The sampling interface—A critical Theory of Sampling success factor in process sampling and PAT

K.H. Esbensen[a] and V. Sivalingam[b]
[a]KHE Consulting, Copenhagen, Denmark. E-mail: khe.consult@gmail.com
[b] Dept. of Process, Technology, University of South-Eastern Norway, Porsgrunn, Norway

In the process industry arena, analytical endeavors are today well served by the Process Analytical Technology (PAT) framework, offering a plentitude of on-line analytics, mostly spectroscopic: UV-VIS, NIR, RAMAN, NMR, 'acoustic chemometrics' a.o. This context is generally considered satisfactory, but there is a major catch. The present paper focuses on a fundamental feature in process sampling, the *link* between an on-line analytical technique and a moving stream of matter, which is to be characterised, monitored, and analysed. A process sampling operation can serve many objectives; process regulation, product or interim product characterisation, optimisation of raw material consumption, maximizing energy efficiency, pollution, and environmental impact management – or process, plant, company, corporation decision making at supervisor and management levels concerning economics and risk management. All these objectives can only be meaningfully undertaken if based on *relevant information*, that is, if based on *reliable analytical data*, which in turn is 100% dependent upon *representative samples*, or *representative sensor signals*. The singular common element in all of this is the *process sampling interface*. This paper develops a critique of most current interfaces, with a scathing verdict: the dual role of the process sampling interface as both delineating an appropriate aliquot volume while facilitating the specific analytical technique, has largely been overlooked, with a significant negative effect that is spelled out in full detail across implemented solutions across widely different application sectors. Most current process sampling interfaces do not comply with TOS' demands, putting representativity severely at risk for heterogenous materials.

Introduction – setting the scene

The Theory of Sampling (TOS) has always been front and center concerning **how to** extract representative samples from moving lots (process sampling), e.g. from a conveyor belt or from ducted material streams. TOS' preeminence is that identical Governing Principles (GP), Sampling Unit Operations (SUO) and Sampling Error Management (SEM) can be applied identically to both stationary and to moving lots. The 70-year application history in TOS' traditional realms is highly successful regarding physical extraction of increments and samples, especially in the Mining, Minerals and Metals processing and refining industry (M3), in cement production and in many other process industry sectors dealing with broken ores, particulate matter, minerals and slurries. A comprehensive, up-to-date introduction to TOS in these sectors can be found in three recent TOS textbooks, and in the wealth of references to the extensive literature found herein[1-3]. These are conveniently presented in a comparative overview, easily available[4].

Recently in these industry sectors, a pronounced interest can be observed for applying modern *sensor technologies* for on-line monitoring because of the obvious advantages for fast, real-time information used for process monitoring and control allowing optimized decision making, all with significantly added economic results. These approaches take a broader view on available analytical techniques that can be applied on-line, always closely related to the characteristics of the materials and lots involved (physical, chemical, other). A key example outside TOS' traditional realm is the use of Near InfraRed (NIR) spectroscopy in the broad pharmaceutical manufacturing context ('pharma' for short). In terms of societal importance and total turnaround economy, one can meaningfully equate M3 and pharma.

In the process industry arena today, analytical endeavors are well served by the Process Analytical Technology (PAT) framework, offering a plentitude of on-line mostly spectroscopic analytics: UV-VIS, NIR, RAMAN, NMR, 'acoustic chemometrics' a.o. A comprehensive introduction to this field can be found in the "PAT bible (2010)"[5].

The present paper focuses on the fundamental feature in common in process sampling, the *link* between an on-line analytical technique and the moving stream of matter, which is to be characterized, monitored, and analysed. A process sampling operation can serve many objectives, process regulation, product or interim product characterization, optimization of raw material consumption, maximizing energy efficiency, pollution, and environmental impact management – or process, plant, company, corporation decision making at supervisor and management levels concerning economics and risk management. All these objectives can only be meaningfully undertaken if based on *relevant information*, that is if based on *reliable analytical data*, which in turn is 100% dependent upon *representative samples*, or *representative sensor signals* from the start.

doi: https://doi.org/10.1255/tosf.141

The singular common element in all of this is – the *process sampling interface*.

Below it is shown that this commonality is the critical success factor w.r.t. representative sampling *sensu* TOS. To reveal the critical role of the process sampling interface, examples shall be presented from TOS' traditional realm (not many are needed taking into account the WCSB10 context), but deliberately also from what at first sight may appear as very different industrial sectors, pharma and *similar* industries (food/ feed processing, production and manufacturing for example), but which are never-the-less dealing with the same matter, i.e. heterogeneous aggregate mixtures, only often from a finer grain-size realm (powders, slurries). From these industrial sectors focus shall be on NIR spectroscopic applications, which are all in the realm of PAT.

While at first sight the M3 and pharma/food/feed (PFF) industry sectors are likely to be found disparate in the extreme in their process technology contexts and w.r.t. the wildly different materials involved, it turns out they are closely related in their *identical need* for representative process sampling to conduct reliable process monitoring and process control, and QC/QA of interim/finished product. In fact, for PFF the need for appropriate sampling Quality Management (QM) is *identical* to those for M3. This means that M3 and PFF industries are fully aligned and joined precisely at the juncture which is the focus for the present paper – the *process sampling interface.* The common *role* of the process sampling interface is to bring a specific on-line analytical technique in appropriate contact with the material/product/streaming material that is to be characterized. The role of the process sampling interface is to facilitate the analytic *interaction* with the moving material – without giving rise to sampling errors to any degree more than absolute necessary. This turns out to be a formidable challenge.

