



Adequacy and verifiability of pharmaceutical mixtures and dose units by variographic analysis (Theory of Sampling) – A call for a regulatory paradigm shift



Kim H. Esbensen^{a,1}, Andrés D. Román-Ospino^b, Adriluz Sanchez^b, Rodolfo J. Romañach^{b,*}

^a KHE Consulting, Copenhagen, Denmark

^b P.O. Box 9000, Department of Chemistry, University of Puerto Rico – Mayagüez, Campus, Mayagüez, Puerto Rico

ARTICLE INFO

Article history:

Received 16 August 2015

Received in revised form 8 December 2015

Accepted 12 December 2015

Available online 18 December 2015

Keywords:

Theory of Sampling

TOS

Blend uniformity analysis (BUA)

Variographic analysis

Measurement uncertainty

Sampling paradigm shift

Regulatory science

Quality control

ABSTRACT

In spite of intense efforts in the last 20 years, the current state of affairs regarding evaluation of adequacy of pharmaceutical mixing is at an impressive standstill, characterized by two draft guidances, one withdrawn, and the other never approved. We here analyze the regulatory, scientific and technological situation and suggest a radical, but logical approach calling for a paradigm shift regarding sampling of pharmaceutical blends. In synergy with QbD/PAT efforts, blend uniformity testing should only be performed with properly designed sampling that can guarantee representativity—in contrast to the current deficient thief sampling. This is necessary for suitable in-process specifications and dosage units meeting desired specifications. The present exposé shows how process sampling based on the Theory of Sampling (TOS) constitutes a new asset for regulatory compliance, providing procedures that suppress hitherto adverse sampling errors. We identify that the optimal sampling location is after emptying the blender, guaranteeing complete characterisation of the residual heterogeneity. TOS includes variographic analysis that decomposes the effective total sampling and analysis error (TSE + TAE) from the variability of the manufacturing process itself. This approach provides reliable in-process characterization allowing independent approval or rejection by the Quality Control unit. The science-based sampling principles presented here will facilitate full control of blending processes, including whether post-blending segregation influences the material stream that reaches the tableting feed-frame.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Despite intense efforts over more than 20 years, the current state of affairs regarding adequacy and verifiability of pharmaceutical mixing is at an impressive standstill. The situation is characterized by two draft guidances, one of which one has been withdrawn, and the second never approved (U.S. Department of Health and Human Services, 2002, 2015). This can only be characterized as a colossal negative outcome after trying so long to formulate a guidance, or to build a consensus for a problem that has attracted intense attention (Berman and Planchard, 1995; Boehm, 2001; Chang et al., 1996; El-Hagrasy et al., 2006; Garcia et al., 1998, 2001; Hwang et al., 1998; Kræmer et al., 1999; Muzzio et al., 1997, 1999, 2003; Prescott and Garcia, 2001; Romañach,

2015; Wu et al., 2000). The debate has encompassed working groups composed of regulators, statisticians, industrial scientists and academicians (Bergum et al., 2014, 2015; Boehm et al., 2003, 2004; Garcia et al., 2015).

The Theory of Sampling (TOS) approach described in this document is generic, and not tied in with a priori fixed sampling plans (pre-defined location, number and size of samples), and may even be applied when different operating principles are used in mixing. This approach has been used extensively in other process industrial contexts for more than 60 years (e.g. in the mining, metallurgical, petroleum, agricultural and environmental industry sectors). A number of textbooks and a plethora of articles have been published describing TOS in many highly regarded peer reviewed journals and a substantial experience base has accumulated. Most recently TOS is the sole subject of the recently approved standard DS 3077 (Danish-Standards-Foundation, 2013), which is the world's only standard on the necessary and sufficient principles of representative sampling. However, the use of TOS has been very limited in the pharmaceutical industry (Colón et al., 2014; Green et al., 2005; Sánchez-Paternina et al., 2015). Here we

* Corresponding author. Fax: +1 7872653849.

E-mail addresses: khe.consult@gmail.com (K.H. Esbensen), rodolfoj.romanach@upr.edu (R.J. Romañach).

¹ www.kheconsult.com

show how TOS provides a new rationale for sampling in the pharmaceutical sector and offers new directions for R&D. This rationale can be illustrated by focusing on the key underlying phenomenon behind every problem encountered within the full pathway from material to analysis, *heterogeneity*.

TOS treats the concept of ‘homogeneity’ slightly differently, but much more rationally than the traditional approach within pharma where homogeneity is defined by a relative standard deviation below 5% as expressed by a specific set of extracted samples. There is a subtle but crucial distinction in (trying to) define homogeneity (or heterogeneity) as a statistical population of analytical results only, which is futile (as shown below), and TOS’ concepts of heterogeneity, which is much more valid characterization of real-world materials DS 3077 (Danish-Standards-Foundation, 2013). Also, TOS defines homogeneity as the theoretically and practically limiting case of zero heterogeneity: In a homogeneous sample every constituent unit (the smallest physically separable entity (“unit”) of the lot, which is not affected by the sampling process) would have to be identical. While most of the constituents of samples are intact molecules, grains, particles . . . they can also be fragments hereof, if/when produced by the sampling process itself. By ingeniously calling both the original, unaffected units, as well as the occasional fragments hereof generically: *fragments*, TOS is able to deal conceptually with all effective constituent units of a lot. If such a thing as a homogeneous material did exist, sampling would not be needed (as all sampling errors would be zero). However all naturally occurring and industrial materials are in fact significantly heterogeneous, it is only a matter of at what scale the heterogeneity manifests itself usually at more/all scales simultaneously.

TOS defines two conceptual components: constitutional heterogeneity (CH) and distributional heterogeneity (DH), to understand, describe, manage and counteract . . . heterogeneity. CH describes the heterogeneity dependent on the physical or chemical differences *between* individual grains or particles (collectively termed ‘fragments’ as per the above). CH can only be reduced by altering the physical state of the material, usually by crushing, i.e. by a preferential reduction of the top particle size(s) (termed comminution) which brings the lot material into a principally new state. DH is defined on the larger scale corresponding to groups-of-fragments (the *increments* in practical sampling, see Table 2 below). DH describes the complementary aspect of heterogeneity, which is dependent upon the spatial distribution of the individual groups-of-fragments in the lot (stratification or segregation and the formation of local groups, of varying coherence, of groups with a significant higher/lower concentration of the critical element). Because of the change in volume support in the definition of CH (fragment scale) vs. DH (increment scale), DH is numerically a fraction of CH; full conceptual definitions and relationships can be found in the fundamental TOS literature, e.g. DS 3077 (Danish-Standards-Foundation, 2013).

DH can be reduced in practice with less effort than CH, either by mixing before sampling, and/or by using a *composite sampling* approach, which – crucially – must cover the entire lot. After mixing, there will always be a minimum residual heterogeneity, and if needed this can only be further counteracted by composite sampling. Dependent on the effectiveness of the sampling process DH can in special cases become close to practically negligible, although never nil; it is absolutely necessary however to estimate the degree of (residual) heterogeneity when pharma is but hoping for “homogeneity”.

These few introductory concepts, together with a more in-depth introduction to TOS below, will elucidate the full nature and sources of heterogeneity observed in pharmaceutical mixtures. If such heterogeneity effects are not properly counteracted and managed, the sampling process will generate unnecessarily

inflated sampling errors, which constitute the often overlooked major contribution of the total between-sample standard deviations.

The presently proposed approach may be applied to tumble and/or convective mixers in both batch and continuous mixing operation modes. This approach bypasses all the problems that have been shown regarding the lack of adequacy and reliability of all tests based on physical sampling from blenders etc. The proposed approach is focused on how samples should be obtained to be representative of the end-state of mixtures. We show here that it is impossible to rely on the currently mandated in-process sampling strategy and come up with necessary-and-sufficient new solutions: It is in fact not necessary, and not desirable, to sample from *within* a blender—it is better to do the required in-process and final product sampling for testing further downstream. Indeed this postponed sampling will contribute to the desired end-state QC. For example, the case below with a batch process reveals a process in need of significant improvement.

The focus of the 20-year debate has ranged from equipment (advantages/disadvantages of various types of *sampling thieves a.o.*) to variability criteria (adequate statistical specifications for drug distribution in blends—NB based only on considerations for analytical values, not their full CH/DH context). This debate has not recognized that the levels of sampling errors involved actually disallow meaningful inspection of the results of the action of blenders. There are severe limitations regarding thief sampling because of material drag-down and alteration of the in-situ powder micro-structure as a thief is inserted, as shown powerfully in (Muzzio et al., 1997). In another case, 343 tablet cores were analyzed obtaining results between 97% and 108% of label claim, while 92 of 93 blend samples were found to be below label claim and showed a wider variation (Berman, 2001). A third study found that electrostatic charge caused locations deeper in the blender to be approximately 115–120% of label claim (Garcia et al., 2001). But even using carefully designed thief samplers (Muzzio et al., 2003; Susana et al., 2011) only a very limited number of pre-selected, fixed locations in a blender can be analyzed. With this approach the pervasive *assumption* is that the small set of extracted samples are *representative* of the entire lot—Unfortunately this assumption is but a pervasive *myth*, as shall be shown here. Most of the earlier approaches have been focused on determining the drug distribution *within* the blender and very little effort has been directed at the blend *after* outflow (Benedetti et al., 2007; Chang et al., 2004; Mateo-Ortiz et al., 2014; Popo et al., 2002).

Despite the current lack of progress in estimating the adequacy of mixing, there is total agreement on the need to establish a science- and risk-based approach to evaluate blend and content uniformity. We here offer a new framework to design proper sampling and to minimize the total sampling errors involved: this approach is also applicable to the current process analytical technology (PAT) drive in pharma. The new approach is linked to *variographic process monitoring* (Esbensen and Julius, 2009; Esbensen and Paasch-Mortensen, 2010; Esbensen and Romañach, 2015; Petersen and Esbensen, 2005). *Variographic analysis* allows to decompose the effective total sampling and analysis errors (TSE+TAE) from the variability associated with the manufacturing, or the mixing process itself, which is the true monitoring objective. Variographic analysis can also be used to estimate the sampling errors associated with specific process stages, e.g. procuring in-process or final product samples respectively (Esbensen et al., 2007).

1.1. Regulatory requirements – building a rationale for a paradigm shift

Section 211.110 of the Good Manufacturing Practices provides comprehensive requirements for sampling and testing of in-

Table 1

Synoptic overview of the Theory of Sampling (TOS). Six governing principles and four sampling unit operations (SUO) are necessary and sufficient for representative sampling, [Esbensen and Wagner \(2014a,b\)](#); DS 3077 ([Danish-Standards-Foundation, 2013](#)); see also Section 3 in text.

Representative sampling: Theory of Sampling (TOS)	
1. FSP: fundamental sampling principle	
2. SSI: sampling scale invariance	
3. PSC: sampling correctness (bias-free sampling)	
4. PSS: sampling simplicity (primary sampling + mass reduction)	
5. LDT: lot dimensionality transformation	
6. LHC: lot heterogeneity characterization (0-D, 1-D)	
7. SUO: composite sampling	
8. SUO: comminution	
9. SUO: mixing/blending	
10. SUO: representative mass reduction (sub-sampling)	

process materials and drug products. These include written procedures for sampling and in-process tests (211.110(a)), in-process specifications (211.110 (b)), and the acceptance or rejection of the material by the quality control unit (211.110(c and d)).

Section 211.110 (a) requires procedures both for in-process controls and tests for appropriate samples of in-process materials of each batch. The control procedures “shall be established to monitor the output and to validate the performance of those manufacturing processes that may be responsible for causing variability in the characteristics of in-process material and the drug product.” These control procedures include the issue: “Adequacy of mixing to assure uniformity and homogeneity”. The requirements are re-emphasized in the process validation guidance ([U.S. Department of Health and Human Services, 2011](#)) which states that manufacturers “should:

- understand the sources of variation,
- detect the presence and degree of variation,
- understand the impact of variation on the process and ultimately on product attributes,
- control the variation in a manner commensurate with the risk it represents to the process and product.”

The detailed understanding of measurement and process variation should be the source of valid in-process specifications

Table 2

TOS' definition of sampling errors.

Fundamental sampling error (FSE)	Error due to the constitutional heterogeneity of the material. FSE is constant for a given material in a given physical state. Only comminution can reduce FSE.
Global estimation error (GEE)	Also called the <i>total</i> “Measurement Uncertainty”. GEE is comprised of two contributions: The total analytical error (TAE) and the total sampling error (TSE): $GEE = TSE + TAE$.
Grouping and segregation error (GSE):	GSE is related both to the spatial heterogeneity as well as to the sampling process. This error exists because of two material factors, segregation and grouping (mixing and/or composite sampling reduces these error effects). The magnitude of GSE also depends on the sampling process, especially on the total number of increments deployed, which consequently also impact on the total sampling mass. A higher number of increments, Q , will always reduce the magnitude of GSE.
Increment delimitation error (IDE)	A correctly delineated process increment consists of a complete cross-section of the stream of matter (or a stationary elongated stockpile), as defined by two parallel boundaries. This ensures that no geometrical part of the lot is represented in higher or lower proportions <i>across</i> the stream than any other in successive increments along the principal 1-D direction.
Increment extraction error (IEE):	All material located <i>within</i> the correctly delineated boundaries must also be <i>extracted</i> in practice. This means that all fragments having their center of gravity inside the delimited increment must be included in the extraction process. This principle is known as the center-of-gravity rule (or the rebounding rule for fragment in a falling stream intersected by the moving wall of the sampling tool). IEE quantifies the errors associated with physical increment extraction from the delineated sampling volume. Any deviation from the center-of-gravity rule will inevitably lead to an extraction error, which can be substantial with some types of particulate materials in which fragments display significant sizes compared to the opening of the sampling tool, or in slurries. Obviously errors such as IDE and IEE, exert more influence on the TSE the higher the heterogeneity of the material in question.
Increment preparation error (IPE):	The third ISE error arises if/when the sample is in any way <i>altered</i> after extraction, for instance by absorbing moisture (or by evaporation), by spillage, by cross-contamination, or by anything else that inadvertently happens to a sample after selection and materialization that affects the analytical result; IPE also results from deliberate sample alteration e.g. from sabotage or fraud. The term increment preparation error (IPE) also includes unintended effects induced during processes commonly known in the analytical laboratory as ‘sample handling and preparation’.
Process integration error	(PIE, or CE (continuation error)), is comprised of three parts, of which only two are new compared to the 0, 2 and 3-D cases. The reader is referred to the TOS literature for a complete introduction to the process errors, e.g. to the literature cited in DS 3077 (Danish-Standards-Foundation, 2013). Pitard (1993) gives a complete overview of process sampling, as does Gy (1998) .

consistent with the drug product final specifications to comply with Section 211.110 (b), and serve as the basis for approval or rejection by the quality control unit (Section 211.110 (c)).

