were obtained for the classes covering the central interval of the training data sets grain-size span. Those classes which could not be modeled well are characterized by a relatively low fraction of very large particles, which will have an adverse impact on the modeling possibilities if these classes are not very well represented in the training data set. Image sampling can be considerably compromised as a consequence of segregation: large particles might either show up on the surface images recorded, or not, with obvious serious consequences regarding their representativity.

Also the extremely fine-grained classes experienced difficulties. This is the exact scenario perceived related to the grain size ratio characterizing a particular material: if the ratio of the largest to the smallest effective grain diameters, $\Phi := \theta_{max} / \theta_{min}$, surpasses a certain critical threshold, IA/AMT will experience increasing difficulties w.r.t. fidelity of the image information in relation to the bulk, 3-dim, concentration reference values. This is not an innate weakness for the IA/AMT method as such, but an unavoidable consequence of the fact that there will necessarily always be a limit to the size contrast which can be modeled by a surface image alone. A completely analogous situation exists within acoustic chemometrics, in which the exact similar ratio $\theta_{max} / \theta_{min}$ brackets

the possibilities to simultaneously acquire informative acoustic spectra related to the specific grain-size setup of free-flowing mixtures [20].

Segregation is of critical importance in image analysis and can only be avoided if all particles are identical in size, weight and shape. Nevertheless, segregation is often *assumed* to be negligible, and checks on this assumption is very rarely carried out which can be fatal. In all practical applications segregation does occur – at some scale. The experiments examined in this paper are no different, witness that even the series 3 results indicated that *some* segregation had taken place: the model data summarized in Figure 20 showed that, on an overall basis, better results are obtained when a larger surface is used during the imaging. This was the case even though a rather large number of illumination sources were used when the small surface was examined. The present sample "depth thickness" is very small indeed compared to the gamut of industrial conveyor belt situations for example, in which one is naturally most often interested in the maximum mass flux possible, i.e. a high load per unit conveyor belt length := a thick sample depth. Thus, in the typical industrial setting, segregations will have a (much) larger effect than realized here – a fact to contemplate in all IA projects.

A technique similar to bed blending was used during pouring of the samples in this study, expressly in order to minimize segregation during pouring as illustrated in Figure 23.

Figure 23. Bed-blending technique as used in this study. Whenever this technique is used, the material should become as evenly spread out as physically possible compared to when the material is poured into the sample container in one single pouring operation - which is a sure way of introducing maximum flow - and other segregation effects.

Whenever the bed-blending technique is used, the variability of re-pouring results should be significantly reduced. Still the results using the two different surface areas were quite different in this study.

Another way to minimize the imprecision associated with prediction of the actual lot value will be to use a large number of (physical reference, or imaging) replicates. This is in complete accordance with the principles of process sampling according to TOS [8-9,21-22]. In the present study, four imaging replicates were used, a number that preferably could be even higher in real-world implementations. However, the present study was undertaken to investigate whether this extended image analytical technique can at all be used for quantifying the frequency of particles having a certain size, wherefore it was of prime importance to use a consistent number of replicates between model alternatives. The results also indicate that it will be beneficial to use several illumination sources (imaging replicates) to minimize the image sampling – and hence the final prediction error. The balance and number of physical and imaging replicates should always be optimized relative to the specific practical usage of course.

It is important to use an optimal type of camera for each specific application context. Two different cameras were tested in the initial experiments and the high-resolution camera from QImaging turned out to give marginally better models, which was expected due to the 3x increase in resolution compared to the camera from Sony. Other factors, such as the particle size in pixels (the effective

pixel size) and the fraction of the total particle number shown in each image (again, the effective pixel size), has always played an important role in determining if AMT can be used for feature extraction from images. All images obtained in the final experiments encompassed a large number of particles within the Field-Of-View (FOW) with particle size ranging between approximately 1 and 50 pixels (linear dimensions). In this study it was of paramount importance to have a large number of particles in each image. If an image displays only a few particles, the AMT algorithm will mainly extract features from the *surface* of each particle only, which in this case will contribute to the total amount of noise, whereas the feature extraction should rather focus on the differences in size *between* the particles. A large fraction of particles shown in each image will also contribute in making the images more representative of the material to be sampled, always leading to better prediction models.

The present feasibility prediction models could almost certainly be further optimized using software/hardware filters in order to enhance for example image contrast better. In the initial modeling phase various software contrast-enhancing filters were tried out with varying success, but improved prediction models *were* obtained. If a color camera is used, this contrast enhancement must be done on each color channel separately. This often leads to differences in predictability for the different channels, but only very small variations were observed for the present study (series two).

Two different illumination angles: 28° and 45° were examined, where the low angle option was found to be marginally better, absolutely in line with what would be expected; a low angle gives rise to larger shadows contributing to the overall more informative AMT feature extraction. The illumination angle should always be optimized locally in the pertinent problem-dependent context. Even though the differences are marginal in the present experiments, large differences in model predictability have been observed in previous experiments [23].

It *might* be argued that one could also have included interaction terms in the underlying experimental design used for investigating the many factors influencing the quality (contrast, definition) of the X-images of the sample surface. While we have made no formal ANOVA evaluation here, Fig.s 18-22 in point of fact indeed were evaluated based on simultaneous considerations of the effects from all factors included. As essentially all individual prediction models were already highly acceptable in their own right (main factor effects), it was decided to rest the case and not to expand an already very large experimental campaign further: very nearly all alternative models already displayed slopes and r^2 well above 0.90. It was decided to let potential, additional second-order improvements be part of the salient industrial, technological or scientific problem-context into which any IAS facility will have to be implemented under all circumstances.

The AMT algorithms have not yet been optimized for on-line implementation, for which reason it is quite time-consuming to calculate both the gradient corrected images and the AMT-spectra. However, these algorithms *can* be optimized – possibly even for on-line use – after which a powerful computer can turn each image into a spectrum within a fraction of a second instead of minutes. With variable selection in the subsequent multivariate data analysis and a good camera for imaging, this methodology has shown good potential towards on-line implementation.

> It was observed that auto-scaling does *not* improve the prediction models for the material examined, contrary to very many general chemometric experiences. Recent, parallel work on the general features of the AMT method (by the *AMT-consortium*: results to be presented shortly in the proceedings from SSC10) has lead to a full understanding of why this is so: It is not the relative information potential between X-variables that is important (as is the case for e.g. spectroscopic X-data) – In fact different X-variables in the AMT "complexity spectrum" are fundamentally different in that they represent different scale levels, between which there is no continuity analogue vis-à-vis spectroscopic spectra. As a consequence, the information content present in the absolute levels of AMT's MA as well as the MDY spectra carries a very important part of the information to be correlated to the Y-data – hence no improvement by applying auto-scaling which relativize along the variable dimension of the X-matrix.

Why Does AMT Work in Characterizing Particle Size Distribution?

AMT has been shown to describe textures in images very well (as a function of many features: intrinsic chemical particle properties, particle mixtures, particle surface characteristics, other) [4-5,16, 23-25]. The success in describing particle sizes and distributions emerges from the use of a unilateral, low-angle illumination, which creates intricate light and dark regions dominantly caused by particle surface protrusions and their shadows. An increasing particle size results in a larger dark region in the image. After unfolding, these regions are described using the all AMT-scales, where the compound, very complex, pattern of occurrences of dark regions will result in spectrum signatures representing increased complexity. It is AMT's greatest tribute that it is able to characterize such complex light-dark-shadow patterns at all scale levels simultaneously, and to be able to embed this complex information into a 1-dimensional complexity spectrum. This pattern of dark and light regions in gray-scale images has here been shown to be quantitatively related to the particle size and - distribution, but only because of the powerful use of PLS-R employing the full-spectrum advantages well-known from e.g. NIR-spectroscopic experiences [18]. This operative proportionality is partly seen here only because the Theory of Sampling was used to counteract the segregation issues, thereby increasing the 2-dim / 3-dim matching.

CONCLUSIONS

A final, optimized data set consisting of 22 natural sands extracted from different beaches along the western Danish coastline were subjected to a new image analytical technique, Image Analytical Sampling, for simultaneously characterizing the salient spanning particle size classes as well as D_{50}.

The technique was first developed through optimization of the experimental imaging parameters. For the final modeling and prediction experiments a QImaging camera and a LHS-500 lamp with unilateral illumination at an angle at 28° was used. The red color channel was chosen for subsequent multivariate AMT + PLS-regression. The results were clearly improved *after* removal of a linear illumination gradient on all images. Such a gradient should always be avoided because it increases computational demands severely. However, for studies on particle size/shape, a unilateral oblique illumination source must be used, since this illumination setup is necessary for optimal definition of sample surface morphologies.

Implementing the Theory of Sampling (TOS) was demonstrated to be critical for the type of aggregates (sand) samples examined; the prediction models were clearly improved by use of a larger surface area (composite sampling), expressly counteracting segregation of each sample as it was presented (poured) to the camera.

Acceptable, indeed very good prediction models were obtained for the following parameters:

- D_{50}
- Size classes: 1.400, 1.000, 0.710, 0.500, 0.355, and 0.255 mm.

The remainder of the 16 original size classes can most likely be modeled with equal success subject only to a better calibration span that what was available in this naturally generated data set. There is no loss of generality by natural sands not necessarily displaying a full grain size span in each of all 16 size classes routinely sieved for in the sedimentological laboratory. The salient conclusion is two-fold: 1. IA/AMT is fully successful for classes falling within the specific grain size ratio brackets for the material at hand. 2. There are limits for application of this new technique, most specifically an average grain-size contrast ratio bracket, which is material-dependent. These findings are

identical to those earlier presented but using acoustic chemometrics [20,26]. Careful calibration, using fully representative sample training sets, is the only way forward pushing this envelope.

The developed image analytical technique has shown its full applicability to particle size determination in general (size-fraction mixtures), but this is not a conclusion specific only for sand(s). The conclusions can be carried over and even expanded into several other scientific and technological research areas as well as industrial production sectors for a range of solids and aggregate materials, e.g. in powder technology and bulk materials handling, where characterization in which grain-size distribution or size-fraction mixing determination is on the agenda. Because the AMT algorithm always extracts all potentially interesting scales from an image, to be calibrated for one or more Y-features of interest in each specific application case, it will be easy to expand this technique to such applications. During the last five years of AMT applications, different parameters have been modeled for, such as: powder flow properties, growth rate, mixture composition and water content, which have all been successfully modeled for materials including solids, suspensions and liquids [4-5,16,23-25]. The carrying-over potential would appear significant.