In this paper, the analytical specifics in themselves are not of interest (see the abundant PAT literature). Below it is all, and only about what comes *before* analysis. Also, the analytical competence *per se* is very rarely an issue; it is almost universally well taken care of in the form of thoroughly validated analytical procedures.

So, are the M3, PFF, and all similar particulate matter realms fully competent regarding the specific process sampling issues? After all, solutions to this task can look at the unparallel success for historical TOS applications over ~70 years in science, technology, industry, trade, society and in the environmental sciences. For more than five decades the Theory of Sampling (TOS) is rightly famous as the framework that covers all necessary elements, i.e, design of sampling procedures and equipment, correct implementation, correct practical operation, sufficient maintenance, a necessary and sufficient theoretical competence framework.

Well, it turns out there are still aspects of TOS application that can benefit from a general perspective - and none more so than the *process sampling interface*.

TOS vs. PAT – an exemplar contrast from which to learn

Within TOS' traditional realms it is well known **how to** sample particulate matter and mixtures, both from stationary lots and from moving streams of matter. When it comes to **how to** extract physical increments and samples in a representative fashion, the foundation of TOS reigns supreme, and all new applications always take their point of departure from this framework and the from the wealth of accumulated practical experience, as witnessed by the historical library of Proceedings from the World Conferences on Sampling and Blending WCSB1-WCSB10. Sampling by physical extraction is very well served indeed, see[6-9] for example.

By contrast, for 'sensor sampling' i.e., Process Analytical Technologies (PAT), which is *claimed* to be a 'no sampling needed' approach, there is no similar theoretical framework for the sampling step, which is rather often treated in a somewhat stepmotherly fashion. Instead, a pronounced practical approach rules in this realm[5], in which the question of **how to** achieve representative sensor signals is the first item on any agenda, here intimately related to the design and implementation of the specific analytical instrumentation, which is to be put in operation by way of a *sensor interface* connecting to, and interacting with, the streaming flux of matter.

This endeavor is closely related to the specific analytical technique involved (the use of NIR is dominant, but other analytical techniques find their use as well in specific cases). This is always based on specific *calibration* approaches; calibration is always needed as PAT instrumentation is multi-channel, multi-wavelength – producing *multivariate spectral data*. There can be observed a drive: "get the multivariate spectral signals – and all will be well", which is based on the overwhelming success of the data analytical discipline *chemometrics* in the form of *multivariate calibration* of process sensor signals from multi-channel analytical instruments[10]. One often gets the distinct impression that the specifics of a PAT sensor interface is more driven by the needs and requirements of the sensor analytics than of considerations for the analytical support (volume/mass) that can be realized. This is where the issue of the effective *increment volume* crops up.

Preluding what is more fully illustrated in systematic graphics below, PAT solutions make use of three principal types of interfaces, three sampling *modi*, which are very well known in TOS. Pat interfaces are almost always of the type (1,2) *instead* of (3):

1. Taking part of the stream, all of the time; *or*
2. Taking part of the stream, part of the time; *instead of*
3. Taking full cross-steam increments part of the time (the only TOS-compliant sampling approach)

These sampling modi have been illustrated with particularly clear graphics in a companion contribution in these proceedings, see Figs. 2-4 in Novaes & Hidding (2022)[11].

With modus (3), focus rightly is on sampling the entire stream, *full slices* of the stream, while sampling *in process time.* While with (1,2), the full width/breadth/depth of the stream is never covered in full, leading to TOS-incorrect sampling, massive IDE, thereby breaking one the two necessary-and-sufficient requirements for representative process sampling (*unbiasedness* and *sufficient* precision).

Based on such a failing PAT focus, the issue of representativity here is mostly, or sometimes wholly, only related to spectral and reference sample measurement uncertainty (MU) – but data modeling errors and uncertainties are also acknowledged if the chemometrics is not up to the job (data model under – or over fitting). With this PAT focus, the realm 'before sensor signal acquisition' is often unwittingly ignored, even though this is the very domain where *sampling* errors abound. This issue constitutes the central theme of this presentation: In the PAT realm sampling errors are incurred in the exact *same fashion* as when extracting physical samples from moving streams. In the PAT realm this is largely a surprising insight – indeed this parallel is also able to shed light on how process sampling sometimes is subpar, even in the TOS arena.

For these reasons, sampling error effects impacting on the quality of sensor and probe signals largely constitute a *terra incognito* in the PAT realm. This situation invites all parties to take a serious look at the need to comply with TOS' stipulations for representativity, no matter with what interface configuration sampling is performed. It turns out that there is a complete *duality* between TOS and PAT, which must be acknowledged and followed where- and whenever the goal is representative sensor signals, see chapter three in[5], and also[10,11]. This duality is the central theme for this paper.

Be aware that a dear friend goes under many names. Thus, what in chemometrics is termed '*multivariate'* (multi-variate, meaning *many variables*), in many other scientific and technological fields has acquired the name '*hyper-spectral*' a.o. A good example of the latter, and a good example of the emerging awareness of the spectral (multivariate, hyper-spectral) advantage by frontline players, would be Mittrup et al (2017)[12], which is also among the pioneers applying NIR spectroscopy in M3.