The observable variability of in-process material and drug product characteristics is affected by many *sources*. The variability is not only related to the manufacturing process, and to errors stemming from analysis, but is also attributable to effects created by the sampling process itself producing *sampling errors* when interacting with a heterogeneous material. The bias introduced by poorly designed or executed sampling cannot be corrected for, e.g. as shown in DS 3077 ([Danish-Standards-Foundation, 2013](#)) and references herein for which reason proper sampling processes have to be designed so as to eliminate the bias (see further below). Understanding the specific sampling process in use and the errors incurred is essential. 60+ years of experience with TOS as applied across a wide range of sectors in science, technology and industry has led to the conclusion that sampling errors are considerably larger than the variability introduced by analysis, indeed they are very often dominating with factors of 10–20–50 times the total analytical error, *ibid*.

A professional understanding and characterization of such sampling error contributions is essential for development of *valid* in-process specifications as required by Section 211.110 (b).

The main thrust of this paper: ‘Sources of variation’ cannot be comprehensibly understood if *thief sampling* is used because it negates comprehensive and valid in-process observability. The sampling error effects embedded in the analytical results characterizing in-process and final product material variability have to be eliminated. Thief sampling is an example of *grab sampling*, a type of sampling that is incurring the greatest adverse effect of all possible sampling procedures in the form of the *inconstant sampling bias*. The subtraction of bias, which is a direct extension of experiences in the analytical laboratory, unfortunately turns out to be a mirage when matters are taken over to the sampling bias realm, see ([Esbensen and Wagner, 2014a,b](#); [Minkkinen and Esbensen, 2009](#)). The sampling error will be *different* every time that a grab sample is obtained, making it impossible to subtract a bias effect from the result obtained. Only the Theory of Sampling formulates a comprehensive understanding of the reasons behind the sampling bias, and only TOS

formulates the principles with which to *eliminate* a sampling bias from any sampling process.

1.2. Theory of Sampling (TOS) and pharmaceutical manufacturing

In TOS the definition of a 'correct' sampling process or sampling procedure requires exclusion of the elements that contribute to sampling bias and that all potential increments of the lot in practice have an equal, non-zero probability of ending up in the sample, while elements foreign to the lot must have a zero probability of being selected. Such sampling conditions are critical success factors for obtaining sampling accuracy. TOS uses the term 'sample' for a correctly extracted amount of lot material, obtained by an *accurate* sampling process. The term 'specimen' is used for a non-correct selection process; a specimen is structurally *biased*.

The Theory of Sampling can be formulated as a set of six governing principles (GP) and four sampling unit operations (SUO), as laid out in Section 3 below and shown in Table 1. As an example, the logical demand that all potential increments from any lot must have an equal probability of being sampled is called the fundamental sampling principle (FSP).

The FSP is in complete contradiction to current efforts in the pharmaceutical industry where sampling plans invariably involve a few pre-selected sections of the blender only. But heterogeneity is a characteristic of any material or lot over the entire volume.

To the credit of foresighted researches within the pharmaceutical industry, focus on 'homogeneity within blenders' has already identified a number of adverse effects associated with *sampling thieves* though (Berman, 2001; El-Hagrasy et al., 2006; Muzzio et al., 1997; Prescott and Garcia, 2001), but these criticisms have unfortunately not yet provided a practical and universal approach for pharmaceutical manufacturing control.

TOS defines 'the lot' (the sampling target), as the entity of the complete volume/mass of the original material being subject to sampling, e.g. a stockpile, a process stream, a reactor, barrel, lorry load, train wagon or a plot in the field, or in the present context: a mixing vessel. Lot dimensionality is defined in the TOS literature; distinction between 1-D and 2-D (3-D lots) follows an intuitive

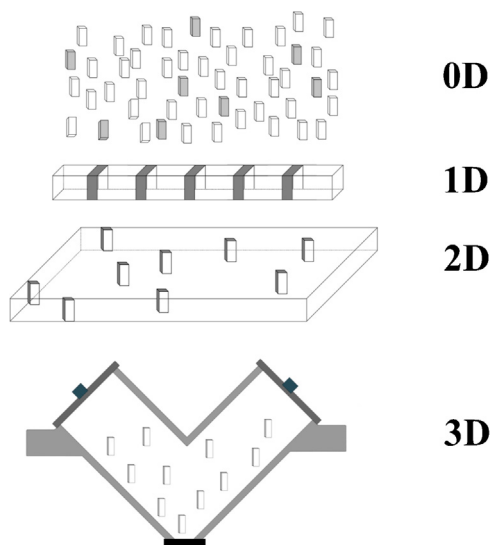


Fig. 1. Schematic overview of TOS' principal lot dimensionality definitions. Grey shaded increments cover the full lot volume in all cases (as best as possible with a given number of increments, Q). Specification of Q results from a replication experiment or via variographic analysis, DS 3077 (2013). It is appreciated how the 1-D lot configuration allows any sampling intensity from low to, in principle, complete inspection, cfr. text.

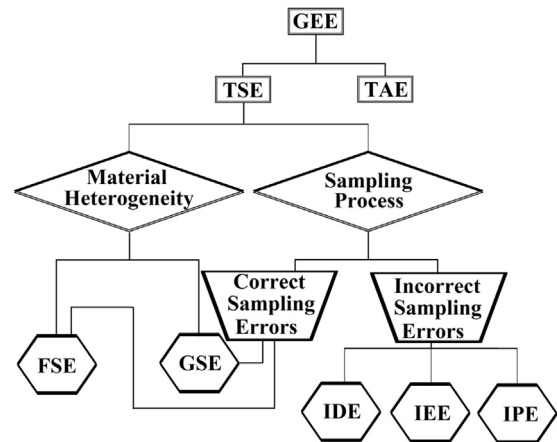


Fig. 2. Interrelationships between the basic five sampling errors originating partly because of the heterogeneity of the lot material (FSE, GSE) and due to incorrect sampling procedures (IDE, IEE, IPE). This setup applies to sampling from lots of all dimensionalities, cfr. Fig. 1.

geometric basis as shown in Fig. 1. Thus, a tumble blender would be considered a 3-D lot, but the outflow of this blender as it moves through a chute to a compressing machine would be considered a dynamic 1-D lot. Tablets from a tablet press can be considered as a linearized set samples also making up a 1-D lot. Linear/linearized 1-D lots offer a superior advantage over any 3-D lot configuration, and potential complete inspection.

Fig. 2 presents a synoptic diagram of all possible sampling errors that can be incurred when sampling stationary lots (3-D, 2-D, non-moving 1-D lots), further discussed in Table 2.

In TOS, composite sampling reigns supreme when the goal is to estimate a reliable (representative) average concentration for the entire lot. If/when the objective is to survey, map or monitor the heterogeneity of the lot, the individual samples must be representative (see below for the appropriate demands). Even if/when a lot is sampled correctly, the lot heterogeneity alone is responsible for the magnitude of the Fundamental Sampling Error (FSE), hence its name. FSE can be considered the practical minimum sampling error when all other sampling error components have been eliminated (which is not necessarily an easy task—See further below). This is the error due to the constitutional heterogeneity of the material, which is constant for a given material in a given physical state (only comminution/crushing will reduce FSE). To deal effectively with GSE, TOS stipulates composite sampling based on full lot coverage with a given number, Q , of increments. The *increment* is thus the fundamental sampling unit, specifically destined to be combined with other increments to provide the final, composite sample. Increments are groups of a number of spatially correlated or spatially coherent fragments (grains), which act as *independent* units during sampling operations. The group size depends on the sampling process, i.e. the sampling tool (volume) as well as how this is operated. Increments can either be a single or several groups of fragments depending on the size of a composite sample. Increment would also be the term specifying an across-stream cut from a 1-D moving product stream, or from an elongated stationary 1-D lot as shown in Fig. 1. Process sampling can be understood as a series of juxtaposed ordinary increments, treated as samples, arranged linearly *along* a dominating transportation dimension. Here the ordered auto-correlation between units comes into play, which determines the extraordinary effectiveness of sampling error reduction/elimination that can be achieved in the 1-D situation. TOS' Governing Principle no 5 stems from this understanding: Lot dimensionality

transformation. Plainly understood, it is (far) better to plan, design, implement and perform sampling on a 1-D transportation manifestation of an otherwise, much more difficult-to-sample (perhaps even impossible) 3-D configuration of the same lot. This approach follows the so-called ‘golden rules of powder sampling’ described by Allen: 1. A powder should always be sampled when in motion; 2. The whole of the stream of powder should be taken for many short increments of time in preference to part of the stream being taken for the whole of the time (Allen, 2003).

The Total Sampling Error springs from both the material properties (heterogeneity) as well as the sampling process itself—in the way a sampling process *interacts* with a heterogeneous lot. This compound nature emphasizes the futility of possible improvement of global estimation results by optimizing the analytical process only, which completely disregards the dominating uncertainty contributions stemming from various sampling errors.

There are two types of sampling errors encountered in all sampling operations, the so-called ‘correct sampling errors’ (CSE), i.e. the fundamental sampling error (FSE) and the grouping and segregation error (GSE) and the incorrect sampling errors (all due to faulty—or badly executed sampling processes); these latter are IDE, IEE and IPE, Fig. 2. The former are called “correct sampling errors” because they exist even after all “incorrect sampling errors” have been eliminated; incorrect errors (in fact the errors producing the sampling bias) are defined in Table 2.

Independently of the nature of the material that makes up the lot, TOS defines the seven basic sampling errors as described in Table 2. A set of five basic errors are related to sampling of 0-, 2- and 3-D lots, aka *stationary* lots. The additional two specific process errors only arise for dynamic 1-D lot sampling, aka process sampling. A full description of process errors is found in standard 3077 (2013) (Danish-Standards-Foundation, 2013).

All the above errors combine to form the total sampling error, TSE. At all sampling stages, these combine to the total sampling error (TSE), which can be written out in full (0-D and 1-D): $TSE = (FSE + GSE) + (IDE + IEE + IPE) + (\text{process errors, if appropriate})^1$.

At all sampling steps, sampling correctness should be controlled by eliminating the incorrect sampling errors, to ensure *unbiased* samples and a reduction in the sampling variance. Once the sampling process is unbiased, the precision (TOS favors the term ‘reproducibility’) can be controlled as in any other statistical setting, by a higher number of increments making up a composite sample, or by a higher number of samples, or by a higher sampling frequency as the specific case may be. The be-all of sampling concerns accuracy (unbiasedness): all sampling processes must be correct—and after this demand has been met, the variability of the now unbiased sampling process is subsequently easy to bring under further control. TOS defines ‘representative’ as the quality attribute of a sampling process which is *both* accurate and sufficiently precise (mathematical definition given in the literature) e.g. DS 3077 (Danish-Standards-Foundation, 2013).

1.3. Current state of disarray in pharma

On August 7, 2013 the FDA withdrew the Draft Guidance originally published in October 2003 proposed by the Product

¹ For process sampling there can be envisaged a problem if the chosen sampling frequency *happens* to be equal to, or a multiple of, an underlying process, or product, periodicity. Sampling essentially at the same phase stage in a cyclical measurement series, would then result in an apparently much more reduced variability (i.e. a much lower sill). TOS is very much aware of this possibility, which will always be checked for; if this case is identified, stratified random sampling is the standard cure, easily implemented, see the basic process TOS literature.

Quality Research Institute (PQRI) in December 2002 (Boehm et al., 2004). This draft guidance was never approved but was followed in many companies that will now need to justify their sampling practices. A new working group was formed with personnel from FDA, academia, and industry (Bergum et al., 2014), sponsored by the International Society for Pharmaceutical Engineering (ISPE). Modifications to the revised guidance have been proposed (Bergum et al., 2015; Garcia et al., 2015). There is a critical need to resolve this dilemma, as it could delay the validation and approval of new products needed by patients.

The latest overview (Bergum et al., 2014), places the current situation in a clear focus: “A question that must be addressed with any sampling plan and acceptance criteria is whether any or all specific problems (blend- non-uniformities, segregation, agglomeration) must be targeted with specific customized sampling plans, or whether the approach can allow these potential problems to be uncovered by other means”. We here present a universal means of checking the efficiency of the last mixing operation (the generic ‘blender unit operation’), with a surprising twist: checking is futile if forced to take place based on sampling from *within* the blender. Obviously we owe the reader another solution.

One of the reasons for withdrawing the draft guidance appears to be FDA’s interest (Bergum et al., 2014; Garcia et al., 2015) in having three replicates from each location of the blender and to subject the data for variance component analysis (ANOVA). This approach will certainly lead to problems due to the limitations of thieves. The samples analyzed will not be ‘true replicates’ because they originate, at best, as but back-to-back samples, taken from a heterogeneous material; we analyze this issue in detail in section 4. In contrast, the proposed TOS approach provides a thorough evaluation of the complete variance of a pharmaceutical blend from within its entire 3-D volume as *mapped* by a 1-D outflow stream. This approach contributes to the goals of the process validation guidance because the variographic method proposed here elucidates all the sources of variation in pharmaceutical processes (U.S. Department of Health and Human Services, 2011).

FDA’s and industry’s ‘vision’ has been to find the area(s) of *presumed* highest blender heterogeneity by introducing a sample thief in locations that *could* be suspected of being a ‘dead spot’ – or *could* constitute ‘places of incomplete mixing’ (Timmermans, 2001). But any localized sampling approach from a 3-D heterogeneous medium will incur the maximum adverse impact from the sum of the fundamental sampling error and the grouping and segregation error (FSE+GSE), as it is in essence grab sampling, which is analyzed and found insufficient in Section 3. In addition thief sampling is fraught with highly significant ISE.

2. Heterogeneity (the history of heterogeneity in pharma)

Blending of powders may be considered at both macro and micro-mixing scales (Osorio et al., 2014). The macro level may be used to describe drug content between tablets as in the content uniformity test currently performed. The micro-mixing scale describes how the excipients and the drug are arranged within a formulation. The formulation may include 100 mg well distributed throughout a 400 mg unit dose, or a series of large drug aggregates. The differences in drug distribution within blends have been investigated in various spectroscopic imaging experiments (Ma and Anderson, 2008; Osorio et al., 2014; Susarla et al., 2013) that have shown a scale-hierarchy of heterogeneity of blends at both macro and micro scale. The scales of scrutiny of interest for pharma are, by necessity of the order of magnitude, mg–g–kg, spanning 6+ orders of magnitudes, yet the Theory of Sampling offers a *scale-invariant* approach that will simplify matters greatly. TOS considers that all types of materials are

heterogeneous only at two different *operative scales*, which gives rise to the two features: compositional heterogeneity (CH) and distributional heterogeneity (DH).