ACKNOWLEDGEMENTS

We express our most sincere thanks to the Danish Coastal Inspectorate (DCI), *Søren Bjerre-Mikkelsen* & *Axel Toksvig* for permission to use the original set of 60 natural coastal sands from DCI's extensive reference sand library. For complete objectivity, all samples were selected by DCI exclusively (SB-M), following our request for "a widely spanning, representative sand sample set illustrating the natural sorting and other aggregate characteristics found in beach sands" (the present authors specifically wanted no influence on the selection of the basic data set).

References

[1] N. Zuech and R. K. Miller, Machine Vision Kluwer Academic Publisher 1989.

[2] J. C. Russ, The Image Processing Handbook, (4 edn) CRC Press 2002.

[3] R. Jain, R. Kasturi, and B. G. Schunck, Machine Vision McGraw-Hill 1995.

[4] J. Huang and K. H. Esbensen, Chemometrics and Intelligent Laboratory Systems, 57 (2001) 37-56.

[5] J. Huang, Developments in Applied Chemometrics - AMT, Acoustic Chemometrics and N-way Analysis, Ph.D. thesis, the Norwegian University of Science and Technology (NTNU), ISBN: 82-7984-187-3 (2001).

[6] K. H. Esbensen and P. Geladi, Chemometrics and Intelligent Laboratory Systems, 7 (1989) 67-86.

[7] P. Geladi and H. Grahn, Multivariate Image Analysis, John Wiley and Sons Ltd. 1996.

[8] P. Gy, Sampling for analytical purposes, John Wiley and Sons Ltd. 1996.

[9] F. F. Pitard, Pierre Gy's sampling theory and sampling practice, (2 edn) CRC Press Ltd. 1993.

[10] L. Petersen, P. Minkkinen, and K. H. Esbensen, Chemometrics and Intelligent Laboratory Systems, 77 (2005) 261-277.

[11] L. Petersen, C. K. Dahl, and K. H. Esbensen, Chemometrics and Intelligent Laboratory Systems, 74 (2004) 95-114.

[12] P. Zamperoni, Chemische Technologie der Cellulose, 43 (1989) 133-143.

[13] The Quantitative Imaging Group, Delft University, *Shading Correction*. http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip-Shading-2.html, 27 Aug. 2004.

[14] R. Andrle, Mathematical Geology, 26 (1994) 83-97.

[15] B. K. Alsberg and B. G. Remseth, Journal of Chemometrics, 6 (2005) 135-150.

[16] P. P. Mortensen and K. H. Esbensen, Chemometrics and Intelligent Laboratory Systems, 75 (2005) 219-229.

[17] K. H. Esbensen, Multivariate Data Analysis - In Practice, (5 edn) CAMO ASA 2002.

[18] H. Martens and T. Næs, Multivariate Calibration, (1 edn) John Wiley and Sons 1989.

[19] P. Geladi and B. R. Kowalski, Analytica Chimica Acta, 185 (1986) 1-17.

[20] M. Halstensen, Acoustic Chemometrics - Experimental multivariate sensor technology and developement of system prototypes for industrial multi-phase characterisation: selected forays, Ph.D. thesis, HIT, 2001.

[21] P. Gy, Chemometrics and Intelligent Laboratory Systems, 74 (2004) 39-47.

[22] L. Petersen & K. H. Esbensen, Journal of Chemometrics, 19 (2005) 625-647.

[23] C. K. Dahl, L. Petersen, and K. H. Esbensen, Quantitative Moisture Prediction by Image Analysis of Water-Induced Powder Agglomeration - AMT (Angle Measure Technique), in Ph.D. thesis (Casper K. Dahl).

[24] J. Huang and K. H. Esbensen, Chemometrics and Intelligent Laboratory Systems, 54 (2000) 1-19.

[25] K. H. Esbensen, K. H. Hjelmen, and K. Kvaal, Journal of Chemometrics, 10 (1996) 569-590.

[26] Halstensen, M. & K. H. Esbensen, Journal of Chemometrics, 14 (2000) 463-481.

Paper V

Lars Petersen*, Casper K. Dahl and Kim H. Esbensen
"Representative Mass Reduction in Sampling - A Critical Survey of Techniques and Hardware"

Chemometrics and Intelligent Laboratory Systems
Volume 74, issue 1, pp. 95-114, 2004.

Available online at www.sciencedirect.com
SCIENCE DIRECT
Chemometrics and Intelligent Laboratory Systems 74 (2004) 95–114
Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

Representative mass reduction in sampling—a critical survey of techniques and hardware

Lars Petersen*, Casper K. Dahl, Kim H. Esbensen

Applied Chemometrics, Analytical Chemistry and Sampling Research Group, ACACSRG, Aalborg University Esbjerg, Niels Bohrs Vej 8, DK-6700 Esbjerg, Denmark

Received 3 September 2003; received in revised form 16 February 2004; accepted 1 March 2004
Available online 24 June 2004

Abstract

We here present a comprehensive survey of current mass reduction principles and hardware available in the current market. We conduct a rigorous comparison study of the performance of 17 field and/or laboratory instruments or methods which are quantitatively characterized (and ranked) for accuracy (bias), reproducibility (precision), material loss (external as well as internal loss), user-dependency, operation time, and ease of cleaning. Graphical comparison of these quantitative results allow a complete overview of the relative strengths and weaknesses of riffle splitters, various rotational dividers, the Boerner Divider, the "spoon method", alternate/fractional shoveling and grab sampling.

Only devices based on riffle splitting principles (static or rotational) passes the ultimate representativity test (with minor, but significant relative differences). Grab sampling, the overwhelmingly most often used mass reduction method, performs appallingly—its use must be discontinued (with the singular exception for completely homogenized fine powders). Only proper mass reduction (i.e. carried out in complete compliance with all appropriate design principles, maintenance and cleaning rules) can always be representative in the full Theory of Sampling (TOS) sense. This survey also allows empirical verification of the merits of the famous "Gy's formula" for order-of-magnitude estimation of the Fundamental Sampling Error (FSE).

Keywords: Mass reduction; Sampling; Riffle splitter; Shoveling; Boerner Divider; Rotational divider; Grab sampling; Representativeness

Sampling is nothing but representative mass reduction. [Pierre Gy]

1. Introduction

The archetype error of ill-reflected sampling is to focus on getting to the final sample volume much too early in the sampling process. Instead of only focusing on securing as quickly as possible the desired representative samples (which cannot be evidenced from the physical samples themselves) of the final sample volume/mass, the Theory of Sampling (TOS) stipulates that only a properly designed and controlled sampling process can facilitate this. Only TOS tells comprehensively how and how much material to extract from a lot. For many types of heterogeneous material often the extracted primary sample has to be of a quite substantial size in order to be representative, and this places stringent demands on the sampler (the sampling instrument), for instance if the sample is used for chemical analysis, where typically only 1 g, or a fraction hereof is required. Usually, there is a very long way from the size of the initial lot—via the primary sample—to the final analytical sample mass (Fig. 1). Typically mass reductions of the order of 1:1000 to 1:100.000 have to be invoked. It is therefore of the utmost importance that all sampling processes make do only with representative mass reduction. Unfortunately many designs and implemented hardware look at mass reduction as a pure materials handling reduction in terms of weight per se. It's quite another thing to be concerned with the degree of representativity of the reduced mass fractions.

Also, usually emphasis is on getting a valid analytical result, in the sense that the amount of the analyte in the final sample, a_S, makes do—while TOS emphasizes that only the corresponding estimate of the lot concentration, a_L, carries the information sought. There is a world of difference between these two concentration estimates—the entire

* Corresponding author. Tel.: +45-7912-7666; fax: +45-7545-3643.
E-mail address: lape@aue.auc.dk (L. Petersen).

0169-7439/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2004.03.020

Fig. 1. Do not focus on directly getting the final analytical volume too early in the sampling process—representative mass reduction does all the work.

1:100 to 1:100.000 mass reduction lies in-between! In the present work we focus on how to reduce the size of any sample (lot, or primary sample) without sacrificing the crucial representativity prior to analysis.

Here we shall not discuss the issues concerning how to do the primary sampling, as this is amply covered in the basic sampling literature [1–14]. Here we are exclusively oriented towards the subsequent mass reduction process(es) involved (principles, methods and hardware: design and maintenance). Even when the extracted primary sample is representative of the lot, it will still be up to the subsequent mass reduction process whether the secondary-, tertiary-, laboratory- and instrumental sample preparation sub-sampling leads to the desired results or not, i.e. whether the mass reduction is "correct" or not in the full TOS sense [1–3,14].

We intend to show that making sure that all mass reduction steps are correct allows for a certain indispensable freedom in the sampling process in the sense that one is now free to make the primary sample mass, MS, of any size necessary (due to the heterogeneity of the material, etc.). This means that having to take a large primary sample is no longer a problem. One simply has to reduce this mass before transportation, storage or analysis in order to save time and money. Thus proper mass reduction comes to the fore at all stages in a compound sampling/mass reduction staged process. The principles and procedures examined here are all operational over this entire range, from reducing the primary sample (orders of magnitude span 1 kg–1 ton, or more) all the way down to an analytical mass of the order of grams, micrograms or even smaller.

In slightly more detail: In order for all mass reduction methods or devices to work properly it is critical to respect all the key principles of TOS, primarily that all constituent fragments of the lot must have equal, non-zero, probabilities of ending up in the final sample. This necessitates complete randomness in the selection process of the constituent fragments (units, groups or sub-samples). We here refer to the literature on proper sampling [1–14].

The present paper focuses on 17 current methods and devices commonly used for mass reduction, which have been tested and assessed with regard to a number of characterizing parameters, among which the most prominent are accuracy and reproducibility (precision), constituting the definition of representativity [1,2,15–17]. But in the present comparison study we are in fact interested in the quality of both the average composition estimates resulting from mass reduction operations as well as in the variances of repeated assessments of the performances of the various instruments employed (replicating the entire mass reduction process). Also, other, more practically related parameters such as operating time consumption, user-dependency and device cleaning requirements, etc., are included in the final overall presented below.