The contradiction between TOS and PAT has been highly frustrating, indeed a situation that has lasted for the better part of the last 10-15 years. The scope can best be illustrated by the fact that the first edition of the famed source book: "Process Analytical Technologies (PAT)" (2007) did not contain a specific chapter on sampling (*sic*). While the second edition, out just three years later (2010)[5], features a comprehensive chapter, aptly named: "Process sampling: Theory of Sampling (TOS) – the missing link in Process Analytical Technologies (PAT)" – the title says it all!

Focusing

In many industry sectors, there is today an emerging rush to take advantage of 'modern sensor technology' (including 'advanced' statistics, machine learning and other artificial intelligence approaches), to provide real-time process information to support and optimise business decision making. Although this trend manifests itself slightly differently in different process industry sectors, typically because the sector viewpoint is intimately related to one dominating specific analytical approach, e.g., XRF, PGNAA, NIR, it is rare to find awareness of the overarching *hidden* commonality involved, the role of process sampling interface. Focusing on this issue makes it possible to review and critique a multitude of current applications to avoid making the same mistakes over and over in many disparate realms in science, technology, industry, trade, society, and the environment.

In medias res - PAT

Process monitoring and control in technology and industry is *incomplete* without full understanding of all sources of variation, causing bias and impression. It is particularly important to be able to *decouple* sampling and analytical variations caused by bias-generating errors because these can, and shall be reduced maximally, indeed preferentially eliminated. This is the first, well-known mandate of TOS: elimination of sampling bias. By contrast, the analytical bias is fundamentally different, not always clearly acknowledged as is a major point in[1]. The analytical bias is treated very well within the specific analytical realm and literature.

A case in point: It is not enough to make use of Multivariate Statistical Process Control (MSPC) or similar QCQAQM monitoring approaches based on charted analytical results, because these may well be compromised by *unknown* significant sampling errors, if not adequately identified, quantified, and reduced to below a relevant *a priori* acceptance threshold: This situation results in unnecessarily large bias and impression.

More precisely: Process data are affected both by analytical measurement errors as well as extraction sampling errors or sensor signal acquisition errors. This understanding is of critical importance, because the latter two categories of sampling error effects are all too often unrecognized, or unknown. But worse, they typically dominate over analytical errors by factors 5-10-20+ the more heterogeneous materials targeted and/or if proper sampling competence is not brought to bear in the design, implementation, maintenance, and operation of the *total* process measurement system. Alarmingly, almost always the Total Analytical Errors (TAE) constitute only a very small, sometimes vanishingly small proportion of the total Measurement Uncertainty (MU) – What makes "all the difference" quantitatively and qualitatively, is the Total Sampling Error (TSE). Below we are only interested in **how to** minimize TSE in process sampling.

In the PAT and in the process technology context in general, it is not sufficiently known that sensor signal acquisition gives rise to *identical error types* as does physical sample extraction.

Below we bring forth the critical analogy between Process Analytical Technologies (PAT) and conventional physical sample extraction and show how three QCQAQM tools, the Replication Experiment (RE), interleaved sampling (IS), and variographic process characterisation (VC) can be involved. RE, IS, VC constitute a necessary and sufficient on-line approach for total variability decomposition. Without proper sampling error management (error *identification*, *reduction* and *elimination*), on-line analytical instrument calibration and validation will unavoidably incorporate significantly inflated data uncertainties, the bias complement of which is inconstant and without correction possibilities, all of which assuredly compromises the ultimate process monitoring and control objectives significantly.

The presentation below reflects the close interaction between TOS, PAT and chemometrics, a competence overlapping that is not much in focus in the IPGSA realm. Here, space consideration does not allow more than referral to a select suite of key, comprehensive literature sources for interested parties who wish to go deeper with respect to fundamental theory with all three scientific fields1[1,5-10].

A practical starting platform for the present objective shall be a fundamental sampling *duality*.

The fundamental duality between physical and sensor sampling

Below all efforts have been made to make illustrations completely anonymous; no identification of persons, locations, instrument brand or companies is intended. Only the didactic value of correct vs. incorrect, faulty, or fatal non-compliance with TOS is of interest.

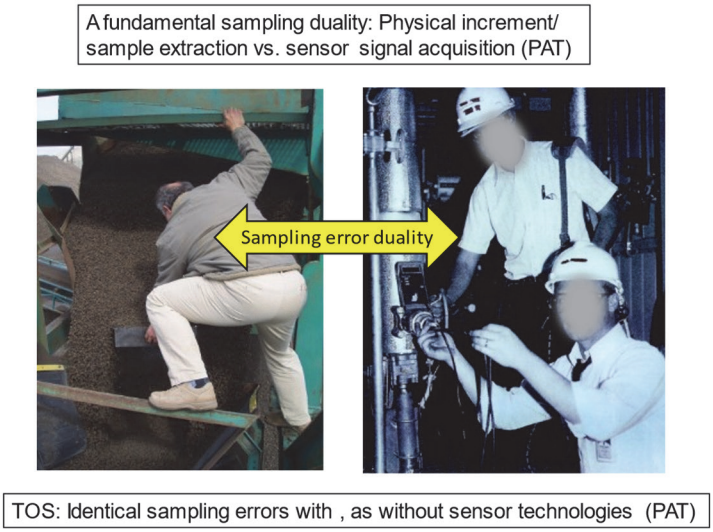

Fig. 1. The fundamental sampling duality between physical increment/sample extraction (left) and sensor signal acquisition (right). Installing PAT sensors has often been claimed to be equivalent to 'no sampling needed' – alas! Illustration copyright KHE Consulting teaching collection, used with permission.