TOS and the theory of mixing share a common focus on determination of variance in mixtures (Danckwerts, 1952; Ghaderi, 2003; Gy, 1981). TOS provides thoroughly tested methods that permit de-convolution of the sampling and analytical error variances related to the composition of a powder mixture. TOS facilitates evaluation analogous to the fundamental principles of Danckwerts, who proposed that the “goodness of mixing” be evaluated through the use of the correlation coefficient calculated over a number of locations in a linear pattern (Danckwerts, 1952). A low auto-correlation was considered indicative of ‘randomness’ in the local mixture, i.e. absence of ‘regularities’, for example vestiges of periodicity inherited from screw feeders to blenders. Higher correlation coefficients, from sample points closer together, could be considered to derive from “more coherent localised clumps”. Danckwerts devised the spatial correlogram as a plot of the correlation coefficient versus the distance between sample locations, and discussed how the geometric shape of ‘clumps’ would affect the form of such correlograms. Subsequent developments in mixing theory indicated that an ideal blender with free flowing non-segregating powders would reduce the variance of the mixture to that of a “random mixture” (Pernenkil and Cooney, 2006), regardless of whether batch or continuous mixing is performed. N.B. the operative prerequisite here is ‘non-segregation’ of powders. This mixing theory, the result of ~65 years of evolution, thus provided a desired conceptual end-state for a mixing process, the *random mixture*.

There is a direct line of evolution from the theory of mixing to TOS’ variographic analysis introduced here, but there is also a flaw in the above historical reasoning because there hardly exists pharmaceutical powders that truly are non-segregating, indeed even the concept of a random mixture is at odds with physical reality. Only the Theory of Sampling (TOS) has reached the fully realistic understanding that all mixtures will have as their end-destination from mixing the state of a ‘residual heterogeneity minimum’, which is material-dependent. This is not equivalent to the assumed “perfect random mixture” as e.g. argued by (Paoletti and Esbensen, 2015). Mixing beyond the minimum heterogeneity state will only lead to a steady-state mixing/segregation heterogeneity configuration centered on this minimum residual heterogeneity. TOS also demonstrates that all sample extraction from within the mixer potentially will be afflicted by up to five sampling errors (CSE+ISE), making up a combined minimum error level which is equivalent to a principal threshold below which any further mixing efficiency simply cannot be detected. This is precisely the result of variographic decomposition of the same process data as are obtained for process monitoring and control purposes.

TOS’ fundamental sampling principle (FSP) is systematically broken in all fixed sampling location plans, e.g. six fixed locations (Berman and Planchard, 1995), or ten fixed locations in the conventional V-blender geometry (Hwang et al., 1998). The fixed location plans will certainly not always select representative “hidden zones” that could be associated with variations in drug concentration in a *comparable manner* for all the many types of compositionally different mixtures and/or the range of different dimensions in current industrial and experimental blenders (geometry vs. mixing dead-zone scale-up issues). Given the differences in the physical properties of excipients and active pharmaceutical ingredients common in industry, it is highly unrealistic that such fixed locations will always constitute the worst case scenario with respect to patient protection against

non-uniformity, which was in fact the very idea behind developing these rigid, fixed sampling *schemes*.

3. Theory of Sampling (TOS) – a practical sampling vade mecum

TOS has analyzed sampling processes for all types of significantly heterogeneous lot material, backed up by a plethora of practical experience, and has reached the universal conclusion that composing a number of *increments*, Q , is the only- reliable, reproducible fashion to produce a *representative sample*, see DS 3077 (Danish-Standards-Foundation, 2013) and the foundational literature referred to here. *Composite sampling* is also intuitively obvious from a simple geometrical point of view: a spatially heterogeneous lot (the following argument is scale-invariant) cannot under any circumstances be represented—by one single randomly chosen ‘specimen’ (archetype grab sampling); there must be used as many increments as possible, or as needed (for example as determined by a suitable pilot lot heterogeneity characterizing experiment DS 3077), *ibid*, over the entire lot geometry/volume. There can be no exceptions from this geometrical demand, which is codified as the fundamental sampling principle (FSP). The extraction of only one increment, *erroneously* considered a sample, is a tradition (it is indeed a very convenient approach) that has developed in many industry sectors. However, grab sampling is the most unreliable approach, totally unable to provide the necessary heterogeneity compensation (Esbensen and Wagner, 2014b; Minkkinen and Esbensen, 2009). This conclusion is universal when the objective is a representative characterization of a significantly heterogeneous lot, irrespective of its size (mass), geometry and of the conditions under which it can be sampled – the FSP must be upheld. But if the purpose of sampling specifically is to *map* the variability of sample-concentrations over the geometric support (volume) of the lot, the demands for representativity simply applies to each individual sample (see further below).

For practical sampling to be reliable, an *optimal* increment number, Q , must be scaled commensurate with the empirical heterogeneity encountered, not with the total lot mass (a prevalent, but unfortunately another myth). First after this requirement can other factors be admitted in the cost/benefit analysis, factors such as logistics, economy, ergonomics. These interrelated issues were discussed in full detail in the industrial practice in the benchmark KeLDA study (Esbensen et al., 2012a,b; Minkkinen et al., 2012). Q is usually determined by a pilot experiment, called a Replication Experiment, DS 3077, or by a variographic analysis, *ibid*. Sometimes this parameter is known from previous knowledge, obviously because of intimate process/product knowledge accumulated over a substantial process time.

‘Sampling correctness’: A correct sampling process is accurate, if *on average* the effective sampling bias is zero. When correct, any sampling process has now earned the term: *unbiased*. If the incorrect sampling errors (ISE) are not eliminated however, a sampling bias different from zero will arise. More importantly, a sampling bias is manifested *differently* every time an incorrect sampling process is performed. When the sampling process interacts with a heterogeneous material at a different location, the sampling bias will appear with a different magnitude because of the irregular spatial heterogeneity, DH. Thus an incorrect sampling process interacting with a heterogeneous material results in an ‘inconstant’ sampling bias that changes in magnitude every time it is realized, or estimated. For this reason it cannot be subject to a conventional statistical “bias correction”. TOS draws the logical conclusion that the sampling bias must be eliminated, which in practice is effectuated by designing the sampling procedure (and equipment) so as not to produce ISE.

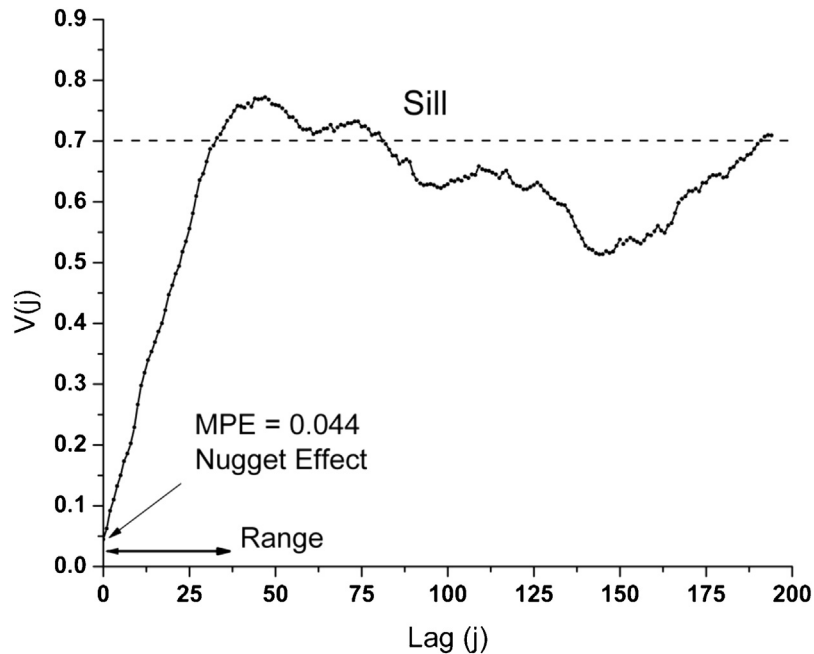


Fig. 3. A process is characterized completely by three parameters: range, sill and nugget effect, which are all estimated from an experimental variogram. While the sill is estimated with respect to the variance level for all lags above the range, the nugget effect is estimated by back-extrapolation of the general trend at the lowest lags below the range. The lag, although measured in the same units as distance, does not represent the linear extension dimension (geometry), but the *inter-increment* distance between pairs of measurements. The variogram is a domain transformation from the distance (time/profile length) to the pair-wise inter-distance domain. The variogram shown has a particularly small nugget effect, cmp. Figs. 4 and 7 below.

3.1. The variogram

Variographic analysis is conducted on a 1-D lot to estimate the magnitude of the total sampling and analytical errors. The *variographic experiment*, in which a number of increments is extracted from a 1-D lot, usually requires a minimum between 60 and 100 increments²; in pharma contexts it will often be possible to use a considerable higher number thanks to the PAT analysis. Each increment is at first treated as an individual sample, with the clear understanding that one of the outcomes of the variographic analysis often is the realization that a consecutive number of these increments, Q , can advantageously be combined to form a composite sample, specifically the number which will reduce the total sampling and analysis error to below an a priori threshold level. This is typically in the case where one is interested in extracting a process sample with a specified, below threshold (TSE+TAE). Composite sampling with precisely this Q_{optimal} number of increments will then provide a fit-for-purpose sampling plan (Danish-Standards-Foundation, 2013; Esbensen et al., 2007).

Another prime objective for variographic analysis is to characterize the 1-D heterogeneity of the stream of matter under investigation, as shown in Fig. 3 for a pharmaceutical mixture. The variogram is constructed by calculating according to the *variogram function* $V(j)$:

$$V(j) = \frac{1}{2(Q_{\text{total}} - j)} \sum_{q=1}^{Q_{\text{total}}-j} (h_{q+j} - h_q)^2 \quad (1)$$

² 60–100 increments represent an extensive practical experience with variographic analysis of very many types of processes from science, technology and industry. “Usually” 60–100 increments is enough.

$V(j)$ is a function of the distance between extracted increments, the *lag*. Q_{total} is the total number of analytical results; j is the *lag* and h is the heterogeneity contribution of the analyte measured in each increment. The *lag* is defined as the inter-distance between pairs of analytical results (the *lag* is the inverse of the sampling frequency)³. The *lag* is calculated as a dimensionless unit: $j = \theta / \theta_{\text{min}}$ where θ_{min} is the fixed minimum sampling interval along the 1-D direction, and runs in the interval $[1, Q_{\text{total}}/2]$. Low values of $V(j)$ result from a series of broadly *similar* drug concentrations and thus indicate high correlation between increments (tantamount to a low difference variability), as is illustrated in the case histories below.

The variogram in Fig. 3 shows the values of the variogram function $V(j)$ as a function of the lag. A lag of 1 corresponds to the unit distance (time or distance) between any two consecutive analytical results, irrespective of location in the ordered measurement series; *lag*=2 characterizes the doubled inter-distance between every other measurement etc. for higher lags. Low $V(j)$ values reflect a reduced heterogeneity in the blend, and is thus a reflection of good mixing. Low $V(j)$ values are usually observed for samples obtained at contiguous locations, i.e. obtained within close intervals of time. Higher $V(j)$ values are often observed when the samples are obtained at larger distances or sampling times, and the sill variance level represents the random variation of all process measurements when estimated without consideration of the ordering (auto-correlation) of the measurement series. Note that a variogram strictly speaking is but a set of $V(j)$ values, one for each lag (one can thus calculate $V(j)$ at maximum for $Q_{\text{total}}/2$ lags). A connecting line (Fig. 3), or a mathematical model approximating a smooth continuity functional relationship *between* each $V(j)$

³ The mass of each individual increment is either a function of the accumulated process knowledge (a situation almost universal in pharma), or it can be optimized during a pilot experimental campaign.

(geostatistics) has no physical meaning other than expressing an assumptive hierarchical variance continuity between neighboring scales (lags). This assumption is underpinning all of *geostatistics* which has developed variographic modelling independently of TOS⁴ (François-Bongarçon and Pitard, 2015; Matheron, 1966), but it is not needed within the realm of TOS, where the objective of the variogram is (i) to characterize and interpret the variance vs. lag relationships and (ii) to derive so-called auxiliary functions (also a function of the lag) with which to be able to estimate the Total Sampling Error (TSE) of a particular sampling plan (Danish-Standards-Foundation, 2013; Gy, 1998; Pitard, 1993). The variogram could be calculated while a process is occurring, as a sequential updatable variogram every time 100 new process data have been acquired. This is one of the primary features and great benefits of the on-line variographic measurement system quality assurance (see further below).

Variographic analysis allows estimation of the three key parameters (range, sill; nugget effect). The nugget effect can be interpreted as the minimum practical error (MPE). MPE is equivalent to the sum of all sampling and analytical errors associated with the measurement system employed. Thus, the *effective* TSE can be calculated since the analytical error (TAE) is always well-determined for pharmaceutical methods. The MPE is calculated by extrapolating $V(j)$ to intercept the Y-axis, effectively at “lag 0”. This combined TSE + TAE information is not known for pharma processes; the first variogram for a pharmaceutical mixing process was published as recently as 2014 (Colón et al., 2014).

A systematic introduction to variographic analysis can be found in many sources, with full background and detailed application contexts as well in a trio of didactic papers published in 2012 (Esbensen et al., 2012a,b; Minkkinen et al., 2012). For the benefit of the reader Appendix A gives an operative brief of variographic analysis sufficient for this paper.

The variogram can be used to improve sampling processes. The sill provides the maximum variance difference between the units sampled which reflects the maximum heterogeneity observed between samples if/when sampled at random. The sill is thus a reflection of the total process, or product, variation *including* the sampling and analytical errors. The systematic use of variographics would provide information that is not currently available to pharmaceutical QC labs and to pharmaceutical scientists designing pharmaceutical processes.

4. 60 years evolution of a theory of mixing

This section summarizes the evolution of a theory of mixing as background for a proposed new mixing index based on variographic analysis. The detailed exposé will show the relationship between a decades' long theoretical quest and the variogram approach provided by TOS.

The historical evolution outlined here is based on five contributions summarized below, which also contain further descriptions and references to fill out the complete picture. These papers have produced a successively more focused understanding of the enormous complexity of mixing processes, as the objective has been, per force, to cover all types of mixtures, all types of mixing and blending methods, e.g. equipment, sampling and monitoring approaches (whether characterized by direct physical sample outtake, or based on modern PAT based spectral acquisition). Unfortunately these key references have not yet been able to formulate a practical mixing efficiency index that could serve as basis for regulatory specifications: (Danckwerts, 1952; Ghaderi,

2003; Gy, 1998, 1981; Jia et al., 2009; Muzzio et al., 1997; Pernenkil and Cooney, 2006). A great many local achievements (theoretical, practical, empirical), have accumulated a sufficiently comprehensive understanding and conclusions, that have inspired and guided us to our current proposal.

Danckwerts provided a first theoretical approach to understand a generic mixing and express the efficiency of mixing (Danckwerts, 1952). A.o. he defined two phenomenological concepts, a segregation *scale* and segregation *intensity* respectively, used to describe variations in terms of a characteristic volume with a heterogeneity intensity deviating significantly from the average of an *ideal* random mixture. Thus:

“A knowledge of the scale and intensity of segregation in a mixture will provide a good deal of information about its texture, in a quantitative form which allows the different degrees of mixing to be compared. The two quantities are virtually independent and represent ‘goodness of mixing’ that cannot be defined by a single quantity. The determination of the scale and intensity, even by the simplest methods suggested, involves in effect making a considerable number of measurements and averaging them. This is inevitable, as the only useful information which can be given regarding the degree of mixing is of a statistical nature. It seems unlikely that any simpler program of measurements (than those suggested here) would yield quantities which would be of any value in a systematic investigation of mixing processes.” . . . “It is suggested that the statistical treatment of mixtures (. . .) may be useful in developing a general theory of mixing processes. For instances, if turbulence is used to bring about mixing it seems likely that power is most economically applied in producing turbulence with a scale comparable to the scale of segregation.” (Danckwerts, 1952).