This study is complementary to the one by Gerlach et al. [17], who performed a survey of five field-sampling techniques. Gerlach et al. was interested in testing robust, quick and efficient methods for soil splitting in the field (methods included were riffle splitting, paper cone splitting, fractional shoveling, coning and quartering and grab sampling, three of which are also covered here), whereas we are more oriented towards major undertakings associated with industrial and routine laboratory sampling in general. One major difference is that whereas Gerlach et al. only used synthetic samples, we use naturally occurring materials making up 99.90% of all compositions investigated.

2. Material system and analytical procedures

Which material system for comparison purposes would be optimal? Should the material system reflect one dominant situation (necessarily with a relatively smaller range of potential applications fields) or should one strive for as general a material system as possible? What would constitute the latter? This issue is intimately related to the very purpose of mass reduction—here mass splitting is specifically used for the purpose of representative sampling, so the possibility to make generalizations from our survey is of prime importance. Accordingly we have laid down the following criteria for the design of an optimal comparison material system:

(1) The system must reflect both major concentrations, intermediate as well as trace concentrations. For this purpose we have chosen the following levels: 89.9%, 10.0% and 0.1% (1000 ppm).
(2) The material system must be sensitive to flow segregation (indeed also to all other manipulation segregations as far as possible: roll segregation, etc.). This is in order for the system to exhibit a significant degree of segregation as an inherent part of the mass reduction process. We have chosen one component (0.1%) with a very smooth surface (the trace concentration component), one smooth component (10.0%)

and one with slightly softer surface characteristics (89.9%).

(3) It is equally important in the present context that at least some (one, two) of the chosen components also show a significant propensity for "rebounding" when impacting on hard surfaces, as this is an inherent weakness in the design of some mass reduction tools (while being better counteracted by others).

We have stipulated these requirements in order for the comparison system to represent a fair *worst case scenario*; we wanted to test the 17 approaches to be compared exclusively from the point of view of their performance in such a realistic, difficult situation. It is of course trivial to generalize to less adverse situations.

The result was a system of mixed wheat grains, rape seeds and glass spheres, with concentrations 89.9%, 10.0% and 0.1% w/w, respectively. We deliberately chose glass spheres as the trace component, in order to represent, e.g. an impurity (an artifact component), so we did not object to this being an artificial component. We also took great care in designing a material system in which the average grain size, and density, for all three components were not significantly contrasted, in order not to end up in pathological situations (extreme size and/or density agitation segregation). The average grain sizes of wheat, rape seed and glass spheres were: 6.0 (by 3.0 as a "cylinder"), 2.6 and 1.0 mm, respectively. Their average densities were: 0.75, 0.77 and 2.60 g/cm^3. We believe that the chosen system does a good job standing in for a very wide range of industrial and laboratory material systems of aggregate materials and powders with respect to these physical design characteristics. It is admittedly very sensitive for flow segregation, but so much the better when the objective is to test the performance of purported universal mass reduction tools.

Mixing of this lot material prior to all mass reduction experiments (always carefully weighed in completely identical proportions) was carried out by randomly shaking a plastic bucket for 2 min (mechanical shaking and mixing). A lot mass of 2 kg were to be mass reduced to get either 100 or 125 g in the final sample, depending on the nature of the method or device (i.e. dependent upon which split ratios could be obtained with the specific methods). After every pass of mass reduction, the composition of the resulting subsamples was determined, using a screening system consisting of two sieves and a bottom collecting pan, all mounted on a shaking table (Fig. 2), which collected the wheat, rape seed and glass, respectively. The screen sizes were 2.8 and 1.5 mm. We initially performed a set of screening verification experiments; the results showed that the efficiency of separating the three components used was completely satisfactory since the three components were fully separated.

After separating the different fractions of the final reduced samples—as well as the very important fractions of the left-over material (i.e. material rebounded out of the receptacle bins, etc.) were weighed individually by a laboratory analytical weight. Weighing was chosen as "analysis" because of the minimal error associated with this, estimated at 0.01% relative. The masses were used to calculate the analytical result, a_S.

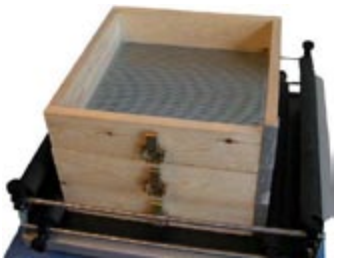


Fig. 2. The screening system.

The same mass reduction/sub-sampling/weighing procedure was repeated 20 times in blocks of 10 by two operators (the two first authors) for all methods and devices investigated. A replication rate of 20 allows for highly trustworthy statistics, which is deemed necessary in order to reach significant conclusions as to reliable, accurate and precise comparability and ranking. Inclusion of two operators in all experiments represents inclusion of inter-operator errors in the overall mass reduction errors estimated in our survey, adding to the validity of a most realistic working situation. If anything, the experimental setup was stacked to reflect a (very) difficult situation indeed.

3. Devices and methods

3.1. Riffle splitting

The most well-founded method for mass reduction is riffle splitting. Riffle splitters can be constructed in several different ways, of which many are in accord with TOS principles and equally many are not. If designed and used correctly it provides a very stable, reliable and inexpensive method for mass reduction with reasonable speed.

3.1.1. Principle

The general principle is that the sample to be divided is introduced to a rectangular area, divided by parallel chutes leading to two separate receptacles. For this device to work properly it must be designed according to a few essential rules. There have to be an equal number of chutes of which every second leads to the two alternate receptacles. The chutes must all have the same size and form; the wall material must be thin in relation to the wall-to-wall dimensions of the chutes themselves. It is also important that no chute can be over-represented when introducing the sample

into the device, for instance by a non-correct design of the sample holders or by a cone-shaped inlet collar in the longitudinal direction. The higher the number of chutes, the better the device splits the sample, both in terms of the splitting bias between the two splits as well as with regard to the variance of repeated splits, as shall be amply demonstrated below. The width of the chutes also has to have a certain minimum width which depends on the particle size, in order to prevent blocking (large particles) or bridging (powders) [18]. An empirical rule-of-thumb stipulates that chutes must be wider than three times the maximum particle size or two times this plus 5 mm, since even extremely small particles should not be split using smaller chute width than 5 mm. The general literature on TOS has exhaustive analysis and discussions of correct design principles of riffle splitters, which must be consulted before acquisition of any riffle splitter [1,2].

3.1.2. Use of riffle splitters

It is, however, not enough to have access to a correctly designed riffle splitter. In order to obtain a representative mass reduction, the device also has to be used—and indeed cleaned and maintained—correctly. There are a few simple rules that must be followed, which may be summarized as follows:

1. The sample must be spread out equally over the whole length of the feeding tray.
2. The feeding tray must have exactly the same width as the rectangular receiving region of splitter; there is thus no need for inclined inlet collars, etc., in the longitudinal direction.
3. The sample must be fed perpendicularly to the longitudinal axis along the device; the sample must be fed precisely on to the center axis.
4. No particles can be allowed to bounce out of the receiving trays or the splitter.
5. The split sample (or the portion to be split further) must be chosen at random.

If these rules are obeyed, any split portion should (in theory) not be systematically biased by the splitter. Fig. 3 shows some of the errors than can result from incorrect design and use of riffle splitters. To understand the importance of the design it is important to remember that even though the sample is evenly spread over the width of the feeding tray, it cannot in practice become homogeneous and this will lead to (minor) differences between the feed for the individual chutes.

3.1.3. Device description

In the present work six different riffle splitters of the basic design described above were tested. During the experimental runs several optimizations on existing devices and the design of a new device took place. This is described in further detail in a later section. The splitters used are named according to design and for easier distinction as follows:

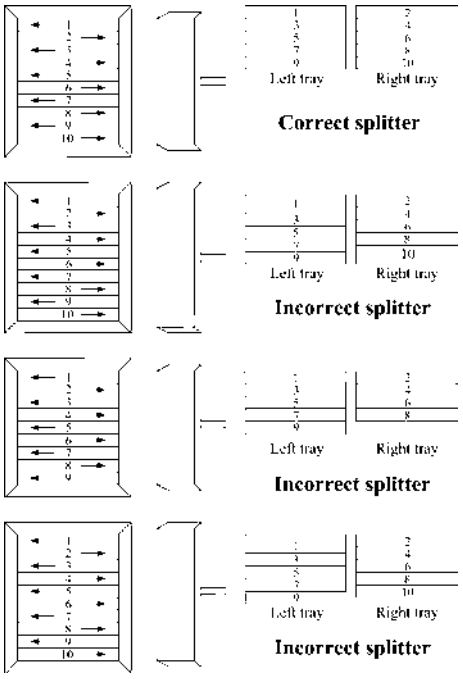


Fig. 3. Schematic illustration of the critical importance of correct riffle splitter design.

- The animal feed splitter
- The seed splitter
- RK 10 chutes/20 mm width splitter
- RK 10 chutes/30 mm width splitter
- RK 18 chutes/16 mm width splitter
- RK 34 chutes/10 mm width splitter

The latter four are manufactured by the same company and three of these are designed in an exactly identical fashion, only scaled-up. The 34 chute splitter differs, since it represents a completely new design resulting from the present work. In the following sections the individual splitters are further described.

3.1.4. The animal feed splitter

This divider (Fig. 4) has 10 chutes and is used by the Danish Ministry of Agriculture's department of animal feed testing. The chutes are 27 mm wide. The design has several apparent errors, but also some advantages. The device has three identical trays, two used for receiving and one for

Fig. 4. The animal feed splitter.

feeding. The trays can be switched freely, making handling easier. The greatest advantage is that it can be taken apart for easy cleaning.

One of the obvious errors is found in the design of the trays, since these are narrower than the top of the divider and the section for the receiving trays. This makes the introduction of the sample to the divider and the reception of the sub-samples unavoidable sources of bias. If the reception trays are placed wrongly, which easily happens, some of the material is lost completely since it does not even hit the reception trays at all. Fig. 5 shows a close-up of the side of the divider when one of the receiving trays is placed wrongly. It is observed that almost half the material hitting the uttermost chute will be lost.