Installing PAT sensors has often been claimed to be equivalent to 'no sampling needed' – alas, this is a complete fallacy. The issues treated below are not novelties for the sampling community – rather they should be trivial but are not entirely. But the main thrust of this paper is directed at related scientific communities, such as chemometrics, process technology, process engineering, processing, and manufacturing plant design – all facing the same issues. For these communities, it has been a rude awakening how the "Eureka moment" in the right panel of Fig. 1. dissolved into a whimper as soon as proper Total Measurement Uncertainty and full system performance validation was invoked. Alas, TSE error effects also abound in the PAT sensor world. "Why - there is literally no physical sampling involved?"

First lesson: Both physical extraction sampling errors and sensor technology spectral acquisition errors are in play – in fact they are identical, as even rudimentary TOS-analysis will show.

The sampling interface

Figure 2 shows the key insight presented in this paper: TOS' mandatory rules that guarantee representative increment extraction from a moving lot (physical process sampling) are *identical* when applied to secure a representative sensor signal from a ducted material stream (sensor or PAT process sampling). There is an often-overlooked subtlety involved here, which is the distinction between a sample container, holder, cup, vial – and a *sampling* interface, see further below.

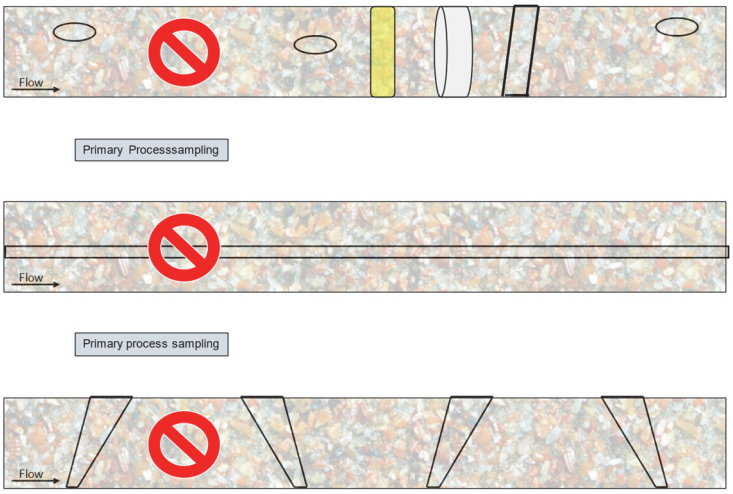

Fig. 2. Key commonality between physical increment/sample volume and sensor signal support in process sampling: Identical rules apply w.r.t. correct increment delineation and extraction (complete and undisturbed volumetric coverage) for both physical extraction as well as for delineation of the sensor signal support volume as facilitated by a sampling interface. This illustration can either be viewed as looking down on a heterogeneous 1-D moving material lot f.ex. on a conveyor belt - or it can be understood as a vertical section through a horisontal conduit transporting a heterogeneous material (indicated by irregular texture). From both points of view, TOS-correct process sampling necessitates recovery of, or spectral coverage of, the full delineated cross-stream increment volume, shown by the three renditions in the upper panel Only fully covering increment volumes with planar-parallel boundaries will result in representative increments or samples. Illustration copyright KHE Consulting teaching collection, used with permission.

Note the non-representative increment geometries indicated in Fig. 2. The upper panel shows fatal 'grab sampling', either as partial 3-D volumes or as superficial grabbed increments from the top of conveyed material only. The central panel illustrates increment delineation/extraction f.ex. resulting from a valve opening capturing a "partial fraction of the moving stream of matter all the time" only (as well exemplified by the infamous Shark Fin valve), while the bottom panel shows flawed cross-stream increment delineations, flawed as they furnish unbalanced increments w.r.t. transverse stream heterogeneity. None of these configurations corresponds to the imperative demand for an uncompromised full slice of the moving stream, none of these deliver TOS-compliant increments.

Thus, Fig.2 shows the non-negotiable demands for correct vs. incorrect increment delineation, which determines whether a sampling process will give rise to Increment Delineation Error (IDE) effects, or not. By implication, the illustration also shows the demands for elimination of Increment Extraction Errors (IEE). These demands on a practical physical sampling system, for example from a conveyor belt, applies with equal criticality for *sensor sampling*, but contrary to what is often *assumed*, there is not a dichotomous divide between physical – and sensor sampling. Fig. 3 shows the most often occurring principal process sampling options of either modality; it is noteworthy that all the shown examples from current process technology do not deliver representativity[1-5].