Danckwerts' mathematical derivations are identical to the formalism of spatial autocorrelation, which is a parallel development to geostatistics (Ghaderi, 2003; Gy, 1981) – which in turn is a parallel to the variographic approach developed by TOS (Danish-Standards-Foundation, 2013; Gy, 1998). TOS includes the statistical treatment of mixtures pointed to by Danckwerts.

In 1997 a seminal paper was published, with very well documented criticism on the sampling practices in powder blending (Muzzio et al., 1997). While leveling this critique in the specific context of pharmaceutical operations, from a TOS point of view the concluding denouncement of thief sampling is much broader, indeed universal, and parallels the experience from TOS for close to 60 years. The TOS literature contains a number of corroborating examples and case histories with quite identical experiences regarding thief sampling. Several types of precisely designed and sharply honed empirical powder experiments allowed to draw the following draconic conclusions:

“In reality, regardless of whether a sample is large or small, a thief samples particles from a large region of the mixture. [. . .] thief samples can generate entirely misleading results. . . . Hence it must be concluded that accurate characterization of the structure of a powder mixture using a thief is not feasible. . . . [. . .] to assume that the state of the blend at a given stage of the manufacturing process predicts the final state of the blend is a risky proposition. Finally [. . .] even if a complete and comprehensive characterization of the state of mixture inside the blender were⁵ feasible, such a characterization would fail to predict the degree of homogeneity of the mixture as it is fed to a press or encapsulator, because mixtures can either mix

⁴ See however Matheron (2015; originally 1966) for an interesting historical intersection.

⁵ *Italics by present authors.* Also the comprehensive experience from TOS shows with complete assurance that such is plainly impossible.

or segregate extensively when discharged, conveyed, and vibrated.” . . . “One overall conclusion [. . .] is that given the state of development of blending and sampling *technology*⁶, the only reliable approach for characterizing the quality of a batch of tablets or capsules is to analyze the tablets or capsules themselves. Regulations mandating extensive analysis at earlier stages are misguided and are likely to result in misallocation of resources. If the goal is to assure that the process consistently generates a quality product meeting specific attributes (i.e. validation), this goal would be best most effectively achieved by increasing the number of tests performed on finished products. Analysis at earlier stages is valuable only as coarse indicators of potential problems.” (Muzzio et al., 1997)

However in this study (Muzzio et al., 1997) as regards terminology, there is no distinction between ‘sample’ vs. ‘specimen’ vs. ‘increment’ vs. ‘composite sample’. Also neither of these first two papers (Danckwerts, 1952; Muzzio et al., 1997) distinguishes between compositional and distributional heterogeneity (CH_{Lot} vs. DH_{Lot}), making comprehension more difficult than necessary.

Ghaderi (Ghaderi, 2003) took up Danckwerts analysis, extending it significantly by invoking a first look at TOS as well as linear time-invariant modeling (LTI). The evolution of mixing in a continuous blender is described by an axial heterogeneity function, which is decomposed into two components, the first describes fluctuations *caused by the feeding system*, while the second is due to the particulate nature of the material being blended (called the random component). Ghaderi took this latter to be modeled as a Gaussian bandwidth-limited white noise, with a variance that could be estimated by a variogram model taken over from TOS (N.B. not a geo-statistical model, see further comments below). This publication also presented many practical, systemic understandings of the mixing process which are important parts of the evolving understanding of continuous mixers (a.o. honoring the oeuvre of de Silva (de Silva, 1997)):

“Continuous mixers are in general unable to deal with fluctuations in feed components. Often expensive control systems become necessary [. . .] in order to ensure that feed fluctuations do not occur. Continuous mixers are very sensitive to malfunction of systems components. The design of a continuous mixer is often restricted to one application, and adaptation to other duties is often difficult. Successive operations are not possible in continuous mixers.”

Ghaderi introduced TOS’ more strict definitions of the concept ‘heterogeneity’ vs. ‘homogeneity’ and ‘uniformity’ (pharmaceutical practices), and was the first to quote Gy’s concept of *heterogeneity contributions* as the effective concept needed, and developed a clearer phenomenological model for how the axial vs. radial mixing is *supposed* to interact in an *ideal* linear continuous mixer by the ability to smooth out feeder axial fluctuations while only permutating parcels of compositional heterogeneity in radial directions. Importantly he also traced the history of the augmented Danckwerts–Weidekötter VRR definition, but notably still stayed within the confines of the end-result of an ‘ideal random mixture.’

The most important contribution of Ghaderi’s work lies in indicating that the variance of random component (the material component) of the heterogeneity function can be quantified by the nugget effect of a *variogram*. Ghaderi brought a key issue of the Theory of Sampling (TOS) into play for the first time, and rightly merits significant recognition for this. A.o. he pointed out that an

output stream exiting a mixer, if not influenced by any periodic variations from the feeder, i.e. if all periodic fluctuations from the feeder *could have been* completely eliminated⁷, would take on the form of a flat “low” variogram, signifying the unreducible, minimum variation attributable to the inherent material heterogeneity (CH_{Lot}) of an ideal feed passing unprocessed through the mixer. There is here an important link to our present proposal for the final effective VRR, see below, especially as it is traditionally *assumed* that there will be no sampling errors⁸, and because Ghaderi’s work only implicitly addresses $\text{var}(FSE)$. For these reasons, Ghaderi did not arrive at the same end-result as the present contribution. His final equation, Eq. (17), (although dressed up in a slightly different terminology than TOS) is technically correct but stops just short of TOS’ understanding of the full meaning of the variogram *nugget effect*, i.e. as a sum of all incorrect (ISE) and all correct sampling errors (CSE), plus the total analytical error (TAE), see Section 3 and Appendix A. In fact, the nugget effect with Ghaderi only appears as a correction to the mixer output measurement series in the limited experimental input-output application of Gy in the form of a small-scale output variogram correction taken to be a proxy *representation* of the CH_{Lot} of the ideal feed (Gy, 1981)⁹.

While Ghaderi’s quotation from (Gy, 1998, 1981), important for his coupling of LTI with a variogram approach allowing a theoretical refinement of the Variance Reduction Ratio (VRR) concept (developing the so-called Danckwerts–Weidekötter equation independently by another line of argumentation), was a significant theoretical achievement, it did not lead to a practical approach to characterize mixer output heterogeneity. Gy’s approach (Gy, 1981) is in practice only applicable in the context of *simulating* the workings of an ideal mixer. This is still a long way from an empirical ‘ VRR_{true} ’ via a practical approach for on-line determination of the output stream heterogeneity to supply comprehensive results for CFR 211.110 process and compliance testing. Ghaderi emphasized that the output variance from any mixer has three components, one that represents the inherited (dampened) feeder fluctuation variance, a second that represents the error variances from the sampling process (TOS’ major tenet) and a third that is attributable to the analytical error. It is the second component, the TSE inflected by the sampling process, and especially the bias-generating sampling errors (ICS) which forms the new crucial element, as forcefully argued in the present work.

The last milestone on the road towards more refined understanding of mixing in general is a review of ca. 50 papers from the 30-year period 1976–2006 (Pernenkil and Cooney, 2006), summarizing the gamut of knowledge regarding the many elements involved in continuous blending. This review gives an authoritative and comprehensive overview of the contemporary theory and practice behind continuous powder blending, establishing an updated platform from which the attributes of continuous blending can be understood across a very wide range of applications. A summary of the most important findings follow:

“Segregation scale and – intensity is reconfirmed as the primary two, independent conceptual elements with which to under-

⁷ Emphasis is on the conditional here, since this is hardly possible in the real world. But this notion can establish a concept of an ideal feeder system, which is useful in theoretical analysis and as an ideal target concept; cfr. footnote 1.

⁸ Truth be told, Ghaderi used the expression: “If sampling can be performed correctly . . .”. In the present TOS context this means: “If all incorrect sampling errors can be/have been completely eliminated . . .”. However this is not enough to let the practical process scientist get a grip on the general situation in which there is always a vestige of incorrect error effects present, which will increase the nugget effect and therefore also the sill.

⁹ Ghaderi acknowledges these shortcomings (pers. com. December, 2014), which will be fully resolved in a future joint paper.

⁶ *Italics by present authors.* The main thrust of the present work is to import the generic knowledge and much wider experiences embedded in the Theory of Sampling (TOS) to bear in parallel, and to illuminate the singular works cited here.

stand and analyze all mixing processes. . . . The key findings by Ghaderi are acknowledged, albeit only in brief [. . .]. One key statement is re-iterated however: An ideal blender's performance is a function of the feeder consistency [we may likely assume that this also holds for a real-world blender]. All theoretical and practical analysis of mixing operations and their efficiency is strongly dependent upon the 'scale of scrutiny' (TOS: 'observation module'). In pharmaceutical manufacturing this scale size is usually chosen to be the equivalent mass of a tablet or capsule, or a few multiples hereof." (Pernenkil and Cooney, 2006).

This thorough review highlighted that continuous blenders act like *low-pass filters* unfortunately allowing the low-frequency feeder fluctuations to pass through. "Blenders that can operate in both batch and continuous modes can be *expected* to have *similar powder behavior*" it was claimed, but the fundamental difference between continuous and batch blending systems (vessels, equipment) is likely of such fundamental importance, that each typical *system* rather should be investigated and characterized on its own merit. Much can be learned from systematic studies of batch blenders that can help in the design of continuous blenders and their operation, and recently a comprehensive treatise presents these matters in great width and in considerable depth (Cullen et al., 2015).

However in all of the above, there is no understanding of the full, adverse effect from sampling errors. A sampling bias will inflate the total sampling error but in an inconstant fashion, decrease the possibility of correctly evaluating the adequacy of mixing and hinder compliance with CFR 211.110.

5. Logical resolution – a call for a paradigm shift

Bringing in the full portent of the Theory of Sampling, the following logical conclusions can be drawn to constitute a framework with which to ascertain the heterogeneity of a blended product as above, or below a regulatory or a scientifically established variance threshold.

This endeavor specifically must be based on an *empirical* approach, not on any preconceived notions or idealized assumptions: mixing end-results may, or may not, follow simple statistical distributions (but if so, after full validation, must represent the full lot volume). Mixing end-results are often of a more irregular disposition than what can be captured by a simple statistical distribution. The fact that one *can* force a very few analytical values, originating from a restricted, rigid 3-D array, into a statistical distribution model is not *ipse facto* proof that this is a sufficient approach with which to deal validly with the residual heterogeneity. The crucial point is that one will never be able to assess the true status of the residual heterogeneity if based on the self-defeating within-blender sampling approach.

Currently the following assumptions rule the day:

- Ideal blenders act like low-pass filters, cutting out high-frequency heterogeneity deviations – It is *assumed* they do this to good measure. But how well do they succeed in removing all such?
- Input deviations, most often conceived of as quasi-periodic fluctuations caused by the specific feeder used, are assumed to be completely removed, leaving only the inherent material-specific heterogeneity variability, CH_{Lot} – this is the hallmark of an *ideal* blender. In the pharmaceutical a.o. literature this desired end-state has often been designated as an *ideal* random mixture. But how realistic are these assumptions?

An honest view of real-world blenders, forces the following modifications:

- A real-world blender will never remove all traces of fluctuations stemming the feeder input characteristics, no matter the effectiveness of the longitudinal and radial back-mixing. There will always remain heterogeneity vestiges in a blender (dead-zones; stagnated material), which reflect the material-dependent residual minimum DH_{Lot} , unavoidably giving rise to GSE error contributions. Crucially these will not be restricted to a set of fixed locations, but will be present over the entire blender volume to a larger or smaller degree. This will assuredly modulate the output material stream characteristics and will therefore leave imprints hereof which conveniently can be observed and interpreted from the information present in a variogram of the blender outflow. Poor mixing must, and can, be detected without the ambiguity from sampling errors.
- It is unlikely (in the extreme) that any mandated, fixed scheme of a small number of sampling locations will be able to capture the existing DH_{Lot} in a representative manner a.o. because of the irregular distribution of the residual minimum heterogeneity within the full lot (blender) volume, and the variation in the physical properties of excipients and API used. This is all the more so when the full impact of biased and non-optimized physical sampling is brought to bear.

However:

- TOS-optimized 1-D sampling guarantees that all deviations from the ideal random mixture signature will be effectively captured by a variographic characterization; this is only a matter of the resolution selected for the basic lag = 1 scale. Only a completely flat variogram with a sill below an a priori determined variance threshold signifies a sufficiently complete blending result, all other variogram morphologies signify *incomplete mixing*, as shown in Fig. 4 .
- The blender outflow variogram will also be influenced by still-existing ISE, in addition to the full impact from the CSE and TAE, conveniently appearing as the sum-total of the nugget effect. The nugget effect must be *estimated* (but this is not a problem with TOS, which furthers several practical approaches). A TOS-correct variographic experiment must be performed, but this is not a problem for any industry as the procedures are very well-known from scores of published works.
- In the blender output variogram the 'true' process variation, i.e. the effective residual variability is not represented by the sill level but by the *corrected sill*: $sill_{output} - n.e._{output}$ as shown in Fig. 4 (notation explained in figure caption). The more a variogram represents the final state of a well-mixed final blend, the smaller this difference and the lower the sill. Any non-zero differences quantify, that less-than-ideal blending has occurred, (Esbensen and Romañach, 2015).
- Thus the realistic, 'true process variability' is shown by the following index, contrast with Ghaderi (2003)¹⁰

$$[VRR_{corrected}]^{-1} := \frac{[sill_{output} - n.e._{output}]}{[sill_{input} - n.e._{input}]} \quad (2)$$

¹⁰ The VRR is defined, *in analogo* to the signal-to-noise ratio, as the ratio of the input variance to output variance; a good mixer increases the VRR. This is partly at odds with its own name, the variance *reduction* ratio. In the present treatment focus is on the practical variance reduction coming out of the mixer, hence the present 'inverse' definition in Eq. (2).