Another error is the open design of the top of the splitter, where many particles (especially rape seed) are observed to bounce out of the device during operation. The Feed Splitter is greatly dependent on the user because of the pouring mechanism. This cannot be avoided in the current design.

3.1.5. The seed splitter

This device (Fig. 6) is used by the Danish Ministry of Agriculture's department of seed testing and has 20 chutes

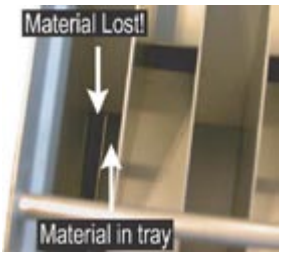


Fig. 5. Unwanted design error for the animal feed riffle splitter. If the tray is placed wrongly, as is very easy, almost half the material hitting the most peripheral chutes is lost.

Fig. 6. The seed splitter.

of width 10 mm. The design again has errors, but also some good features.

The errors consist of the open design of the section between the feeding tray and the chutes. Much material is lost in this region due to particles bouncing out of the splitter. The advantages are found in the feeding mechanism, which makes handling easy and also minimizes the possibility of operator introduced errors. This splitter can also be taken apart for easy cleaning. The relatively high number of chutes (20) makes the splitter more reliable than the previous one. A nice detail is that the leading edges of the blades between the chutes are sharpened to minimize particle bouncing. The feeding tray is nicely aligned with the sides of the splitter and no error is thus induced from this.

3.1.6. The RK splitters

Three of the four splitters from "Rationel Kornservice A/S" (RK) are designed from the same basic principles. The only features changing are the number of chutes, the width of these chutes and the resulting overall dimensions of the devices. The splitters consist of two separate reception trays, a splitter and a swinging, mounted feeding tray (which can be easily dismounted however if need arises). These splitters are delivered with extra plates for insertion over the chutes to minimize sample loss due to bouncing. The reason that these insertion plates are not permanently installed is only that the splitters are also used for grass that has a tendency to clog up the device if this inner clearing is too narrow. These plates were installed on all the splitters used in the current experiments, except the RK 10 chute (20 mm) and RK 34 chute (10 mm) splitters.

There are errors in this design too. The first relate to the fact that the feeding trays are slightly narrower than the top of the splitters. This, however, is possibly only of marginal importance, since the error is the same in both distal ends, and thus both reservoirs are underrepresented from the outmost chutes by approximately the same amount. The splitter is not easily cleaned since it cannot be taken apart. The advantage on the other hand is equally obvious since a

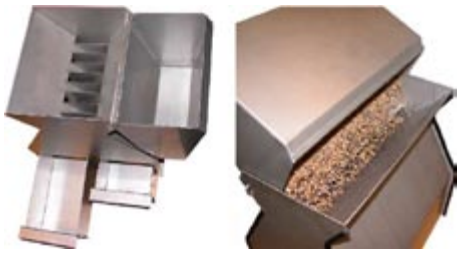


Fig. 7. The 10 chutes/20 mm width splitter (left) and the 18 chutes/16 mm width splitter in action (right). Notice the closed design resulting in minimal material loss.

minimum of particles are lost during use because of the closed design; another advantage would be that the sample is presented to the splitter in exactly the same way every time, because of the fixed feeding tray. Fig. 7 shows the principle design and use of these RK splitters.

The last of the four RK splitters is a result of an ongoing collaboration between ACACSRG and Rationel Kornservice A/S to build a better splitter. The new design has several improvements, mainly in the increased number of chutes (34) and the optimized feeding tray. In the next section this device is described in detail.

3.1.7. RK 34—device optimization

The new design differs from the old mainly in the feeding device and the number of chutes. Designing the new feeding mechanism started by checking out the idea of using the same principle of a mounted feeding tray as the seed splitter. This turned out to be a very constructive idea and the design process continued on this basis. The solution sought had to eliminate the step of pouring material from one of the previous receiving trays into the feeding mechanism, since this introduced rolling and trajectory segregation as well as impact effects. When spreading out the material as evenly as possible in the feeding trays, these effects lead to bias of the results. The final solution was simply to furnish the feeding mechanism with a slot for directly inserting one of the previous used receiving trays (Fig. 8).

When tipped, the sample is poured into the conically inclined interior of the feeding device. This minimizes the effect of segregation drastically. The conical delivery funnel only opens when the tray is tipped all the way and gently brought in contact with the chute area in the splitter due to a small activation/stopper pin. Also, the width of the opening can be regulated by this controlling pin (Fig. 9).

Several lengths of the feeding funnel were tested out, to see if a shorter delivery path along which the particles can segregate would increase splitter precision of repeated operations. This was not the case and therefore the final design was kept to minimize size and weight configuration for both economical and practical reasons. The resulting

Fig. 8. The final RK 34 design, using an insertion slot for a third tray in the feeding device.

splitter looks exactly as in Fig. 8, but with a shorter feeding funnel. As in the previous design there is indeed in principle introduced a very small error since the feeding tray is slightly shorter than the splitter length by the width of two individual chute walls; this would appear almost totally negligible however, since in the longitudinal direction this foreshortening amounts to 1.6 mm/359.6 mm, or 0.4% only. Since we were in fact unable to demonstrate any effect of this error, the design was consequently kept as is. The results of the present overall survey were not known to us at the time when we decided to stop the development process of the present apparatus. Only later it was learned that this prototype RK 34 riffle splitter indeed outperformed all other riffle splitters in the present study, so this error truly must be exceedingly small. For more homogeneous systems the old (long) feeding trays can still be mounted, since these fit into the same socket as the new one. This also opens the possibility of changing to the new feeding mechanism on existing dividers, etc. RK has since the ending of the present work declared that all future splitters (all sizes and models) will be built according to this new design.

Fig. 9. The feeding tray opens "automatically" when tipped and gently pressed against the chutes. This takes place exactly at the center longitudinal axis. The small pin is inserted for controlling the width of the opening.

3.1.8. Experimental run

The riffle splitter experiments were carried out in identical fashions for all methods. The 2 kg were divided four times to get approximately 125 g in the final sample. After each split the tray to be further divided, or the final sample, was selected at random. The practical procedures for the individual splitters were necessarily a little problem-dependent, but all efforts were made to follow the rules of correct use. The material was poured into the feeding tray and spread evenly over the entire width with the greatest care in order to allow all methods to compete evenly and fairly. The material was then fed along the center axis of the chutes with a steady flow in order to minimize bouncing, especially of rape seed. All RK riffle splitters tested were fast and easy to use. The three older designed splitters can all be somewhat user-dependent, since pouring the sample into the feeding tray can vary. By using the improved design this user-dependency is minimized or fully eliminated.

3.2. Revolving splitters

Revolving splitters are based on the same principle as riffle splitters. The revolving feeder distributes the sample material equally (in time) over a number of radial chutes, assuming constant rotational speed. These devices are very easy to use, since one only needs to pour the material to be divided into a hopper, thereby getting one or several reduced splits. They also require very little preparation and clean-up and split the sample very fast. The latter of course depends on the rotating speed as well as the influx velocity of the material through the feeding funnel. The essential principle here is that every second radial chute contributes to one of the two alternative collecting reservoirs.

Since a larger number of sub-samples in this context give more representative samples it can be desired to increase the total number of revolutions. This can be done by using a smaller outlet size of the hopper. Also, size, slope and rotation speed of the inlet tube can be altered to change the outlet speed.

3.2.1. Vario Divider

With a revolving variable sample divider named "Vario Divider" (Fig. 10), it is possible to get a mass reduction ratio as small as 1:100, depending on the nature (mainly particle size) of the material to be split. Some models include the possibility of getting several final samples.

The lot material is poured into a hopper (1) from where it is led to a revolving feeder (2). From here the material is led either to a chute opening (11) or to the bottom as left-over material (9). The divider shown gives two equivalent samples (7 and 8).

When using the model type 1G/1–4 on the current sample composition, it is only possible to get a sampling ratio of 1:9 because of the particle sizes involved. To obtain a sample of 100 g it is necessary to realize a sampling ratio of 1:20. The sample therefore had to be divided in at least two steps, which allows for a realistic testing of the Vario dividing principle.

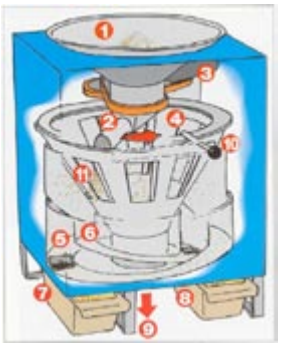


Fig. 10. Vario Divider. 1: hopper, 2: revolving feeder, 3: motor, 4: chute closer, 5 and 6: sample outlets, 7 and 8: samples, 9: left-over material, 10: chute opening handle and, 11: chute [19].

Two different multiple-step settings were tested: sampling ratios of 1:4 + 1:5 and 1:2 + 1:2 + 1:5. Splitting the sample in three steps is of course more time demanding; however, we found this to be only minimal in practical terms. In general, using more steps results in a larger sampling error since every single step is error generating. Again, this is setting the comparison study with the most stringent demands on the performance for this device type. Only a marginal sample loss was observed. This was exclusively caused by rebounding of the rape seeds from the sample boxes. Therefore almost no maintenance or clean-up was necessary, and the method must be classified as very easy and fast to use. Since the user only has to pour the sample into a hopper, neither any user-dependency is observed.

There has been a certain theoretical discussion regarding the possibilities of such devices to deliver correct (representative) split samples, mainly related to the variable portion of left-over material—we here refrain from entering into this discussion, but are the more happy to include examples of these Vario Dividers into the set of devices to be compared and ranked. We will let the empirical performance of these revolving riffle splitters speak for itself.

3.2.2. 32-Divider (fixed ratio)

With the 32-Divider (fixed ratio), Fig. 11, it is possible to get the lot material divided into 32 supposedly equal subsamples (so the design objective claims).

The principle used for this divider is identical to that for the Vario Divider, but without any variability whatsoever. The whole cross-section area is divided into 32 fixed chutes, so that the lot material is split completely, and there is no left-over material at all. The constant rotation of the revolving feeder causes the lot material to be equally divided amongst the 32 chutes giving 32 sub-samples.

Fig. 11. The 32-Divider.