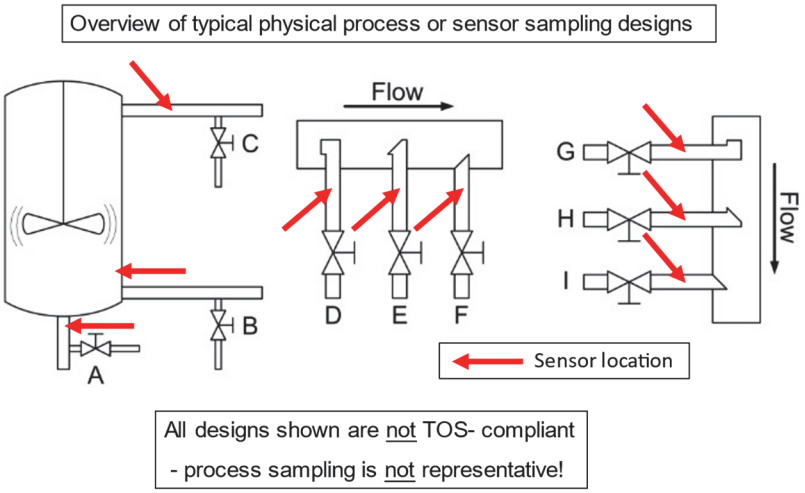

Fig.3. Overview of typical, often occurring sampling system configurations and locations encountered in process technology across diverse industry sectors. "Sampling valves" are designated A through I, while typical locations of sensor technology solutions (PAT) are marked as red arrows. None of the process system solutions shown here are representative however, see[5]. Illustration copyright KHE Consulting teaching collection, used with permission.

Whether designed for physical increment/sample extraction by use of "sample valves" [A through I] from vessels and containers, horisontal or vertical pipelines – or designed to take advantage of the much claimed "no sampling necessary" sensor technology (PAT), *none* of the sampling systems depicted in Fig. 3 will give rise to representative samples or signals, which is easily verified by reference to the 'full slice imperative' shown in Fig. 2.

Figures 2 & 3 and the remarks above, lend full justification to the *duality* argument illustrated in Fig 1.

There is another duality presented below. Illustrative examples of the main message in this paper shall come, apparently willy-nilly, from both the pharma PAT/NIR world and from the traditional TOS process realm addressing M3 for example. This is deliberate, as it is meant to bring forth the universality of the discussion of the process sampling interface that follows.

But first a few reflections on sensor spectroscopic penetration depth, which are of key interest.

Spectral penetration depth also matters

NIR analysis is influenced by how lots, samples, and smaller sub-samples are handled *before* the aliquot is presented to the analyser. An iconic author in the NIR field (P. Williams, 2001) listed ~30 sample-associated factors that affect the accuracy and precision of NIR analysis, but only gave indicative actions for how to avoid *some* of these and emphasized that: "*If [these] are not efficiently carried out, consistent and accurate NIR analysis is not possible, no matter how sophisticated the software.*", (Williams quoted in[13]). We might here add: "… and regardless of the analytical competence". This view is in complete agreement with the principles presented by TOS.

Thus, with respect to analytical errors and their effect (NIR and for almost any other analytical modality), in a specific sense it is all about what comes *before* analysis – because this is where the quality, relevance and validity of the subsequent analytical results originate! Disregard for this realm is like the original sin …

One specific analytical issue can also be seen as impacting the sampling realm, which is of paramount importance for sensor process sampling. This concerns the effective spectroscopic *penetration depth*. We shall use NIR as an *exemplar* for many other spectroscopic techniques, although their individual penetration depths will vary, and can be widely different. However, the principal argument below is universal: penetration depth is a *bona fide* sampling issue, penetration depth uncertainty matters!

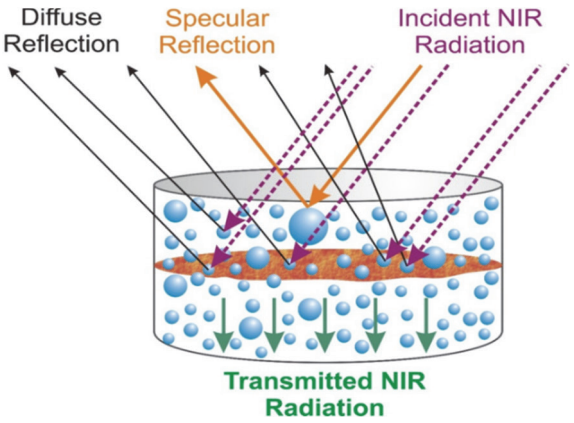

Fig. 4. Schematics for NIR spectroscopy/spectrometry, illustrating the thorny issue of varying penetration depth. Illustration copyright, Rodolfo Romanach teaching collection, used with permission. See also discussion in[13]

The analytical setup in Fig. 4 illustrates the business end of sensor sampling of a powder mixture, based on the interaction of the incident radiation with the material, either in an extracted sample contained in a sample holder (full cylinder height) or in an assumed, correctly delineated volume serving the same purpose as the sample container. This volume is delineated by the sensor sampling interface.

The key issues is that this interaction is simultaneously a sampling operation (sensor sampling), as well as the physical basis for quantitative analysis through a multivariate calibration model, allowing prediction of a specific material property (often a concentration) (Esbensen and Swarbrick 2018)[10], Romañach and Esbensen 2016)[13]. However, this type of chemometric prediction of analytical concentration will be fraught with unnecessarily inflated prediction errors (prediction variances), with varying material-dependent impacts, due an obvious volume mismatch between the *assumed* aliquot volume (full cylinder height) and the physically realized support volume for the spectral signals, which is only a fraction hereof.