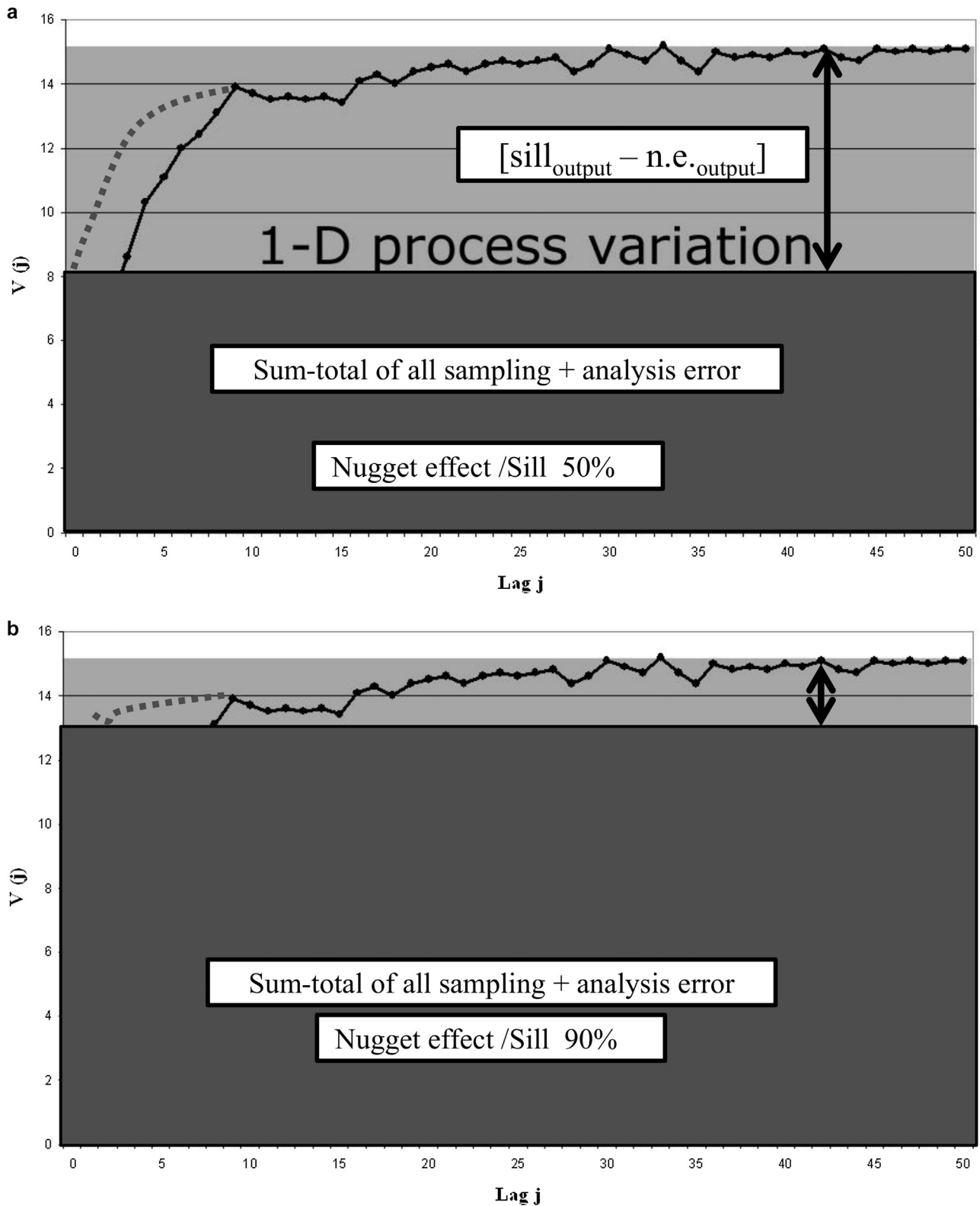


Fig. 4. (a) Variogram expression of a situation pertaining to e.g. an ideal blender in which the measurement system is no longer able to deliver a reliable depiction of the process/product variability. $RSV_{1-dim} \sim 50\%$. (b) Variogram expression of a situation in which the measurement system is completely unable to render any process/product information. $RSV_{1-dim} \sim 90\%$. (n.e. = nugget effect; lower indices "output" or "input" refer to whether the stream characterized is the input, or the output stream to/from the blender).

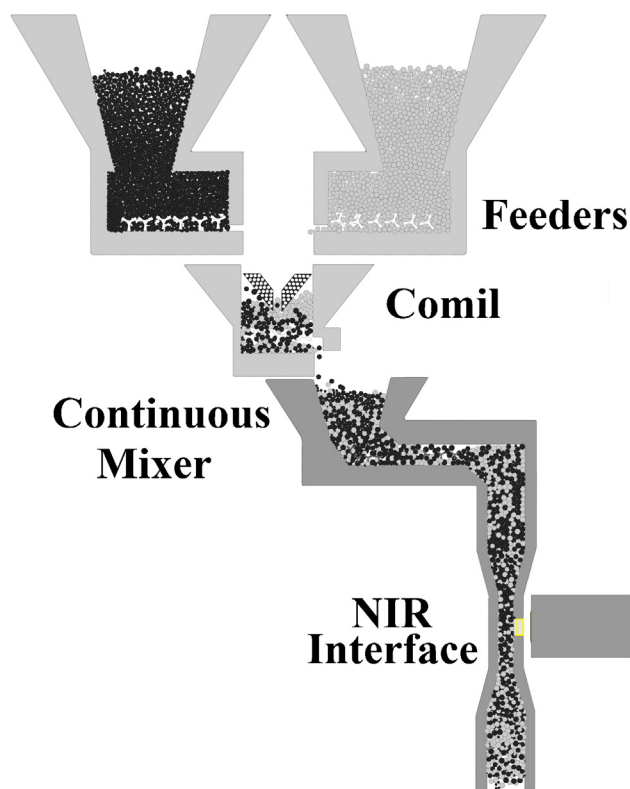


Fig. 5. Schematic of the continuous manufacturing setup at Rutgers and the interface for the FT-NIR system in the chute that brings powder from the mixer into the tablet compressing machine. The NIR analyzes approximately the top 1.0 mm of the 20 mm thick powder flow at the interface. NIR spectral characterization shall ideally correspond to a full slice of the falling stream at all instants.

(input and output with respect to an ideal blender).

- i) There are severe difficulties in implementing the above theoretical derivation, especially the practical approach to characterize the input stream variogram with the same measurement system as is used for the output stream from the blender. The concepts and mechanisms behind the $VRR_{corrected}$ contribute to theoretical understanding – while the information needed for process improvement is available from output stream variogram alone, viz. $[sill_{output} - n.e._{output}]$.
- ii) This index is self-calibrating; the *effective* remaining total sampling and analysis error (GEE) is subtracted from the sill level estimate, leaving the true process (product) variability to be expressed – for all types of mixers and blenders, under all conditions.

Fig. 4 illustrate various cases of $[sill_{output} - n.e._{output}]$, from acceptable to unacceptable. Fig. 3 (above) already illustrated the characteristics of a good total process analytical system.

6. Summary of recommendations

The authors recognize that the approach discussed is quite different to the current practice in the pharmaceutical industry. Our recommendations are summarized below:

1. When the objective concerns a *reliable estimate* of the full lot composition, grab sampling (obtaining only ‘specimens’) from fixed locations within blender must be terminated with extreme

prejudice. This especially applies to thief sampling and PAT methods where only one specific location is analyzed.

2. Instead perform *composite sampling* where all parts (increments) of the lot have the same opportunity of being sampled (FSP) and the utmost care shall be taken to avoid affecting the material microstructure by the sampling process itself.
3. Only accurate (ISE-minimized, i.e. unbiased) sampling procedures to obtain *localized samples* (not specimens) shall be used when this is called for as part of the process monitoring/QC agenda.
4. Develop all sampling strategies based on proper TOS understanding of all sampling errors and how to avoid them (ISE, CSE). Always comply with FSP.
5. Validate all sampling processes using a replication experiment or variographic analysis (Danish-Standards-Foundation, 2013). This can be carried out for both batch and continuous processes.
6. Evaluate possible improvements in the mixing process, always taking note of what is possible with the existing total process monitoring system by comparing MPE to the sill.
7. Use validated sampling methods to establish TOS-based fit-for-purpose specifications for process monitoring and product QC.
8. Evaluate the effective (“true”) process variations based only on TOS-compliant design of sampling approaches.
9. Collaborate with instrument vendors to develop ISE-minimized analytical (PAT) systems that follow TOS principles, i.e. that analyze an entire cross-section of the product or process flow. This is not necessarily a straight-forward development, but certainly doable.

The term “Theory of Sampling” does not imply a solely theoretical approach. On the contrary TOS includes all the necessary practical methods to achieve accuracy and precision, including variographic analysis, that have provided many benefits in very many other process industries, e.g. the mining, metallurgical, petroleum, agricultural and environmental industry sectors – and which can equally well provide similar benefits to the pharmaceutical industry. The case studies below illustrate the implementation of TOS in practice with examples related to pharmaceutical materials and processes demonstrating how both sampling and mixing processes can indeed be improved.

7. Implementation in practice

The following case studies illustrate the proposed solution for two cases, continuous and batch manufacturing:

7.1. Continuous manufacturing example

This example presents an acetaminophen (APAP) powder mixture prepared in a continuous manufacturing pilot plant, equipped with Schenk feeders, a Gericke blender and a total flow rate of 20 lbs/h. Fig. 5 shows the experimental set-up. The first feeder delivers APAP while the second contained silicified microcrystalline cellulose (SMCC) (excipient) already mixed with 1% magnesium stearate (MgSt) (lubricant). A co-mill is placed after the feeders and used for de-lumping of the powders. APAP is a very cohesive drug and is notoriously difficult to mix (Osorio et al., 2014). APAP is used as a realistic model drug (although not formulated here at the high level usually observed in commercial formulations).

Spectra were obtained with a Bruker Optics Matrix FT-NIR spectrometer (Billerica, MA) at a resolution of 16 cm^{-1} , each analytical determination averaging over 16 scans. The powder flows from the blender outlet down a cylindrical chute described in (Singh et al., 2014a,b). FT-NIR spectra are obtained through a sapphire window adaptor connected to the cylindrical chute, Fig. 5. Together with the spectral averaging (16 scans), and a

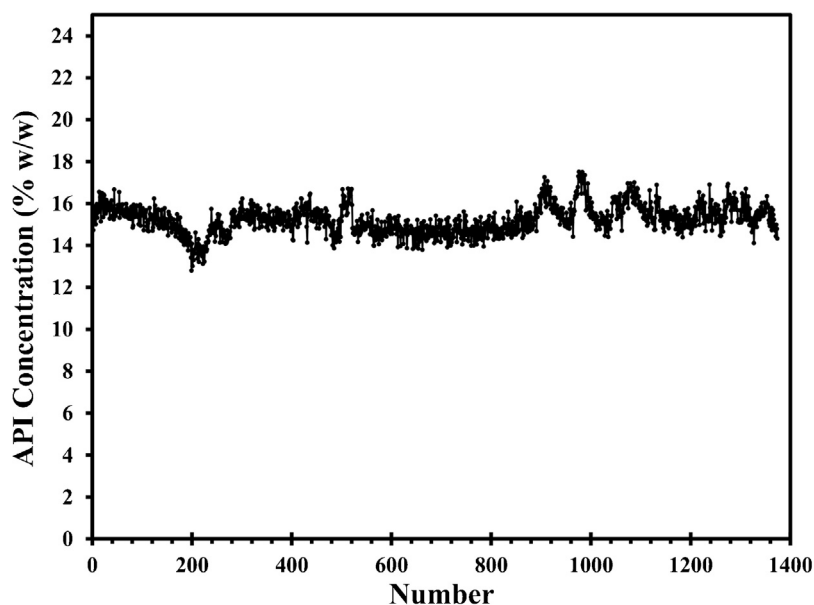


Fig. 6. Real time PLS-prediction results for the experimental 15% (w/w) continuous manufacturing run. The sample mass analyzed by each spectrum was approximately 90 mg.

bulk density of 0.5 g/cm^3 , this translates into a composite sample weight of approximately 90 mg, taking into account the irradiation penetration depth of about 1 mm (Colón et al., 2014). Fig. 5 shows that with this probe the NIR will not analyze the 20 mm thick cross section of the falling powder stream. This is then an increment delimitation error (IDE) that leads to a non-neglectable MPE. Thus, TOS points to an important opportunity to improve future PAT sensors, through the design of systems capable of analyzing an entire cross-section of the sample stream. The present NIR method was developed with full understanding that this limitation could inflate the MPE, and the need for future improvement because of its relevant illustrative and didactic power.

A multivariate PLS calibration model was developed for real time determination of APAP in the setup shown in Fig. 5. The calibration model was developed with powder mixtures of 10%, 12%, 14%, 16%, 18% and 20% (w/w) APAP. Approximately 100 spectra were obtained for each powder mixture flowing down the experimental cylindrical chute shown in Fig. 5. The calibration model was developed from 9000–4500 cm^{-1} with three PLS factors (determined from test set validation): spectra were pretreated with SNV-1st derivative (25 points) using SIMCA 13 software. The calibration model was used to predict a 15.0% (w/w) APAP validation blend (test set validation), resulting in an average value of 15.1% (w/w), i.e. a bias of +0.10%. Two-standard deviation RMSEP was $\pm 0.60\%$ (w/w), and relative coefficient of variation was therefore 4.0% based on the test set validation.

This model was applied to the continuous manufacturing pilot plant resulting in close to 1400 spectra. The PLS-predictions are shown in Fig. 6 where the average drug concentration was now 15.23% (w/w), with a standard deviation of 0.85% (w/w) and an RSD of 5.57%. The departures from a 15.0% (w/w) target concentration could be the result of inferior sampling, inferior compensation by the chemometric model, or both. Regardless, the full complement of probe sampling errors, chemometric modelling errors and the analytical errors is conveniently summed up as the effective MPE as delineated by the derived variographic characterization.

The variogram shown in Fig. 7 Fig. corresponds to the full prediction data series (lag times range from 1 to 700). The unit lag time corresponds to two consecutive APAP concentration determinations. The $V(j)$ values increase slowly until a value of

approximately 0.5 is obtained, a level occasionally exceeded by irregular quasi-periodicities. In spite of these periodicities the $V(j)$ values return to 0.5, which is the value that can be considered as the overall sill for this process, which shows a range of 180 lags. While a completely ‘homogenized’ mixture per definition must correspond to a completely flat variogram, in the present experiment an increasing variogram is observed. Periodicities are observed of a magnitude of ~ 70 lag units (with a less distinct multiple at 140) as well as at higher commensurate lag distances. These irregularities may be related to the fact that the continuous manufacturing system used in this experiment did not have an automated system for regulating the powder flux within the chute where the NIR spectra were obtained (in this pilot experiment, the chute outflow was simply controlled manually). The strongly increasing feature at all lags above ~ 550 is of no consequence; this is because the entire experimental runs in not fully stationary, indeed over the full 1374 measurements a weak, but distinctive small negative trend is prominent. However a de-trended

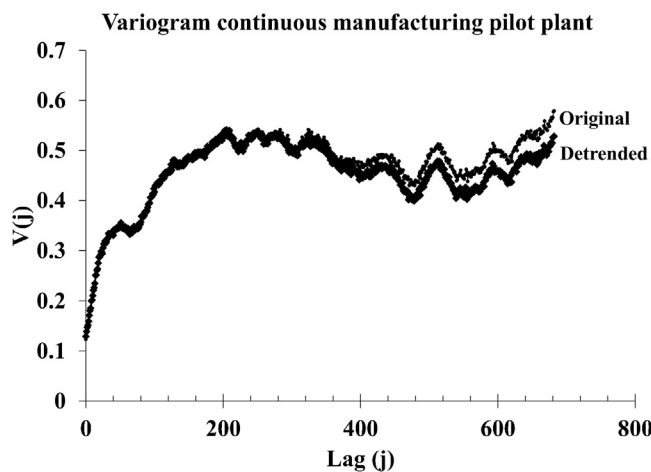


Fig. 7. Continuous outflow variogram for the process data shown in Fig. 6. Observe the magnitude of the nugget effect (approx. 0.12), which compared to the sill level (approx. 0.5) makes for a RSV_{1-dim} of $\sim 20\%$. Also shown is a de-trended variogram, which does not distinguish itself in the important lag interval 1–375.