Two of the 32 sub-samples were joined into one composite sample of approximately 125 g; these two sub-samples were always selected at random from the 32 candidates. The tested divider was fairly easy to use, but required the user to attach plastic bags to each of the 32 tubes. This did of course affect the preparation time for each split, but since this is only a practical, and easily solvable problem, it should not be taken into serious account. Besides this minor attachment work it was regarded as a fast and easy mass divider to use, with no maintenance. If the sub-samples are extracted at random, neither systematic differences nor user-dependencies are expected to influence the splitting results.

3.3. Shoveling methods

3.3.1. Spoon Method

The Spoon Method can be used to achieve relatively low sampling ratios. This method is somewhat related to the principles behind bed blending, but weakly at best [1,2,20]. The lot material is spread out in an even layer: the lot is poured out on a flat surface as a thin string in an s-shape movement in one direction. This is subsequently repeated in the traverse (90° rotated) direction, then this procedure is repeated again and so forth until all of the lot is poured into the tray, as shown in Fig. 12.

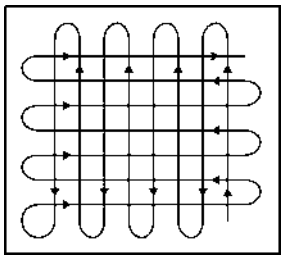


Fig. 12. The pattern used for spreading the lot material in the tray.

Fig. 13. Extraction of sample using the spoon method.

After the laying out step, the spoon method simply consists in extracting a "sample" by using a spatula and a small spoon ensuring that the bottom of the tray is reached (Fig. 13). Several sub-samples, increments, are extracted and joined as the final sample, reducing the effect of the grouping factor [1,2]:

$$\gamma = \frac{N_F - N_G}{N_G - 1}$$

where N_F is the number of fragments in the lot and N_G is the number of groups, or increments, in the lot.

It is important that the final sample is composed of as many increments as possible. To make sure that uncompromised increments can in fact be extracted, and thereby reducing any incorrect sampling errors (ISE), the lot material has to be spread out in a layer of a certain thickness. The extracted sub-samples have to be extracted completely at random from the whole lot to get a probabilistic sample. The method is generally time consuming, tedious, and greatly user-dependent. By following the guidelines mentioned above the user-dependency will be somewhat reduced.

In practice, the final sample in the present experiments was composed of five increments (Fig. 14), which resulted in an average mass of 115 g. Some sample loss always occurred due to problems (selective losses) during increment transfer to the sample box.

3.3.2. Alternate shoveling

The alternate shoveling method can be used to split a sample into two samples with almost equal weight, hope-

Fig. 14. Box after random sample extraction by the "spoon method".

Fig. 15. Alternate shoveling.

fully also of almost equal composition. TOS has made thorough analysis of this general approach [1–3], and there are many pitfalls which are almost universally unavoidable in practice. The equality of the final samples obtained by this method will also be highly dependent on the nature of the material sampled.

The method is based on the principle that all extracted shovelfuls from the original sample are deposited sequentially in two alternative heaps as illustrated in Fig. 15.

It is important that all extracted shovelfuls are selected at random from the initial lot and that all increments have the same (approximate) size. Each heap should consist of an equal number of shovelfuls. One heap should only consist of all even-numbered samples while the other should only consist of all odd-numbered samples. By ensuring that all shovelfuls are carefully selected at random, the condition of sampling equity is preserved thereby minimizing the risk of a systematic bias to some degree.

Four full splits were necessary to obtain a final sample on approximately 125 g in our experimental runs. Some sample loss was observed due to the practical handling of shovelfuls. Different samplers (operators) will definitely have unequal impact on the quality of the final reduced samples, since the shoveling can vary greatly from user to user.

3.3.3. Fractional shoveling

With true fractional shoveling it is possible to divide the lot material into N sub-samples instead of only two. Shovelfuls are extracted from the lot material and deposited into N distinct heaps. In Fig. 16, true fractional shoveling with $N=5$ is shown.

Fig. 16. Fractional shoveling with $N=5$.

Fig. 17. Grab sampling.

The shovelfuls should again all be extracted at random from the lot material and should be (approximately) equal in size. Each heap should consist of an equal number of shovelfuls. All extracted shovelfuls should be alternated from heap one to heap *N*.

To reduce the initial 2 kg into approximately 100 g the sample was first split into five heaps. From these five heaps one was chosen randomly and further divided into four samples at approximately 100 g each. This method can also be slow, tedious and user-dependent.

3.3.4. Grab sampling

Grab sampling is the easy choice for extracting a "sample" and is (unfortunately) very often the preferred choice in practical sampling situations. One sample only is extracted to represent the whole lot. Grab sampling is the archetype sampling error at work. The focus is exclusively on getting the final sample mass directly in one go! However, a grab sample may also result from joining several increments (sub-samples), to produce a composite sample, which in general should result in a more representative sample. Grab samples are typically taken by a

Fig. 18. Boerner Divider.

Fig. 19. Boerner Divider at work.

scoop or a shovel, depending on the size of the original lot, etc.

To obtain a truly representative sample all virtual units making up the lot per force must have the same probability of being selected. With grab sampling this manifestly neither can be, nor hardly ever is, the case. The singular sample to be extracted by grab sampling is of course also often taken from an(y) easily accessible part of a lot, i.e. the top. Grab sampling is therefore often classified as deterministic sampling.

Extracting a sample from the bottom of the lot in Fig. 17 was difficult as in most cases (extraction shown in the figure). The sample domain was divided into six virtual areas. Each grab sample was extracted at random from one of these six areas thereby counteracting unwanted systematic bias. It was attempted to extract the sample from the whole virtual area by pushing the scoop to the bottom and withdrawing it slowly upwards.

If such ''samples'' are erroneously accepted, grab sampling is a fast and easy choice. If the final sample is to be a probabilistic sample it can be time consuming, very often difficult or downright impossible to get correctly sampled increments. Lastly grab sampling is always user-dependent. Following TOS, grab sampling is supposed to perform the worst of all alternative mass reduction approaches. It was therefore a natural must for the present comparison purpose, if nothing else as a (bad) benchmark.

3.4. Other methods

3.4.1. Boerner Divider

Using the Boerner Divider (Figs. 18 and 19) it is possible to divide a sample into two app. equal half-splits. The halved samples can iteratively be divided again, etc., until a satisfactory sample size has been reached.

The initial sample is poured into a hopper. When the hopper's bottom-shutter is opened the sample is directed down onto the top of a cone. Since the sample is flowing downwards by gravity it should be spread out evenly in all azimuth directions. At the bottom of the cone the material is led through a number of alternative chutes which are so connected so as to lead the alternate part-streams into the resulting two half-sample splits.

The Boerner Divider is very easy and very fast to use. Set-up and maintenance is straight forward, but some sample loss can occur during use, depending greatly on the nature of the material to be split. Especially bouncing materials such as rape seed and glass pellets can be a problem. To avoid any systematic differences from one side of the splitter (receiving tray) to the other it is also here important to choose one of the two receptacle trays randomly.

The initial 2 kg material lots in the present study had to be divided four times to get a sample size of approximately 125 g. Some loss of material was observed primarily due to the design of the receptacle tray combined with the nature of the glass and rape seed used.

Two different Boerner Dividers were tested; one will be referred to as calibrated, the other as non-calibrated, mean-

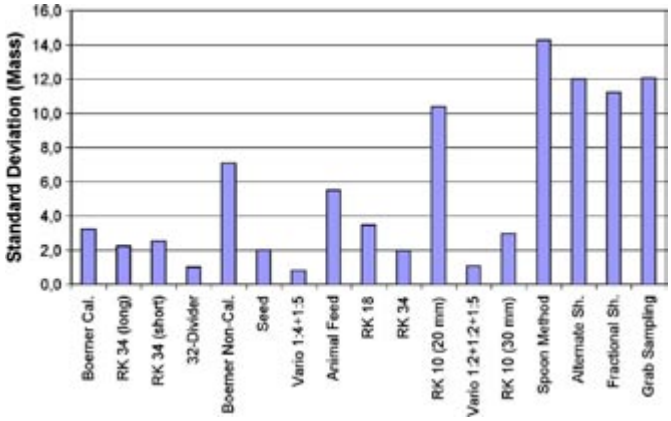


Fig. 20. Standard deviation of the final sample mass.

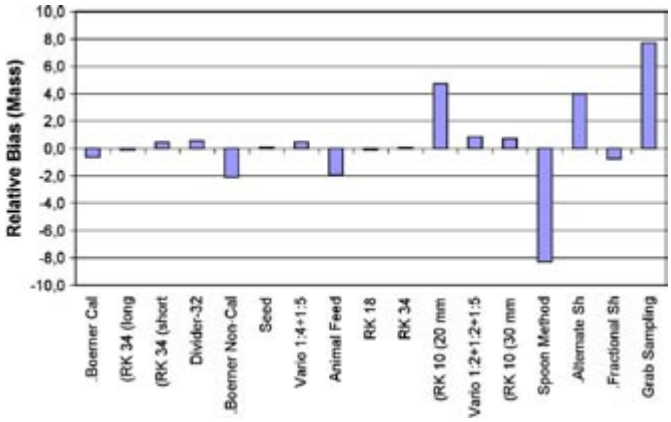


Fig. 21. Relative bias of the final sample mass.

ing that the hopper was slightly off-centered and that the receptacle bins where open. For the non-calibrated divider this could lead to over-representation of some chutes and a fair amount of lost material. Figs. 18 and 19 show the non-calibrated divider.

4. Results and discussion

4.1. Characterizing parameters

In order to evaluate and compare the above methods a system is developed for characterization by a selected set of mass reduction quality parameters. These parameters are described shortly below.

4.1.1. Mean and bias (concentration)

After performing the splits or the individual mass reductions, the results are characterized by the arithmetic average concentration for all three materials after the universal 20 repetitions. This parameter is very important in practice since it characterizes the method's ability to perform splits and leave the sample with the same composition as the lot material.