In this case, sensor sampling is intricately related to the physics of diffuse reflection in NIR spectroscopy. The volume mismatch issue is in general material dependent. As the top layers of the sample will receive more radiation than the bottom layers, less radiation reaches the bottom part of a material on a volume-to-volume comparison. For a given material with a specific physical constitution and different grain-size distribution a.o., the depth penetration also depends on the wavelength used. The lower wavelength radiation penetrates to greater depths at where molecules absorb less NIR radiations. Thus, the *effective sample mass* interacting with NIR radiation also *varies* according with wavelength. A more comprehensive introductory description of these issues can be found in[13] and in "Sampling" column in *Spectroscopy Europe/World* SE-2 (2021)[14], entitled "What are Sampling Errors – and WHAT can be done about them", which has this ingress:

> "The objective of this column is to provide easy-to-understand examples of sampling errors. Prompted by recent participations and presentations at on-line conferences and meetings, we believe there is a need for a more fulfilling introduction and exemplification of the concept and real-world consequences of committing "sampling errors". WHAT is a sampling error? WHAT is the result of sampling errors? WHAT can we do about sampling errors? […] The point of departure will be in the Theory of Sampling (TOS) and in the near infrared spectroscopy analysis and pharma application sectors, but the focus will be developed to be more *general*, so that readers can carry-over to other scientific and application areas of interest."

The vast professional NIR literature is readily available (a plethora of references can be found in the literature referred to above).

The key issue is that the effective volumes are not identical – which according to TOS' framework will unavoidably lead to a bias.

It is sometimes argued that since this mismatch error is 'identical' for all samples (calibration, validation, prediction samples), it will in effect create a 'constant bias', which will influence calibration, validation, and prediction samples identically, that can therefore be corrected for (as can an analytical bias). However, this is a fatal *misunderstanding* because this understanding tacitly *presupposed* that all materials, mixtures, and powders are *identical* when involved in NIR analysis, all differences in sample preparation, handling, pouring, compactification, and presentation a.o. notwithstanding. But this is clearly never the case for the multitude of very different materials subjected to NIR analysis in this world.

These relationships cannot be resolved universally by the traditional route of demanding a "bigger sample". The solution is hoped to be NIR radiation interacting with an increased analytical mass in a well-designed fashion to accomplish practical *composite sampling* and at the same time guarantee sampling correctness (elimination of sampling bias). This objective has been the holy grail for decades of development of suitable 'sample cells' within the NIR realm, far from all of which having reached this goal, however. While the specifical analytical problems of this kind are well recognised, it takes more

than a minimum of TOS competence to look through the immediate manifestations of many 'sample interfaces', salient examples appear below.

This kind of sensor sampling error cannot be corrected for by any physical pre- or post-treatment, any mathematical algorithm or by any software approach. These issues will always affect adversely on the possibility of developing fully optimised calibration models. All NIR spectroscopic methods are similar to a multiple path-length sample cell where particles may interact with radiation more times than once. The mass analysed by the radiation may be estimated in some cases but is generally not known in detail. Most other analytical chemists know exactly the material that is analysed, including its mass. The authoritative "HANDBOOK of NIR Analysis (2020)", chapter 19, characterises this fundamental sensor sampling enigma in the following way (*quote*): "*NIR spectroscopists are spectacularly the most successful analytical chemists that do not know the exact sample mass being analysed!*" Factor in the appropriate density of the analysed material, then mass =~volume in this scathing dictum.

For multivariate calibration truly to be the chemometric be-all, end-all solution to process sampling, it is imperative that 'an increment – is an increment' - in the sense that 'a process sampling interface increment volume – is simultaneously also the physical reference sample volume *of matching size*'. This is obviously not something that can be taken for granted across the gamut of analytical techniques in the realm of PAT.

The process sampling interface is a harsh mistress at this most fundamental scale. It pays to pay attention, also because other issues occur at the scale of the process sampling interface itself.

Spectral averaging is not composite sampling

Development of PAT systems to meet TOS' Fundamental Sampling Principle (FSP) is an ongoing challenge. Full compliance with FSP requires that an entire cross section of a process stream shall be obtainable for it to be a *bona fide* representative increment, Fig. 2. This requirement is probably the most often occurring sin-by-omission, or sin-by-neglect, in many current PAT and in similar process technological applications.

As but one example, the Field-of-View (FOV) of a beam of NIR irradiation only interacts with, say, the top 1-2 mm or so of a moving bed of matter. For a process stream the depth of which may be many, many times greater; there is clearly a massive Increment Delimitation Error (IDE) at work here. This situation must rightly be considered *beam grab sampling*, incurring a significant support volume mismatch. The parallel with physical grab sampling is striking – the duality is complete.

Today's dominating modus operandum is pledging allegiance to a massive averaging of spectral scans acquired during the process flow. While this in some, restricted way appears to constitute an improved procedure, the penetrative question is, how much better is a number of grab samples, scanned or physically obtained grabs, if none of them complies with the necessary principles for increment representativity? In the informed optics of TOS, this 'way out' for PAT analysts will unfortunately not necessarily lead to representative averaged signals even though *apparently* using spectral composite sampling - because it is still only the uppermost part of the streaming matter which is 'averaged'.