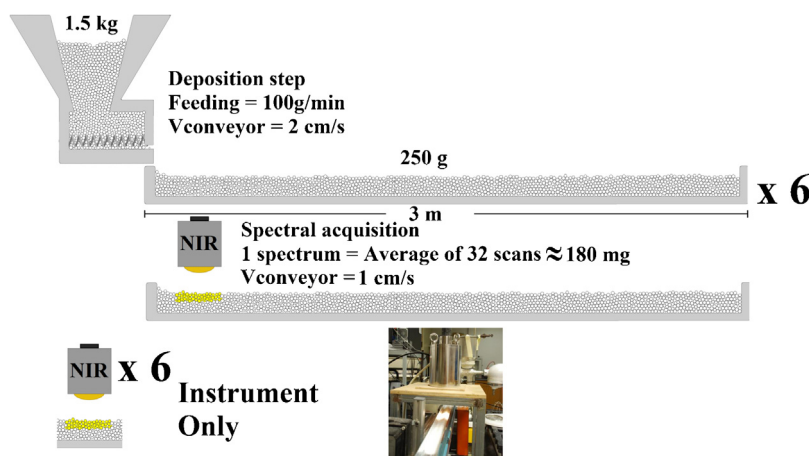


Fig. 8. Replication experiment setup for characterization of the three-stage sampling/deposition/moving PAT analysis process.

measurement series gives rise to the exact same variogram feature for the only important lag segment in the interval 1–375, Fig. 7, so all conclusions are identical de-trending or not.

Fig. 7 provides very important information on the total measurement system error. Following well-established TOS procedures, DS 3077 and references herein (Danish-Standards-Foundation, 2013), the first five $V(j)$ values of the variogram were extrapolated backward to estimate the intercept of the ordinate axis¹¹. This is a practical estimate of the magnitude of the nugget effect. This is a fit-for-purpose estimate of the minimum possible error (MPE), here ≈ 0.12 , which compared to the sill level (approx.0.5) makes for a RSV_{1-dim} of $\sim 20\%$. This then represents the full complement of TSE + TAE. MPE is an appropriate measure of the absolute minimum error that can be engineered in practice using this particular sampling interface and PAT calibration model if measurements were actually taken corresponding to the minimum sampling lag used in the variographic characterization. This variogram is one of the first performed with a continuous mixing system, and future experiments will provide more much valuable information to improve processes.

The sill, range, and MPE constitute the full set of variogram features reflecting the principal sources of variation and constitute an approach to meet the demands of the Process Validation Guidance (U.S. Department of Health and Human Services, 2011). The variogram MPE is the best possible estimate of the 'effective, sum-total uncertainty' in a fully realistic practice, while the range and sill indicate the particular residual variation features associated with a blending process that has not reached full or satisfactory completion. This is because the difference between MPE and sill is related to the residual heterogeneity of the blend. Thus, the present process can indeed be significantly improved, as evidenced by a samplings quality index of slightly above 20%. These results substantiate the considerable challenges of mixing the cohesive APAP particles (Osorio et al., 2014).

The variogram should always be viewed together with a plot displaying the analyte concentration in real time, Figs. 6 and 7, jointly illustrating the variation in the drug concentration throughout the continuous mixing run. Both plots are needed to provide the necessary complementary information.

8. Batch mixing

Batch mixing may also be evaluated through variographic methods. The only requirement is to set-up a 1-dimensional outflow sampling/analysis system after mixing termination. This is not an additional requirement as batch processes are eventually converted into 1-D lots when the powder mixture from a batch blender flows into a compression machine. The 1-dimensional system may for example include prediction drug concentration with a NIR method as powder f.ex. moves on a conveyor belt (Colón et al., 2014; Roman Ospino et al., 2015; Sánchez Paternina et al., 2015), flows down a chute to the compression machine like in the continuous manufacturing example above. N.B. the 1-dimensional system may also analyze a complete series of tablets discharged from the manufacturing process. The essential requirement is to respect the order in which the drug concentration is obtained for samples permitting calculation and plotting of the variogram.

The following example shows the successive development of a batch mixing process using lab scale experiments. Small in-house developed blenders, capable of working with up to 1.5 kg of materials, were used for all experiments. Two mixing procedures were evaluated: (1) Mixing in a tumble blender for one hour – (the T process); 2. 30 min of vibration followed by 90 min of tumble blending – (the VT process) (Roman Ospino et al., 2015; Sánchez-Paternina et al., 2015). The blends were prepared from lactose monohydrate Granulac (Meggler Pharma), microcrystalline cellulose Vivapur 102 (JRS Pharma) and semi-fine acetaminophen (APAP) from Mallinckrodt Inc. (Raleigh, NC). The lactose monohydrate was passed through a U.S. Standard Sieve 60 (250 μm opening) before mixing. The target concentration for the blends was again 15.0% (w/w) APAP. A 1.5 kg powder blend was prepared and an in-house developed screw feeder was used to obtain six 1-D sub-samples of approximately 250 g. Each 250 g sub-sample (1/6 total blender lot volume) was deposited along a rig of 3 meter length, 4 cm wide and 3 cm deep that moved over a conveyor belt. NIR spectra were then obtained along the entire length of the rig.

The spectral acquisition parameters were chosen to obtain a composite sample mass of approximately 180 mg, taking into consideration the bulk density of the powder blend, the acquisition time for the NIR spectra, the linear velocity of the conveyor belt, and the diameter and depth of penetration of the NIR beam (Colón et al., 2014). The FT-NIR system analyzed the top 1.2 mm of the powder bed. This was confirmed by increasing the thickness of a lactose powder bed over a layer of talc, until the talc bands were no longer detected. A total of 32 FT-NIR spectrometer scans were averaged resulting in a total spectral acquisition time of

¹¹ This is only a first order estimate subject to some variability as a function of exactly how many $V(j)$ variogram points to include in a simple regression modelling. Pitard (1993), Esbensen et al. (2010), Esbensen and Paasch-Mortensen (2010), but this is all that is needed to present the essential information as to the relative magnitude of the nugget effect vs. the sill, Gy (1998).

Table 3
Comparison of the two mixing methods.

	Deposition (n = 6)			Replicate of single deposition (n = 10)			Repeatability study (n = 6, at 10 points)		
	VT	T1	T2	VT	T1	T2	VT (n = 6)	T1 (n = 10)	T2 (n = 10)
Average% (w/w)	14.93	15.17	16.39	15.21	14.63	16.24	15.78	15.82	15.54
Spectra (#)	390	258	390	647	570	650	36	60	60
SD% (w/w)	0.78	1.74	1.58	0.34	1.06	2.02	0.14	0.14	0.17
RSD (%)	5.20	11.46	9.62	2.23	7.25	12.46	1.3	0.88	1.12
MPE	0.04	0.9	0.7						
Sill	0.7								
SD Blend ^a	0.75	1.46	1.34						
Inherent Blend SD ^b				0.31	1.05	2.01			

^a After subtracting the MPE and short term instrument error (determined from repeatability study).

^b After subtracting short term instrument error (determined from the repeatability study).

approximately 4.4 s. The powder blend moves at 10 mm/sec while the 32 scans are obtained throughout 4.5 s providing a composite sample. A validated NIR spectroscopic method was used to determine the drug concentration as the powder deposited in the rig moved at a linear velocity of 10 mm/s. The validation efforts were not limited to the analytical method, but also included a Replication Experiment, as described in DS 3077 (Danish-Standards-Foundation, 2013), to study the compound sources of error in the sampling and measurement set-up.

Replication Experiment studies were performed as shown in Fig. 8 (six successive rig depositions, 10 times dynamic PAT analytical determination to-and-fro over just one deposition). The first part of the replicate experiment was conducted by performing six depositions each of approximately 250 g along the 3 m rig. This setup yielded approximately 65 spectra per deposition. The Matrix FT-NIR spectrometer was placed at a height ~10 cm to obtain spectra as the rig moves at 10 mm/s. This constitutes the first sub-sampling, obtaining six sub-samples of 250 g from the 1.5 kg lot. The APAP drug concentration was predicted for each spectrum using the validated FT-NIR multivariate calibration model (Martens

and Naes, 1992). Thus, every spectrum is a de facto sub-sample of the powder blend, with a mass of approximately 180 mg. The first experimental part summarizes the variability stemming from all sampling and measurement steps. The second part of the replication experiment consisted of moving one of the full length 250 g deposition over the conveyor belt to and fro 10 times, obtaining spectra from one end to the other. This part excluded the sample deposition step and determined the precision with which one 250 g sub-sample could be measured, as per a replication experiment.

The final experimental part consisted of an analytical repeatability study whose precision was determined as six consecutive spectra were obtained without moving the powder mixture or the spectrometer under the same operating conditions within a very short time interval (ICH, 2005).

The left side of Table 3 shows the results obtained for the n = 6 depositions in the entire 3 m rig clearly showing that the VT process is superior in reducing the heterogeneity of the powder mixture and obtaining a mixture with concentrations near the 15.0% (w/w) APAP target level. The plot of drug concentration vs. spectrum number shown in Fig. 9 also provides the same conclusion. The breaking up of particles due to vibration also improved the flow properties of the powder mixture. The central graph (T1) shows less drug concentration results due to accidental difficulties in powder flow and deposition onto the 3 m rig. The VT process showed the lowest standard deviation (0.78% w/w APAP), at least half of those obtained with the T process. The VT process also has the lowest sill and MPE. The VT process shows a nugget effect – minimum practical error (MPE) of only 0.04% as shown in Fig. 10.

In the VT process in Fig. 10 (bottom), the MPE in comparison with the target level of 15.0% results in a very low RSV(%), approximately 0.20% only. Thus, the sampling and analysis system is capable of providing a suitably low sum of all correct and incorrect sampling errors + the total analytical error (TAE). However, MPE also depends on the specific heterogeneity of the blend (the correct sampling errors, FSE and GSE, depend on the heterogeneity of the blend). Thus MPE is indeed greater for the more heterogeneous blends T1 and T2 as shown in Fig. 10. The empirical MPE and sill levels can be used to establish a threshold for the mixing process. It is likely that a procedure general enough to catch all mixing processes can be developed based on this principal observation.

The number of samples analyzed by following this approach was much greater than those that could be obtained with a sample thief, and were obtained without affecting the properties of the powder blend during powder flow – yet they represent a simple, routine PAT implementation. The samples analyzed were not based on previous assumptions on areas of possible incomplete mixing within the blender etc. in fact this approach very easily discovered

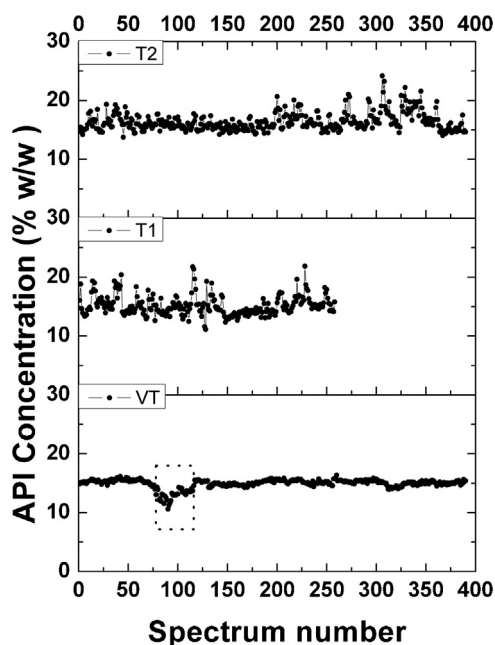


Fig. 9. Real time PLS-prediction results for the experimental 15% (w/w) blends mixed in a batch process but analyzed by deposition into 3-m long rig (1-dimensional system). The sample mass analyzed by each spectrum was approximately 180 mg. Note the interval of spectra removed from the variographic re-analysis described in text and Fig. 11.

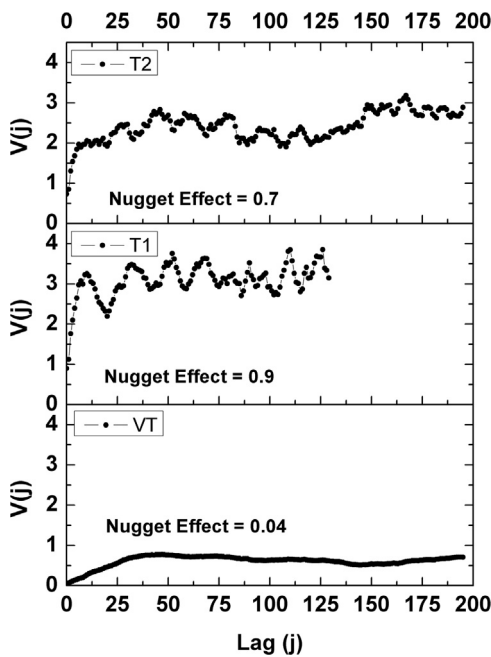


Fig. 10. Variograms for the three mixing experiments evaluated. The VT process leads to a much lower level of heterogeneity and much lower sill and nugget effect (MPE) values.

areas of significantly incomplete mixing in the VT process, as seen in Fig. 9 bottom in the conspicuous drop in concentration observed from spectra #78–116. This was easily observed because the entire manufacturing batch volume was available for inspection.

The outflow approach now allows design of plans to improve the clearly insufficient mixing process and how to establish suitable specifications for a process. A simulation of the improved process was performed by excluding the samples in the interval #78–116 (shown in Fig. 9 bottom). A renewed variogram is presented in Fig. 11 (right), in which the MPE is identical, while there is a very notable reduction of the sill level. Renewed estimation of the RSV_{1-dim} results in a value of 2.6%. The withdrawn draft guidance required the analysis of drug concentration for at least 10 blends from a tumble blender with: (1) a relative standard deviation $\leq 5\%$, and (2) all individual results within 10.0% (relative) of the mean drug concentration (U.S. Department of Health and Human Services, 2003). If the blending process were improved to eliminate the sudden drug concentration drop shown in Fig. 9 bottom, then the RSD in drug concentration reduces to approximately 2.6% and all values would then be within 10% of the mean drug concentration stipulation. Without improvement, the VT process would not meet the requirements of the recently withdrawn guidance. The sill for the improved process (Fig. 11, right) could be a suitable threshold for the process.