The mean is a measure of the accuracy of the method when assessed against the true x_{Lot} (a_{L}). The bias is simply

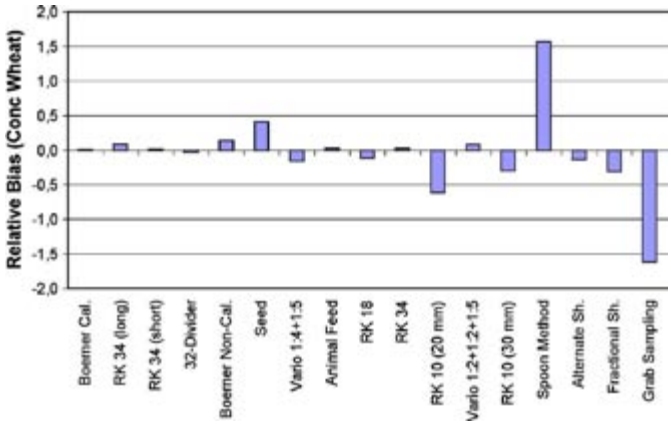


Fig. 22. Relative bias for the concentration of wheat.

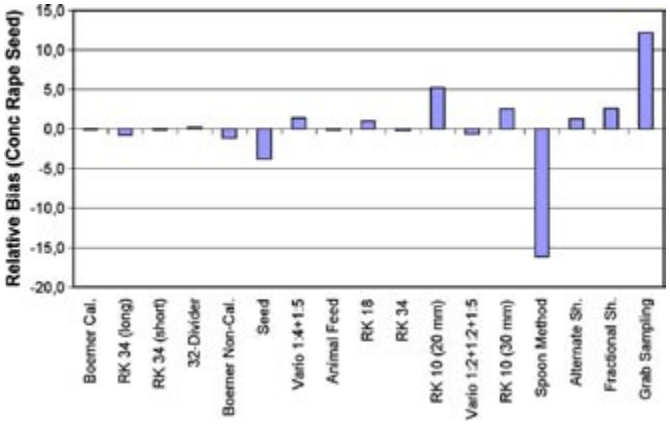


Fig. 23. Relative bias for the concentration of rape seed.

calculated as the average concentration minus the true (known) concentration of the lot, x_{Lot} (or a_{L}). The bias is also presented as a relative value showing the percent wise deviation from 89.9, 10.0 or 0.1% w/w. The relative bias is calculated as:

$$\text{Relative bias} = 100 - \frac{100 \cdot x_{\mathrm{Lot}}}{\bar{x}}$$

4.1.2. Standard deviation (concentration)

In order to characterize the dispersion of the results around the mean, the standard deviations are calculated. The formula used is based on sample statistics (i.e. a statistical sample selected from a large population), since 20 results is not (statistically speaking) really a large number of replicates. The square of the standard deviation, the variance, represents a measure of the reproducibility of the method (precision in statistical terms).

4.1.3. Mean and bias (mass)

The ability to extract the wanted mass of material for the sample is characterized by the mean of the masses. This is an important factor in industrial, laboratory as well as field sampling, where samples often have to be of more or less constant mass.

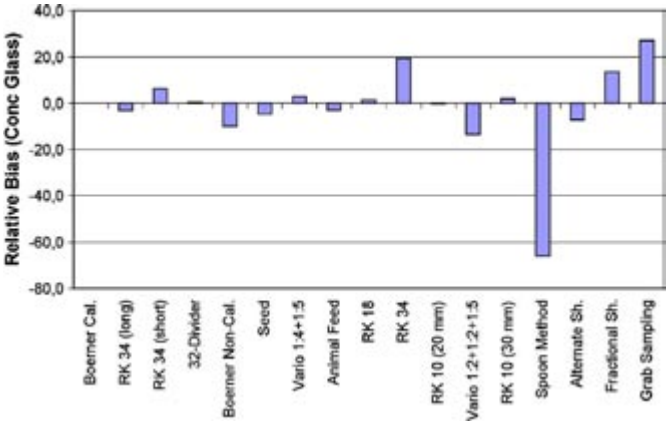


Fig. 24. Relative bias for the concentration of glass.

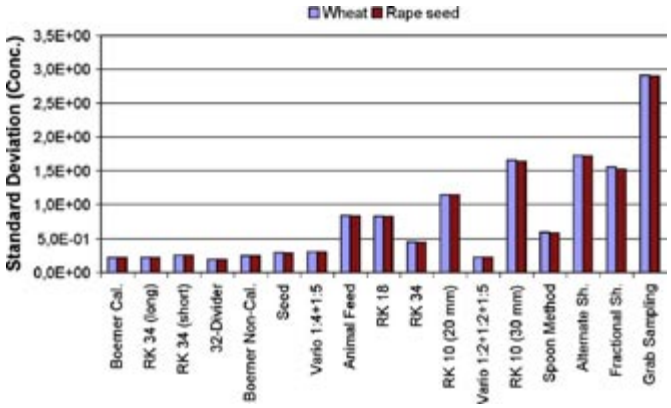


Fig. 25. Standard deviation for the concentration of wheat and rape seed.

4.1.4. Standard deviation (mass)

Again, the dispersion of the individual runs around the mean is presented as the standard deviation.

4.1.5. Loss of material

In order to characterize the method it is also interesting to find out to what degree it causes loss of material during mass reduction. This parameter is also related to the recovered mass of the three fractions in the sample as well as to the residue (material adhering to walls or other surfaces in the apparatus after completed splitting); the residue is estimated as the mass of cleaned-out material. Loss is simply calculated as the initial mass (2 kg) minus the sum of these two recovered weights.

4.1.6. Representativeness

Representativeness is authoritatively defined by Gy [1] and is the only statistic term that includes both accuracy and reproducibility (precision). Many authors, standards and norm writers use the term representative only very loosely and very often without a proper mathematical definition. Normally this characteristic is used for setting a lower limit, with which a representativity statistic for a particular split has to comply in order to be acceptable. In this work

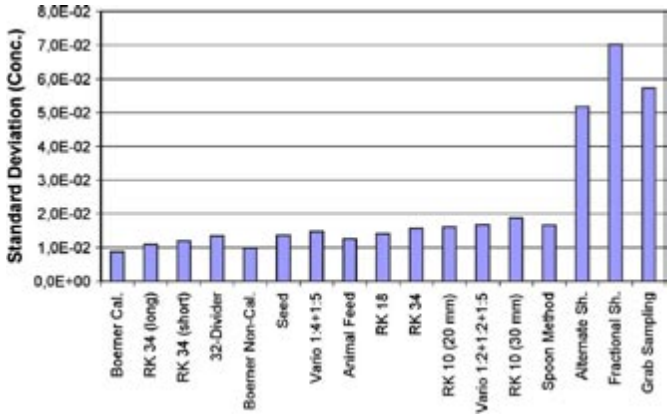


Fig. 26. Standard deviation on the concentration of glass.

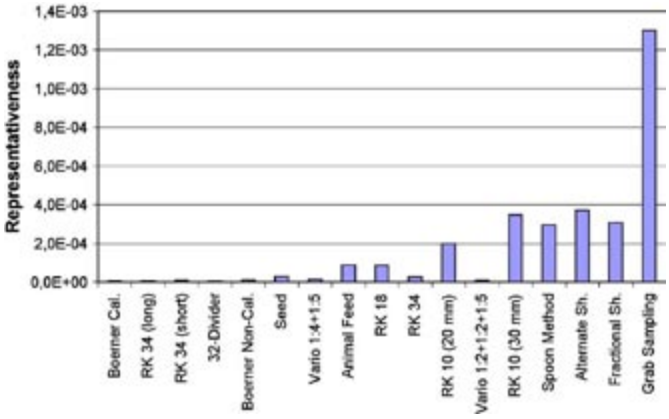


Fig. 27. Representativeness of wheat.

however we use it strictly for comparative purposes. The representativeness is defined as the mean square of the selection error, SE, i.e. the sum of the squares of the mean and standard deviation of the selection error [2]:

$$r^2(\mathrm{SE}) = m^2(\mathrm{SE}) + s^2(\mathrm{SE})$$

where the selection error in turn is defined as:

$$\mathrm{SE} = \frac{a_\mathrm{S} - a_\mathrm{L}}{a_\mathrm{L}}$$

Here a_S is the grade of the critical component in the sample and a_L is the grade of the critical component in the lot material. The latter is known without any uncertainty in our runs, since we carefully prepared the same initial lot composition for each experiment.

4.1.7. Processing time

In order for the method to be attractive, for instance for handling a high throughput of samples (commercial laboratories, government and other regulating bodies, etc.), processing time has to be relatively low. The parameter presented here is simply the average time duration, in seconds, from the beginning of the split until the final sample is at hand.

4.1.8. User-dependency

This characteristic is meant to divide the methods in two types; the first of which is where the person performing the

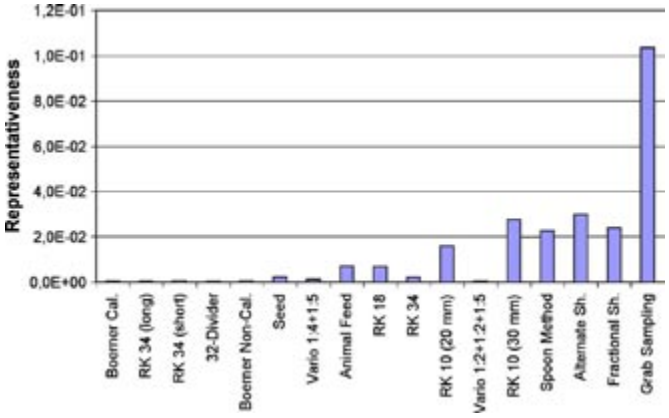


Fig. 28. Representativeness of rape seed.

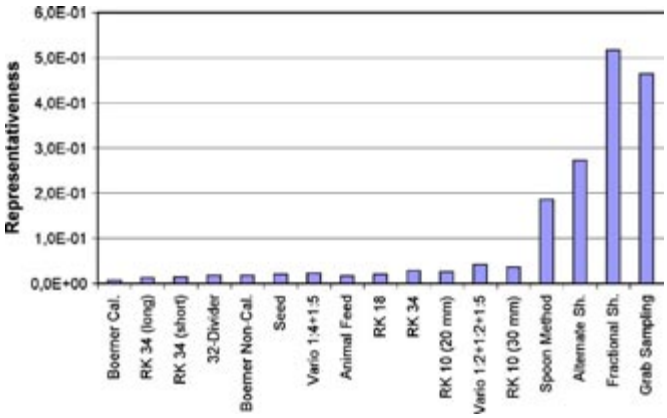


Fig. 29. Representativeness of glass.

sampling has a large influence on the result. The second section is where this influence is minimized if not totally negligible.