A representative composite sample (physical or optical) requires that all constituent parts of the lot moving through the sampling interface, have the *opportunity* to become part of the composite sample. This can manifestly not be obtained by any variation of increments which only represent the top of a moving stream, or any similar series of *partial* increment volumes only, Fig. 2. This is a structural impossibility. Spectral averaging is not by itself TOS-compliant composite sampling, thereby affecting TSE – but spectral averaging does help to get the analytical uncertainty better constrained, thereby affecting TAE (only). This all takes place under the complete boundary condition of the Total Measurement System Error TMSE := TSE + TAE.

From PAT/NIR sensor sampling to physical sample extraction

What is the relationship between the pharma/NIR/PAT powder and powder mixture worlds briefly described above – and the traditional TOS realm of sampling of aggregated particulate materials of geological, biological a.o. origin over a very wide range of grain-sizes, say from decimeters to micrometers? Despite all the drastic intrinsic differences, from the point of view of process sampling *it is all the same* - only the scale changes.

Key distinction

Perusing the relevant literature, one often comes across a misunderstanding concerning the distinction between a "sample cell" and a "sampling cell". The difference is very much not just a semantic difference – it constitutes the essential insight for full understanding of the role of the process sampling interface.

A "sample cell" (noun) (think of a sample holder, a sample cup, a sample vial, a 'petri dish' … …) *presupposes* that 'the sample' has been sampled before being put into this container. All the thorny issues regarding sampling have been left in the analytical ante-room, the realm 'before analysis'. For good order it is noted that this realm is typically comprised by

an integral series of sampling stages: primary -, secondary, tertiary sampling ... the last of which produces the analytical sample. The primary characteristic of a sample cell is that it is *passive*; it 'merely' is filled by the already-sampled analytical aliquot. "Someone else is responsible for the sampling" – heard this before?

A "sampling cell" (noun) is an *active* devise that serves a dual purpose (think of a 'flow through' cell). A *sampling* cell is designed to delineate the operative volume support (sensor increment) from which spectroscopic signals is obtained, thus effectively determining the analytical sample mass - while simultaneously facilitating the spectroscopic interaction with the moving material (sensor field-of-view, analytical duct window opening a.o.). The former function has everything to do with TOS' concepts of IDE and IEE. A broader view is needed than the basic truism: all sampling cells enclose a specific volume, determined by the design of the sampling interface. Here is the most important distinction:

The dynamic analytical support volume delineated by the process sampling interface – must be equal to the volume of a 'full slice' increment, in the sense of Fig. 2. Indeed, physically it must be the *same volume.* No exceptions are acceptable, if representativity is the objective.

Graphics – a picture tell a thousand words

The following figures are presented as principal sketches and generic renditions; they only serve the purpose of illustrating the characteristics of an installed process sampling interface. No identification of instrument brand, OEM, plant, company, corporation, or persons is intended. On the contrary every effort has been made only to present schematic didactic information pertaining to **how to** perform proper *process* sampling.

Figures 5-13 below all focus on the role of the sampling interface. The examples shown cover a wide and dominating range of contemporary process technological implementations and PAT installations.

It is highly significant that nearly all the illustrations are telling the story of **how not to** perform process sampling! The reason is sobering – there are not many practical process sampling equipment and system solutions found on the market today, or in contemporary industrial installations, that factually comply with TOS' requirements for representative process sampling. Unless the gamut of all the world's materials correspond to infinitely diluted chemical solutions or to perfectly mixed composite materials. Not many do (see also the compositional caveat below). For all of these, the following will be a dramatic eye-opener,

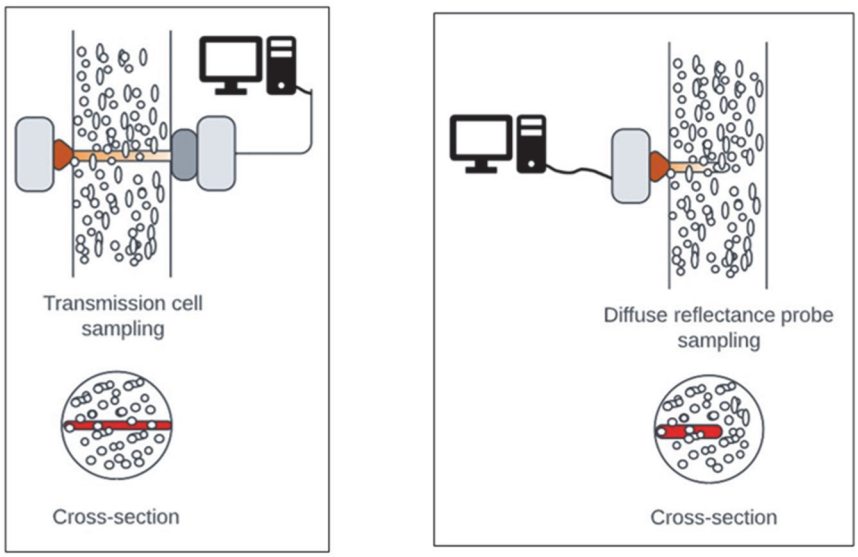

Fig. 5. Principal sketch of process sampling interfaces for analytical techniques purporting to deliver sensor signals that are representative of the streaming matter in conveyed flows, ducted pipes or from reactors. The flow intersecting volumes, depicted in red in cross-section views of the flow, do not comply with the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Illustration copyright KHE Consulting teaching collection, used with permission.