Table 3 also shows the results from the second part of the replication experiment where one of the full length 250 g depositions was moved over the conveyor belt to and fro 10 times, obtaining spectra from one end to the other. The standard deviation of the VT process is now 0.34% (w/w) APAP, while the T process blends show standard deviations of 1.06 and 2.02 respectively. These results reflect the significant differences in the residual heterogeneity of the blends obtained. The standard deviations observed in this second part are lower for the VT and T1 blends than those observed in the first part of the replicate experiment. This result is expected since this second part only involves one sub-sampling operation (performed with the NIR calibration model). The T2 blend shows a higher standard

deviation, which is likely associated to the high heterogeneity observed.

The analytical repeatability study results shown in Table 3 also encompass the contribution from the short term instrument variation. The variance from the repeatability study may be subtracted from the variance obtained in the second part of the replication experiment ($n=10$), as indicated in the row marked “inherent blend SD” in Table 3. The values obtained show that the short term instrument variation (effective analytical repeatability) is relatively low in comparison with the blend heterogeneity.

Table 3 includes a row marked “SD Blend” where the variance associated with the MPE and the repeatability studies have been subtracted. The results shows that for the VT blend the sampling error and the instrument error only have an absolutely minimal effect, as the standard deviation is only reduced from 0.78% (w/w) to 0.74% (w/w). The MPE may also be subtracted from T1 and T2 as shown in Table 3, showing that most of the variation is related to the heterogeneity of the blend and not to the sampling + analysis system.

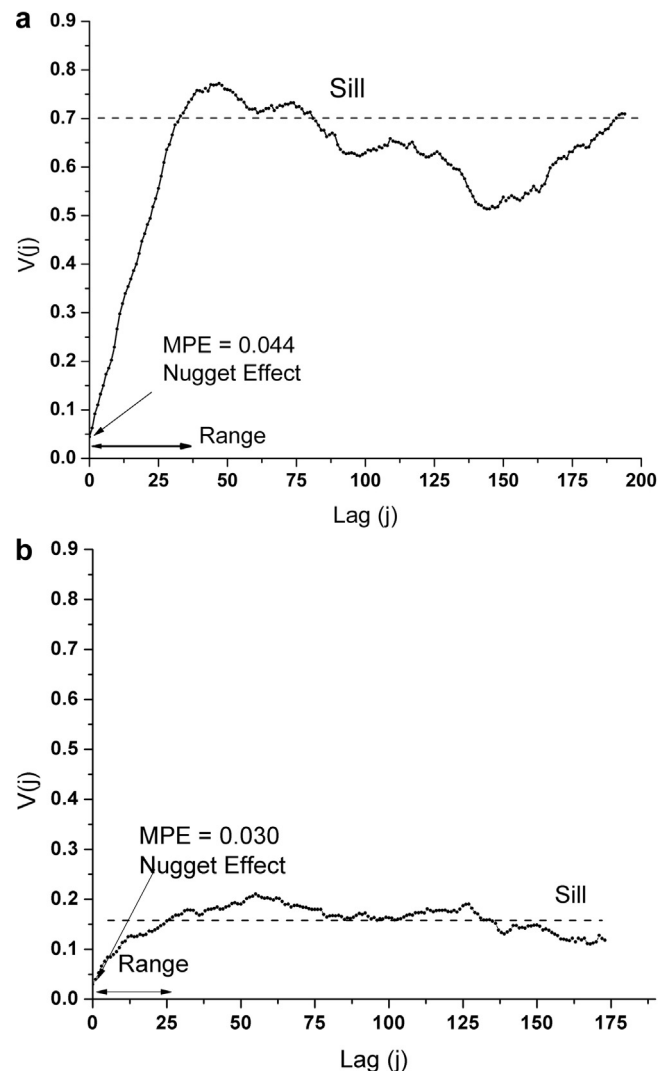


Fig. 11. (a) Variogram based on the total 390 individual analyses of the complete lot (six depositions of the 3 m rig length). The range is 30–35, nugget effect=0.04, sill=0.7. (b) variogram excluding spectra #78–116.

In both the continuous and batch blending examples mixing monitoring may be improved as previously discussed¹²:

$$[\text{VRR}_{\text{corrected}}]^{-1} := \frac{[\text{sill}_{\text{output}} - n \cdot e_{\text{output}}]}{[\text{sill}_{\text{input}} - n \cdot e_{\text{input}}]} \quad (2)$$

This equation and the TOS approach represent a quality by design alternative to evaluate mixing processes. The sampling errors are minimized by proper TOS considerations, and a sampling strategy is designed without unrealistic *assumptions* as to the areas of incomplete mixing within the blender. The sources of variation are also adequately identified as discussed in the process validation guidance (U.S. Department of Health and Human Services, 2011).

9. Pharma application – changes are necessary

The above exposé calls for a rational TOS-based approach which must not be based on sampling from within containers. All sampling from within a pharma mixer – especially if mandated to be carried out by thief/spear sampling – should be discontinued. Only TOS allows full command over the fatal ISE without which all attempts to reach the true minimum residual mixing heterogeneity will remain futile. The proposed approach allows taking samples from throughout the entire lot, and all parts of the lot will have the same opportunity of being selected because of a lot transformation from 3-D to 1-D. The variogram and the TOS based approach can be used in combination with proposed statistical approaches to evaluate the data (Bergum et al., 2015; Garcia et al., 2015) with which to develop suitable specifications. It is not the job of the present work to develop the exact numerical thresholds; this is a task for a concerted action involving regulators, the pharma industry and principal researchers.

After ISE elimination, further output variability reduction can be addressed with standard approaches e.g. better process control, better input/feeder stability, and/or higher increment sampling frequency. If the experimental variographic analysis is based on a significant over-sampling frequency, there is also the extremely powerful tool of increasing the number of increments making up the routine process sample (Esbensen et al., 2007). TOS furthers a plethora of theory, practice, equipment design and solutions, see DS 3077 (Danish-Standards-Foundation, 2013) and the normative reference list herein.

Main thrust of paper: From the above it follows that it is only possible to ascertain with the necessary level of confidence the heterogeneity status of a finished blended product if/when based on the *outflow stream*. This should be viewed as a *linear mapping* of the complete container content, regardless of the inaccessibility of some of the content locations due to blender geometry. This mapping is the only way to guarantee with 100% confidence that the Fundamental Sampling Principle can be upheld. The current regimen with a small number of fixed locations (irrespective of whether with one or three replicate samples from each) is a fatal transgression of FSP, paradoxically almost *guaranteeing* non-representativity and making it impossible to establish adequate specifications.

Since the outflow stream from all blending operations in the pharmaceutical realm will always be subject to *additional* transportation and manipulation towards a tableting or encapsulation operation (be this by conveyor belt, pipeline, or other means), it will *per force* be subjected to some measure of additional *induced* segregating impact, large or small. This makes *in-vessel*

heterogeneity testing (blender testing) structurally flawed and incommensurate with respect to testing of the final product. This is no doubt the reason why regulators demand both in-line heterogeneity testing and final product compliance testing.

The logical conclusion is inescapable: a well-calibrated PAT system that includes real time blend uniformity monitoring can be used to control a pharmaceutical blending process with any degree of fidelity required. The PAT system will be helpful as demonstrated in a number of studies, and future improvements will be obtained with systems capable of analyzing an entire cross-section of the sample. However, the control protocol must also consider analyses *subsequent* to the last impacting segregation event. The real time heterogeneity testing provides control of the mixing process, while the final product testing provides a fully representative testing procedure.

With TOS this conclusion can be forcefully augmented: this quality control can conveniently be carried out as a variographic control. This can either be based on conventional sample outtakes (TOS-compliant of course), but would be especially effective if based on a relevant instrumental PAT approach (NIR, UV, Raman, imaging . . .). This would in principle allow *complete* control of all tablet or dosage units, but would in practice probably be based on a more reasonable standard statistical sampling rate of the linearized lot (consider as a population of tablet-volume units in this case). The principal issue is that by performing both the blending control and the final product control as one-and-the-same output variogram monitoring, the door is open for using any lag definition desired from lag: = 1 tablet/dosage upwards. A lag larger than 1 corresponds to an increased level-of-scrutiny scale (TOS: observation module). Any scale level (higher-than-1 lag level) is possible as this is hardly even a technical issue when combined with an automated PAT system of monitoring and analysis.

Suppose, for the sake of illustration, that a particular system is characterized by an MPE that reflects satisfactory low TSE + TAE levels. Suppose that the sill is only slightly higher than this MPE – i.e. a flat variogram for all lags including all lags below the range. This can only mean that the product is at the desired minimum residual heterogeneity, consistent over all scales, and that the variability observed is indeed only due to the minimized total measurement system error. This will be a reflection of a *fit-for-purpose*, statistically stable manufacturing situation. This will be the desired signal that one cannot mix better, under no circumstances, based on the contemporary total measurement and monitoring system. For fully validated systems (process and sampling) like this, monitoring can be based on any lag, which will perform the necessary QC. Assuming the same total measurement system effect (MPE), an increase in the sill level signifies a higher process variability than this ideal state, i.e. a higher between-unit variability. This may, or may not be within the stipulations for a *fit-for-purpose* threshold, but at a certain too large sill (a too large $\text{RSV}_{1\text{-dim}}$) it is necessary to conclude (on-line QA) that the mixing is no longer up to the required standard.

The crucial issue is that by switching to the proposed variographic blending and final unit control the way forward is completely prepared for the desired level of quality control. The variographic approach allows a full-scale process in the lab (obviously under realistic pilot plant conditions etc.), addressing the full production scale and any scale-up issues present in a separate stage. The fact that the manufacturing process has already been validly optimized in the laboratory, makes this important scale-up issue significantly easier to address, and in any event the Theory of Sampling will be available for any unforeseen sampling issue that might crop up at all larger scales. These features will be especially effective in combination with PAT.

The proposed ‘end-game’ QC/QA paradigm is also the logical answer to the regulatory demand of optimal consumer protection:

¹² Strictly speaking, VRR is not defined for the batch mixing case, since the input heterogeneity is not well-defined, but for quality control purposes, the right hand side of the equation works absolutely according to its purpose, since the output variogram is always defined and as shown easily realized.

'optimal' can be directly translated into the specific lag (i.e. sampling rate) with which to conduct the final product inspection. In all likelihood inspection *may* be satisfactory also if not based on total inspection – this would depend strongly on the overall process control achieved in the preceding manufacturing/processing stage(s). In any event by invoking the present proposal the road is fully open to conduct simple, yet conclusive on-line validations that will demonstrate the degree of process stability achieved. A stable, identical form and level of the variogram from recurrent variographic analyses carried out at regular intervals will signify not only a process with the necessary stability, but in the case of a stable MPE magnitude, that the total measurement system is also demonstrably stable.

The proposed regular variographic monitoring and QC approach is self-checking and self-correcting, in as much, as an example, an increasing measurement system degradation will immediately be observable as an increasing MPE. And the MPE will always be subtracted, to reveal the true effective process variability. As such if a potential measurement system degradation continues to develop, the effective mixing index (or production efficiency index) $[sill_{output} - n.e._{output}]$ will become smaller and smaller, ultimately leading to a complete loss of the possibility of a valid process/product monitoring, see Fig. 4.

10. Conclusions

Variographic analysis allows problem-dependent breakdown of the inherent measurement system uncertainty (comprehensibly summarized as the MPE or nugget effect) and the blend heterogeneity ("true process variation").

Variographic analysis allows unequivocal continuous quality checking of the measurement system, which is only acceptable as long as the MPE remains at the same level as when it was verified during the laboratory development phase.

TOS allowed the definition of a fully realistic 'mixing index' for the first time, presented here in its theoretical form of $[VRR_{corrected}]^{-1} = [sill_{output} - n.e._{output}] / [sill_{input} - n.e._{input}]$. But for practical industrial process monitoring (mixing and blending monitoring), the self-correcting empirical output variogram version suffices: $[sill_{output} - n.e._{output}]$.

We here call for a paradigm shift in regulatory demands re. pharmaceutical mixing process and product compliance testing – we call for a discontinuation of the futile, inherently flawed in-container testing, especially when demanding the use of thief samplers but also to other physical or PAT sampling systems if similarly incurring untoward ISE; exceptions *may* exist, but can only be considered if based on extensive variographic testing as introduced. It *may* for example turn out that *sensor* installations on various industrial mixer types *might* show a low enough, fit-for-purpose error that facilitates the control of the mixing process. However, the subsequent mixer outflow segregation possibilities would not be included and end-product QC is still required. The latest two decades of accumulated knowledge regarding continuous mixing coupled with the impact of The Theory of Sampling (TOS) suggests to continue to improve PAT sensors to control mixing, and to switch to variographic characterization of the final product stream en route to the tableting machine, or by the final tablets/capsules/dosages.

It is hoped that regulatory authorities, as well as leaders in pharma, will see the advantages of a systematic introduction of variographics; all that is needed is professional TOS competence.

Acknowledgements

This work is supported by the National Science Foundation Engineering Research Center for Structured Organic Particulate

Systems, through Grant NSF-ECC 0540855. The authors also thank Dr. Ravendra Singh for his help in running the continuous mixing system at Rutgers, Krizia Karry for calibration spectra in the continuous system at Rutgers, and Elvin Almodovar for help in developing the Matlab code for the variograms. The first author thanks his institution in the period 2010–2015, GEUS, for permission to develop a full GDAS educational program, with wider application vistas, not necessarily only with a geo-prefix; such managerial foresight is rare in institutions with a comparatively narrow utilitarian perspective.

Appendix A.

The variogram is a powerful tool with which to characterize 1-D variations in all kinds of streams of matter (e.g. manufacturing in- and outflows, processing streams) that benefits from inherent auto-correlation between units (increments) sampled with different 'between pairs-of-samples' distances, termed 'lags'. *Variography* is in particular relevant for process sampling (or equivalently, stationary 1-D lot sampling), both instances referring to lot configurations for which one elongated dimension in time or space dominates completely, because the other two dimensions are eliminated by TOS' stipulation that all increments extracted must cover both these dimensions completely – hence TOS' rigid demand only to use *correct* increment delineation and extraction. TOS' approach to both stationary and dynamic 1-D sampling is able to address e.g. moving streams of matter on conveyer belts or in pipelines, units transported as truck loads, railroad cars or tanks vehicles, manufactured or produced units like containers, vessels, bags. Depending on their intrinsic heterogeneity characteristics, such streams are characterized by various degrees of 1-D auto-correlation as manifested by the variogram range and sill levels. A comprehensive description of sampling streams of e.g. irregularly distributed trace concentrations, including a thorough exposé of the versatility of variographic characterisation as a general approach for designing "fit-for-purpose" sampling plans *commensurate* with the empirical lot heterogeneity, can be found in Esbensen et al., 2012a, Minkkinen et al., 2012, and Esbensen et al., 2012b which deal with GMO batches originating as shipments into Europe from international ports of origin. While specifically addressing *feed*, this methodological treatment is rather a complete *exemplar*, which can be applied to all types of lots with *similar* 1-D lot heterogeneities; sampling of output streams from blenders is but one excellent example. There exist numerous, in-depth descriptions, illustrations and very many case histories involving variograms in the general TOS-literature; most of which are referred to in DS 3077 (Danish-Standards-Foundation, 2013) and Esbensen and Wagner (2014a,b).