4.1.9. Cleaning

Some devices are easily cleaned and others are more difficult to dissemble, etc. This is an overall assessment performed by the authors in every case. It is divided into three categories: easy, intermediate and hard.

4.1.10. Initialization

Another parameter that might be of comparative importance would be the time consumption and workload needed in initializing a mass reduction operation. For instance when the same instrument has to be used for different material systems, or for performing different split ratios, initialization time might be important. Again, this is an assessment performed by the authors and it is divided in three groups: quick, intermediate and long.

4.2. Comparison of mass reduction methods

There are two major fields of interest when characterizing a mass reduction method: the mass of the final sample and the composition resulting from analysis of this mass. The ability of the devices/methods to find the correct (target) mass is summarized in Figs. 20 and 21. We notice that the shoveling methods show a particularly bad accuracy

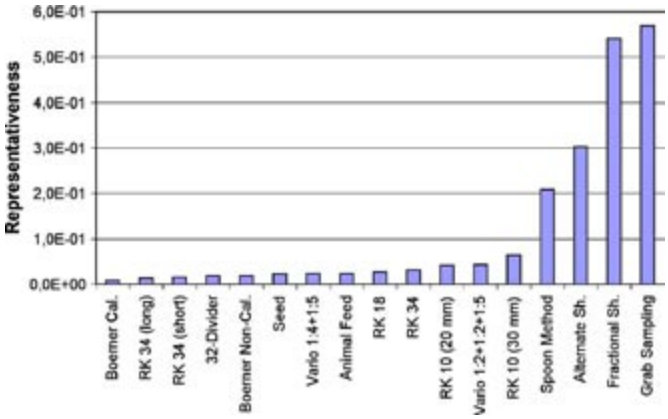


Fig. 30. Sum of representativeness (pooled for wheat, rape seed and glass).

and precision. Only fractional shoveling seems to have a relatively good precision. This method is comparable to alternate shoveling, but differs in having only two mass reduction steps instead of four, which possibly can explain the better precision. The rotational dividers and some of the riffle splitters seem to have good accuracy and precision throughout in finding the target mass. The RK 10 chute (20 mm splitter), however, differ significantly from the rest of the riffle splitters (in an adverse sense). The only possible reason for this must be the missing insertion plates on this splitter.

With regard to the composition of the final sample, the methods are evaluated by the standard deviation, relative bias and the representativeness.

In Figs. 22–24, the methods deviating most from the rest clearly are the spoon method and grab sampling. This is understandable since these two methods are both shoveling methods, and thus expected to be less precise. In general, all shoveling methods and some of the riffle splitters are characterized by bad precision, while all revolving dividers show good precision. The reason for this good precision is the large number of rotations involved, and hence the large number of effective chutes involved in the mass reduction. When the sample takes about 1 min to pass through these devices, the number of revolutions per minute is 40 and the number of openings eight, the effective number of chutes is actually 320 for the Vario Dividers and an impressive 1280 for the 32-divider (which has 32 openings).

The standard deviation is an expression of the precision of a particular method. Fig. 25 shows the standard deviation of wheat and rape seed, which both are present in rather large concentrations in the material (89.9 and 10.0% w/w). All the methods with a large number of chutes or openings have a low standard deviation on both wheat and rape seed. The shoveling methods, again, stand out as terribly imprecise methods, even though the spoon method would appear just within the window for this parameter alone. This is possibly due to the bed blending like preparation of the lot material and the extraction method, which for the experienced operator ensures nicely delimited increments (sub-samples).

Glass is present in very low concentration (0.1% w/w) in the material, and it is expected that the reproducibility (precision) for this material is substantially worse than the components present in larger concentrations. The absolute values in Fig. 26 are not directly comparable with those in Fig. 25, since the standard deviation is a relative value. It is noted that the precision of all the methods is more or less equal for the trace element level, except for the shoveling methods. The spoon method again seems to have an acceptable performance, but the rest of the shoveling methods are distinctly bad, very likely due to the extraction method.

The overall TOS-measure representativeness takes into account both accuracy and precision, and will thus express the overall performance of a method. It is seen in Figs. 27–30 that the methods with the lowest number of chutes or openings and the shoveling methods indeed are worst. This is in accordance with the previous conclusions.

The representativeness of glass has a dominating influence on the pooled sum, since these values are much larger than the values for wheat and rape seed. However, the sum is the best measure for the overall performance of the methods, since it includes constituents present in both high and low concentrations. It is observed from Fig. 30 that the calibrated Boerner Divider has the best overall performance. There is, however, no large difference to be found between

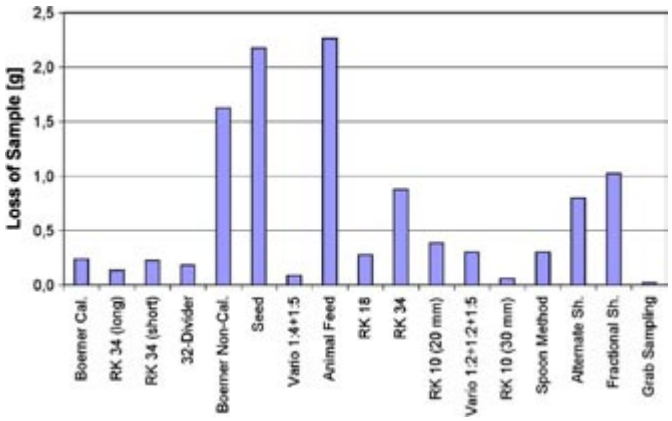


Fig. 31. Total loss of material.

the 10 best of the methods, meaning that all these methods in principle are suitable for mass reduction with regard to representativeness.

The total loss of material during mass reduction is seen in Fig. 31. The loss is high for all splitters with open designs, and for all the shoveling methods with several steps involved. Especially the seed splitter, the animal feed splitter and the non-calibrated Boerner Divider stand out as spilling large amounts of material (especially rape seed is seen to bounce out).

In Table 1, the compound characteristics for the investigated methods are summarized. The different values are here weighed equally and it is left for the reader to apply differential weighing to fit his or her own customs or needs. If for instance operating time is of greater importance than cleaning, or if it is absolutely crucial to have the correct mass in the final sample, these parameters can be weighed on an individual basis.

The resulting sum-scores divide the methods/devices into three groups:

Very Good (sum = 5)
- Boerner Divider (Cal)
- RK 34 chutes of 10 mm short
- RK 34 chutes of 10 mm long
- 32-Divider
- Vario Divider 1:4 + 1:5

Acceptable (sum = 4, 3 or 2) but only under certain, problem-related circumstances:
- Boerner Divider (Non-cal.)
- Seed Splitter
- Rk 18 chutes of 16 mm
- RK 34 chutes of 10 mm normal
- Vario Divider 1:2 + 1:2 + 1:5
- RK 10 chutes of 30 mm

Poor (sum less than 2) not recommended under any circumstances
- Animal feed splitter
- RK 10 chutes of 20 mm
- Spoon method
- Alternate shoveling
- Fractional shoveling
- Grab sampling

The newly developed riffle splitters, the calibrated Boerner Divider, the 32-divider and the Vario Divider outperform all other methods—even though they all can be difficult to clean. The riffle splitters are in general rather slow to use, but this is a relative factor, since the slowest method overall—alternate shoveling—uses only approximately 200 s to reduce the mass by a factor of 20. The devices in the best group are all really good at finding the correct target concentration and mass of a final sample.

This is also the overall conclusion for the intermediate group. Most of these methods have a significant loss however and/or are also rather slow to use.

Table 1
Summary of the characteristics for the investigated methods

	Boerner Divider (Cal)	RK 34 long	RK 34 short	32-Dvider	Boerner Divider	Seed splitter	Vario 1:4 + 1:5	Animal feed splitter	RK 18 chutes og 16 mm	RK 34 Normal	RK 10 chutes of 20 mm	Vario 1:2 + 1:2 + 1:5	RK 10 chutes of 30 mm	Spoon method	Alternate shoveling	Fractional shoveling	Grab sampling
Composition	+	+	+	+	+	+	+	+	+	+	0	0	0	−	−	−	−
Mass	+	+	+	+	0	+	+	0	+	+	−	+	+	−	−	−	−
Loss	+	+	+	+	−	−	+	−	+	0	+	+	+	+	0	0	+
Cleaning	0	0	0	−	0	+	−	+	0	0	0	−	0	0	+	+	+
Initialization	+	+	+	+	+	+	−	+	0	0	+	+	+	0	+	0	+
User-dependency	+	+	+	+	+	+	+	+	0	+	0	+	0	−	−	−	−
Time	+	+	0	+	0	0	+	−	0	0	−	+	0	−	−	−	+
Score	5	5	5	5	2	4	5	1	4	4	0	4	3	−3	−2	−3	1

Table 2
Parameters used to estimate FSE

	M_L [g]	a_L	δ_A [g/cm^3]	c [g/cm^3]	β	f	g	d [cm]
Wheat	2000	0.899	0.75	0.088	1	0.1	0.65	0.35
Rape seed	2000	0.100	0.77	6.914	1	0.48	0.8	0.26
Glass	2000	0.001	2.6	2595.554	1	0.52	1	0.1

δ_A is the density.

The dividers/methods in the worst group include, amongst others, all the shoveling methods and many of the methods which have really substantial losses, show large user-dependency, are (too) slow or have severe difficulties to end up with the correct target concentration or mass. It is, however, important to consider carefully the purpose of the method/device, so as to make the correct choice. In this context users should pay special attention to the parameters of importance in the specific situation.

4.3. Estimation and comparison of the Fundamental Sampling Error

The Fundamental Sampling Error (FSE) is the error that remains when the sampling procedure is rid of incorrect errors and faults. This means that FSE is the minimum sampling error that can be obtained in practice and it is inherent only to the material heterogeneity. For this very reason it is, of cause, method-independent. FSE can be calculated from a series of measurements as the difference between the estimate of the lot grade, a_S, and the actual lot grade, a_L (known in the present experiments):

$$\mathrm{FSE} = \frac{a_S - a_L}{a_L}$$

FSE for a given material can also be estimated beforehand using the so-called "Pierre Gy formula" [1,2]:

$$s^2(\mathrm{FSE}) = cfg\beta d^3\left(\frac{1}{M_S} - \frac{1}{M_L}\right)$$

- c is the constitution parameter expressed in g/cm^3 that accounts for the densities as well as the proportions of the constituents.
- f is a "particle shape factor" (dimensionless) describing the deviation from the ideal shape of a cube. A square will have $f=1$, a sphere $f=0.52$ and an almost flat disc $f=0.1$.
- g is a "size distribution factor" (dimensionless) describing the span of particle sizes in the lot. Default values are estimated by Gy and Pitard [1,2].
- β is a "liberation factor" (dimensionless) describing the degree of liberation of the critical component from the matrix. Totally liberated particles means $\beta=1$ and totally incorporated particles means $\beta=0$.
- d is the "top particle size", defined as the square-mesh screen that retains 5% of the material (dimension of length expressed in cm)—this does not necessarily correspond to the physical particle diameter, as in the case of "cylindrical" particles such as wheat.

The parameters listed in Table 2 were used to calculate FSE for the given materials used here.

The ratios shown in Figs. 32–34 should optimally be around 1.0, which would imply that the methods or devices only have a sampling error in the range of the Fundamental Sampling Error (FSE), implying very low deviation from minimum practical sampling error. Pierre Gy's estimate is

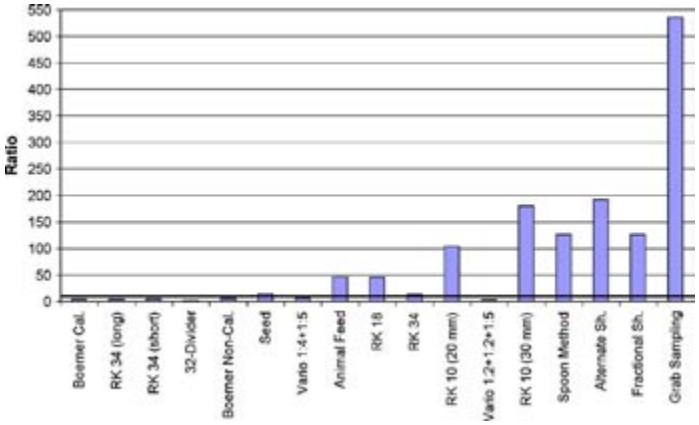


Fig. 32. Ratio between estimate of FSE and FSE from experimental procedure (for wheat). The horizontal line depicts the ratio of 10, indicating that all ratios lower than this is within an order of magnitude from the Pierre Gy formula estimate of FSE.

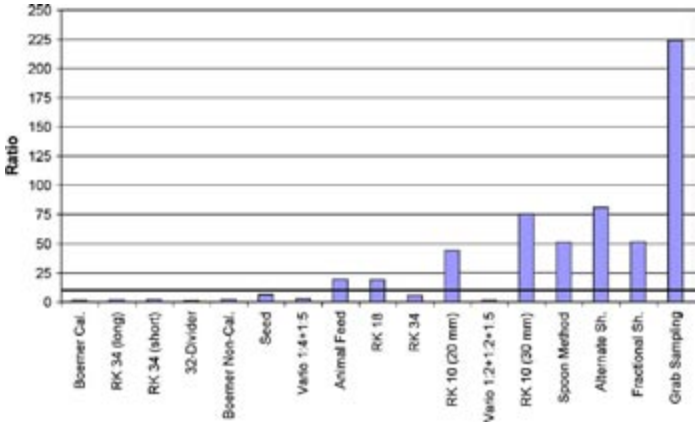


Fig. 33. Ratio between estimate of FSE and FSE from experimental procedure (for rape seed). The horizontal line depicts the ratio of 10, indicating that all ratios lower than this is within an order of magnitude from the Pierre Gy formula estimate of FSE.

meant to give the order of magnitude for the value of FSE; in this case the ratio should maximally be 10. This is marked as the flat line shown in Figs. 32–34. From the figures it can be observed that the estimates fit nicely with the experimental values for all acceptable methods, indicating that Pierre Gy's formula can be used for getting a rough estimate of FSE prior to any experimental procedure. At the same time, it further indicates the great overall performance of the best of the methods. It must, however, be stressed clearly that Pierre Gy's formula only yields an *estimate to an order of magnitude* of FSE, and must not be taken for an absolutely true value.

5. Conclusions

In order to achieve the best possible mass reduction it is crucial that the operator clearly analyses and defines the needs in a specific situation. In the overall characterization of the methods it has been decided to weigh all the characterizing parameters equally, even though this might be unreasonable for some applications in certain situations. We ask the readers to make their own modified conclusions from their specific needs; it will be very easy to consult Table 1 in this context. We have here emphasized the overall pooled characteristics of the methods investigated.

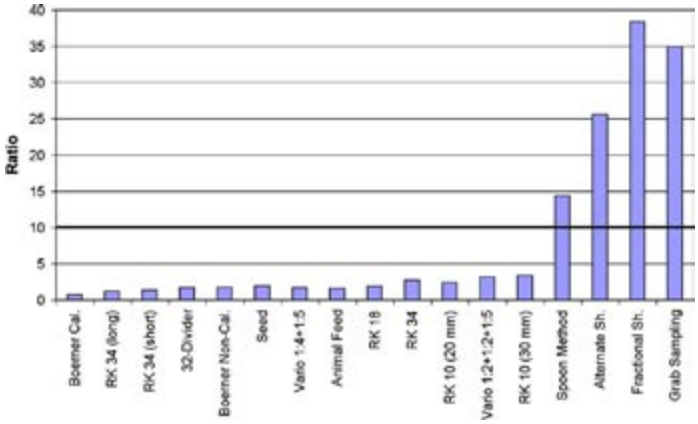


Fig. 34. Ratio between estimate of FSE and FSE from experimental procedure (for glass). The horizontal line depicts the ratio of 10, indicating that all ratios lower than this is within an order of magnitude from the Pierre Gy formula estimate of FSE.

Fig. 35. Grab sampling—the world's *worst* mass reduction/sampling method! If the lot material is heterogeneous and/or segregated (which is most often the case), grab sampling is the simplest and fastest way to get *heavily biased samples*. A miniature riffle splitter can easily be used instead—adding only seconds to the total preparation time, but several orders-of-magnitude to the representativeness. Grab sampling is to be totally avoided!

If all parameters can indeed be equally weighed, the following conclusions can be drawn.

The best overall methods for mass reduction are:

- Boener divider (cal.)
- RK 34 chutes (10 mm) short
- RK 34 chutes (10 mm) long
- Rotating 32-divider
- Vario Divider with splitting ratio 1:4 + 1:5

Of these methods riffle splitters are portable devices to be used both in the field or laboratory, while the latter two are heavy machines that can only be installed permanently for high speed reductions in permanent sampling stations or in laboratories. The Boerner Divider is heavy, though it also can be classified as portable if in a tight spot. All five methods perform excellently in both finding the correct target concentrations as well as having the nominal split-mass in the final samples. This compound criterion includes exactly what should rightly be characterized as the two most important parameters of a mass reduction device or method. Any differences concerning proper mass reduction between these five devices are minor and can be regarded as insignificant.

5.1. Total rejection of the world's most often used method—grab sampling

In general, all grab sampling and shoveling methods must be totally avoided; indeed grab sampling should never be used in practice—with the singular exception for thoroughly homogenized fine powders. It is a sad state of affairs that it is indeed a really fast, easy and cheap method since it "just happens" to be the absolute worst of all mass reduction methods. Fig. 35 tells its own story directly with the utmost clarity.

Acknowledgements

We would like to express our gratitude towards A/S Rationel Kornservice, Esbjerg, Denmark (Knud Klit, Axel Schou and Christian Husted) for their invaluable pieces of advice, time and help within the present work.

We would also like to thank the Danish Ministry of Agriculture's departments of seed testing and animal feed testing (Dot Vittrup Pedersen and Lone Bjørn) for help and general correspondence on the present work.

Peter Paasch Mortensen is thanked for his "magic" illustrative powder mix.

References

[1] P. Gy, Sampling for analytical purposes, Wiley, Chichester, England, 1998.
[2] F.F. Pitard, Pierre Gy's Sampling Theory and Sampling Practice, 2nd ed., CRC Press, Boca Raton, FL, 1993.
[3] P.L. Smith, A primer for sampling solids, liquids and gases-based on the seven sampling errors of Pierre Gy, ASA SIAM, USA, 2001.
[4] P.M. Gy, Coal Mining and Processing, (1981 September) 62–67.
[5] P.M. Gy, Process Control and Quality 1 (1990) 5–22.
[6] P.M. Gy, Process Control and Quality 6 (1994) 97–102.
[7] P.M. Gy, Trends in Analytical Chemistry 14 (1995) 67–76.
[8] P.M. Gy, Analusis 23 (1995) 497–500.
[9] P.M. Gy, HiT Skrift 1 (2000) 255–265.
[10] P.M. Gy, LC-GC 11 (1994) 808–817.
[11] P. Minkkinen, Analytica Chimica Acta 196 (1987) 237–245.
[12] G.J. Lyman, International Journal of Mineral Processing 55 (1998) 95–112.
[13] P. Minkkinen, Chemometrics and Intelligent Laboratory Systems 29 (1995) 263–270.
[14] P.M. Gy, Analytica Chimica Acta 190 (1986) 13–23.
[15] T. Lwin, R.C.A. Flann, G.M. Short, W. Guthrie, International Journal of Mineral Processing 54 (1998) 59–80.
[16] J. Ronalds, Australian Development Assistance Course on the Preservation of Stored Cereals, vol. I, 1981, Proceedings of the conference, Australia, CSIRO Division of Entomology, Canberra, 1983, pp. 333–346.
[17] R.W. Gerlach, D.E. Dobb, G.A. Raab, J.M. Nocerino, Journal of Chemometrics 16 (2002) 321–328.
[18] M. Rhodes, Introduction to Particle Technology, Wiley, Chichester, England, 1998.
[19] Personal communication, with-, and product descriptions from A/S Rationel Kornservice, Esbjerg, Denmark, 2003.
[20] P.M. Gy, International Journal of Mineral Processing 8 (1981) 201–238.