References

- Allen, T., 2003. 1 – Powder sampling. In: Allen, T. (Ed.), *Powder Sampling and Particle Size Determination*. Elsevier, Amsterdam, pp. 1–55.
- Benedetti, C., Abatzoglou, N., Simard, J.S., McDermott, L., Léonard, G., Cartilier, L., 2007. Cohesive, multicomponent, dense powder flow characterization by NIR. *Int. J. Pharm.* 336, 292–301.
- Bergum, J.S., Prescott, J.K., Tekwani, R.W., Garcia, T.P., Clark, J., Brown, W., 2014. Current events in blend and content uniformity. *Pharm. Eng.* 34, 28–39.
- Bergum, J., Parks, T., Prescott, J., Tejwani, R., Clark, J., Brown, W., Muzzio, F., Patel, S., Hoiberg, C., 2015. Assessment of blend and content uniformity. technical discussion of sampling plans and application of ASTM E2709/E2810. *J. Pharm. Innov.* 10, 84–97.
- Berman, J., 2001. The compliance and science of blend uniformity analysis. *PDA J. Pharm. Sci. Technol.* 55, 209–222.
- Berman, J., Planchard, J.A., 1995. Blend uniformity and unit dose sampling. *Drug Dev. Ind. Pharm.* 21, 1257–1283.
- Boehm, G., 2001. Report on the industry blend uniformity practices survey. *Pharm. Technol.* 25, 20–26.

- Boehm, G., Clark, J., Dietrick, J., Foust, L., Garcia, T., Gavini, M., Gelber, L., Geoffroy, J.-M., Hoblitzell, J., Jimenez, P., Mergen, G., Muzzio, F., Planchard, J., Prescott, J., Timmermans, J., Takiar, N., 2003. The use of stratified sampling of blend and dosage units to demonstrate adequacy of mix for powder blends. *PDA J. Pharm. Sci. Technol.* 57, 64–74.
- Boehm, G., Clark, J., Dietrick, J., Foust, L., Garcia, T., Gavini, M., Geoffroy, J.-M., Jimenez, P., Mergen, G., Muzzio, F., Planchard, J., Prescott, J., Timmermans, J., Takiar, N., Whiteman, D., 2004. Results of statistical analysis of blend and dosage unit content uniformity data obtained from the product quality research institute blend uniformity working group data-mining effort. *PDA J. Pharm. Sci. Technol.* 58, 62–74.
- Chang, R.-K., Shukla, J., Buehler, J., 1996. An evaluation of a unit-dose compacting sample thief and a discussion of content uniformity testing and blending validation issues. *Drug Dev. Ind. Pharm.* 22, 1031–1035.
- Chang, R.-K., Guo, X., Lai, J.-W., 2004. The use of a rotary tablet press as a compacting spinning riffler for blending validation: the authors conducted an experiment involving phenylpropanolamine HCl as a model drug to demonstrate that a rotary tablet press can be used as a compacting spinning riffler. *Pharm. Technol.* 78.
- Colón, Y., Florian, M., Acevedo, D., Méndez, R., Romañach, R., 2014. Near infrared method development for a continuous manufacturing blending process. *J. Pharm. Innov.* 9, 291–301.
- Cullen, P.J., Romañach, R.J., Abatzoglou, N., Rielly, C.D., 2015. *Pharmaceutical Blending and Mixing*. John Wiley & Sons, Ltd.
- Dankwerts, P.V., 1952. The definition and measurement of some characteristics of mixtures. *Appl. Sci. Res. Sec. A Mech. Heat Chem. Eng. Math. Methods* 3, 279–296.
- Danish-Standards-Foundation, 2013. DS 3077(2013), Representative Sampling – Horizontal Standard. Danish Standard s Foundation, 1–42.
- de Silva, S.R., 1997. Mixing and Segregation in Industrial Processes. *Int. Fine Part. Res. Inst. SAR* 12–20.
- El-Hagrasy, A.S., D'Amico, F., Drennen, J.K., 2006. A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending. Part I: D-optimal design for characterization of powder mixing and preliminary spectral data evaluation. *J. Pharm. Sci.* 95, 392–406.
- Esbensen, K.H., Julius, L.P., 2009. Representative sampling data quality, validation – a necessary trinity in chemometrics. *Comp. Chemom. Chem. Biochem. Data Anal.* 1–4, C1–C20.
- Esbensen, K.H., Paasch-Mortensen, P., 2010. Process Sampling: Theory of Sampling – The Missing Link in Process Analytical Technologies (PAT), Process Analytical Technology. John Wiley & Sons, Ltd., pp. 37–80.
- Esbensen, K.H., Romañach, R.J., 2015. Proper Sampling, Total Measurement Uncertainty & Fit-for-Purpose Acceptance Levels for Pharmaceutical Mixing Monitoring. *Proceedings WCSB7* 25–30.
- Esbensen, K., Wagner, C., 2014a. Why we need a Theory of Sampling. *Anal. Sci.* 30–38.
- Esbensen, K.H., Wagner, C., 2014b. Theory of sampling (TOS) versus measurement uncertainty (MU) – a call for integration. *Trac-Trends Anal. Chem.* 57, 93–106.
- Esbensen, K.H., Friis-Petersen, H.H., Petersen, L., Holm-Nielsen, J.B., Mortensen, P.P., 2007. Representative process sampling—in practice: variographic analysis and estimation of total sampling errors (TSE). *Chemom. Intell. Lab. Syst.* 88, 41–59.
- Esbensen, K.H., Paoletti, C., Minkkinen, P., 2012a. Representative sampling of large kernel lots I. Theory of Sampling and variographic analysis. *Trac-Trends Anal. Chem.* 32, 154–164.
- Esbensen, K.H., Paoletti, C., Minkkinen, P., 2012b. Representative sampling of large kernel lots III. General considerations on sampling heterogeneous foods. *Trac-Trends Anal. Chem.* 32, 178–184.
- François-Bongarçon, D., Pitard, F., 2015. Comparison Between Samples with Constant Mass and Samples with Constant Fragment Population Size (and Calculations of Their Sampling Variances). *5. TOS Forum*, pp. 231–238.
- Garcia, T.P., Taylor, M.K., Pande, G.S., 1998. Comparison of the performance of two sample thieves for the determination of the content uniformity of a powder blend. *Pharm. Dev. Technol.* 3, 7–12.
- Garcia, T.P., Wilkinson, S.J., Scott, J.F., 2001. The development of a blend-sampling technique to assess the uniformity of a powder mixture. *Drug Dev. Ind. Pharm.* 27, 297–307.
- Garcia, T., Bergum, J., Prescott, J., Tejwani, R., Parks, T., Clark, J., Brown, W., Muzzio, F., Patel, S., Hoiberg, C., 2015. Recommendations for the assessment of blend and content uniformity: modifications to withdrawn FDA draft stratified sampling guidance. *J. Pharm. Innov.* 10, 76–83.
- Ghaderi, A., 2003. On characterization of continuous mixing of particulate materials. *Part. Sci. Technol.* 21, 271–282.
- Green, R.L., Thurau, G., Pixley, N.C., Mateos, A., Reed, R.A., Higgins, J.P., 2005. In-line monitoring of moisture content in fluid bed dryers using near-IR spectroscopy with consideration of sampling effects on method accuracy. *Anal. Chem.* 77, 4515–4522.
- Gy, P.M., 1981. A new theory of bed-blending derived from the theory of sampling – development and full-scale experimental check. *Int. J. Miner. Process.* 8, 201–238.
- Gy, P., 1998. *Sampling for Analytical Purposes*, first ed. Wiley, New York.
- Hwang, R.C., Gemoules, M.K., Ramløse, D.K., 1998. A systematic approach for optimizing the blending process of a direct-compression tablet formulation. *Pharm. Technol.* 22, 158–170.
- ICH, 2005. *Validation of Analytical Procedures: Text and Methodology Q2(R1)*. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 1–17.
- Jia, Z.Y., Davis, E., Muzzio, F.J., Ierapetritou, M.G., 2009. Predictive modeling for pharmaceutical processes using Kriging and response surface. *J. Pharm. Innov.* 4, 174–186.
- Kræmer, J., Svensson, J.R., Melgaard, H., 1999. Sampling bias in blending validation and a different approach to homogeneity assessment. *Drug Dev. Ind. Pharm.* 25, 217–222.
- Ma, H., Anderson, C.A., 2008. Characterization of pharmaceutical powder blends by NIR chemical imaging. *J. Pharm. Sci.* 97, 3305–3320.
- Martens, H., Naes, T., 1992. *Multivariate Calibration*. Wiley.
- Mateo-Ortiz, D., Colon, Y., Romanach, R.J., Mendez, R., 2014. Analysis of powder phenomena inside a Fette 3090 feed frame using in-line NIR spectroscopy. *J. Pharm. Biomed. Anal.* 100, 40–49.
- Matheron, G., 1966. Comparaison entre les échantillonnages à poids constant et à effectif constant. *Revue de L'industrie Minérale* 609–621.
- Minkkinen, P.O., Esbensen, K.H., 2009. Grab vs. composite sampling of particulate materials with significant spatial heterogeneity – A simulation study of correct sampling errors. *Anal. Chim. Acta* 653, 59–70.
- Minkkinen, P., Esbensen, K.H., Paoletti, C., 2012. Representative sampling of large kernel lots II. Application to soybean sampling for GMO control. *Trac-Trends Anal. Chem.* 32, 165–177.
- Muzzio, F.J., Robinson, P., Wightman, C., Brone, D., 1997. Sampling practices in powder blending. *Int. J. Pharm.* 155, 153–178.
- Muzzio, F.J., Roddy, M., Brone, D., Alexander, A.W., Sudah, O., 1999. An improved powder-sampling tool. *Pharm. Technol.* 23, 92–110.
- Muzzio, F.J., Goodridge, C.L., Alexander, A., Arratia, P., Yang, H., Sudah, O., Mergen, G., 2003. Sampling and characterization of pharmaceutical powders and granular blends. *Int. J. Pharm.* 250, 51–64.
- Osorio, J.G., Stuessy, G., Kemeny, G.J., Muzzio, F.J., 2014. Characterization of pharmaceutical powder blends using in situ near-infrared chemical imaging. *Chem. Eng. Sci.* 108, 244–257.
- Paoletti, C., Esbensen, K.H., 2015. Distributional assumptions in food and feed commodities-development of fit-for-purpose sampling protocols. *J. AOAC Int.* 98, 295–300.
- Pernenkil, L., Cooney, C.L., 2006. A review on the continuous blending of powders. *Chem. Eng. Sci.* 61, 720–742.
- Petersen, L., Esbensen, K.H., 2005. Representative process sampling for reliable data analysis—a tutorial. *J. Chemom.* 19, 625–647.
- Pitard, F.F., 1993. *Pierre Gy's Sampling Theory and Sampling Practice*. Heterogeneity, Sampling Correctness, and Statistical Process Control. CRC Press.
- Popo, M., Romero-Torres, S., Conde, C., Romanach, R.J., 2002. Blend uniformity analysis using stream sampling and near infrared spectroscopy. *AAPS PharmSciTech* 3, E24.
- Prescott, J.K., Garcia, T.P., 2001. A solid dosage and blend content uniformity troubleshooting diagram. *Pharm. Technol.* 25, 68–88.
- Roman Ospino, A., Zuñiga Ortega, C., Sanchez, A., Ortiz, S., Romañach, R.J., Esbensen, K.H., 2015. In: Esbensen, K.H., Wagner, C. (Eds.), *Estimating Total Sampling Error for Near Infrared Spectroscopic Analysis of Pharmaceutical Blends Theory of Sampling to the Rescue*, 5. WCSB7 IM Publications, TOS Forum, Bordeaux, France, pp. 71–75.
- Romañach, R.J., 2015. *Sampling and Determination of Adequacy of Mixing, Pharmaceutical Blending and Mixing*. John Wiley & Sons, Ltd., pp. 57–78.
- Sánchez-Paternina, A., Roman Ospino, A., Ortega Zuñiga, C., Alvarado, B., Esbensen, K.H., Romanach, R.J., 2015. When “homogeneity” is expected – Theory of Sampling in Pharmaceutical Manufacturing. *Bordeaux, France Proceedings WCSB7 IM Publications*, 5. TOS Forum, pp. 67–70.
- Singh, R., Sahay, A., Karry, K.M., Muzzio, F., Ierapetritou, M., Ramachandran, R., 2014a. Implementation of an advanced hybrid MPC-PID control system using PAT tools into a direct compaction continuous pharmaceutical tablet manufacturing pilot plant. *Int. J. Pharm.* 473, 38–54.
- Singh, R., Sahay, A., Muzzio, F., Ierapetritou, M., Ramachandran, R., 2014b. A systematic framework for onsite design and implementation of a control system in a continuous tablet manufacturing process. *Comput. Chem. Eng.* 66, 186–200.
- Susana, L., Canu, P., Santomaso, A.C., 2011. Development and characterization of a new thief sampling device for cohesive powders. *Int. J. Pharm.* 416, 260–267.
- Susarla, R., Sievens-Figueroa, L., Bhakay, A., Shen, Y.Y., Jerez-Rozo, J.L., Engen, W., Khusid, B., Bilgili, E., Romanach, R.J., Morris, K.R., Michniak-Kohn, B., Dave, R.N., 2013. Fast drying of biocompatible polymer films loaded with poorly water-soluble drug nano-particles via low temperature forced convection. *Int. J. Pharm.* 455, 93–103.
- Timmermans, J.H., 2001. A report of the PQRI workshop on blend uniformity. *Pharm. Technol.* 25, 76–84.
- U.S. Department of Health and Human Services, F.D.A., 2003. *Guidance for Industry Powder Blends and Finished Dosage Units-Stratified In-Process Dosage Unit Sampling and Assessment*, pp. 1–18.
- U.S. Department of Health and Human Services, F.D.A., 2002. *ANDA's: Blend Uniformity Analysis; Withdrawal of Draft Guidance*.
- U.S. Department of Health and Human Services, F.D.A., 2011. *Guidance for Industry Process Validation: General Principles and Practices. Current Good Manufacturing Practices (CGMP)*, 1–22.
- U.S. Department of Health and Human Services, F.D.A., 2015. *Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance – Production and Process Controls*.
- Wu, L.-S., Pang, J., Chen, J.-G., Hussain, M., 2000. Dry blending process scale-up for a very low dose drug candidate. *AAPS PharmSciTech* 1, 5–9.