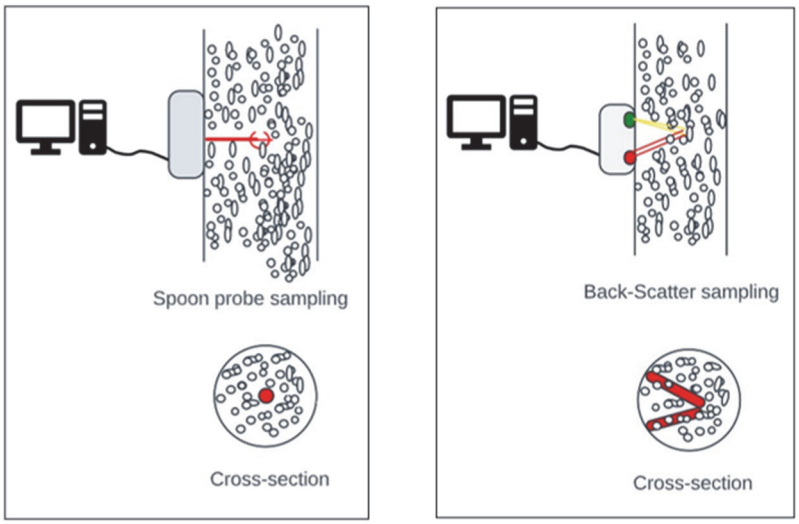

Fig. 6. Principal sketch of process sampling interfaces for analytical techniques purporting to deliver sensor signals that are representative of the streaming matter in conveyed flows, ducted pipes or from reactors. The flow intersecting volumes, depicted in red in cross-section views of the flow, do not comply with the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Illustration copyright KHE Consulting teaching collection, used with permission.

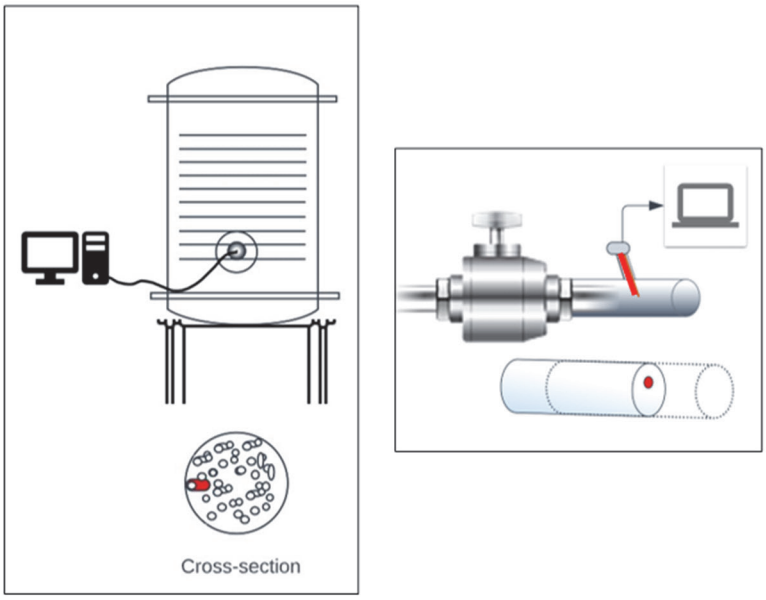

Fig. 7. Principal sketch of process sampling interfaces for analytical techniques purporting to deliver sensor signals that are representative of the streaming matter in conveyed flows, ducted pipes - or from reactors. The flow-intersecting volumes, depicted in red in cross-section views, do not correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Illustration copyright KHE Consulting teaching collection, used with permission.

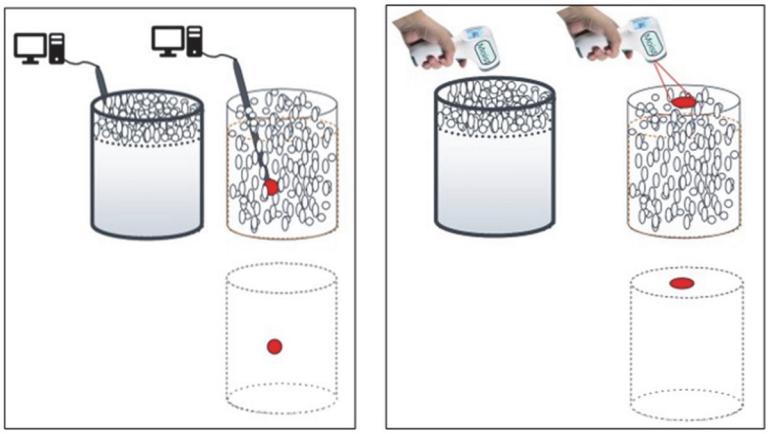

Fig. 8. Principal sketch of process sampling interfaces for analytical techniques purporting to deliver sensor signals that are representative of the lot material residing in storage containers and similar. The realized analytical support volumes, depicted in red in cross-section views, do not correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. A series of containers or drums, for example from a production line or forming an incoming raw product line, also constitute a process sampling situation, whether sampled by an installed sensor (left) or a handheld device (right). Illustration copyright KHE Consulting teaching collection, used with permission

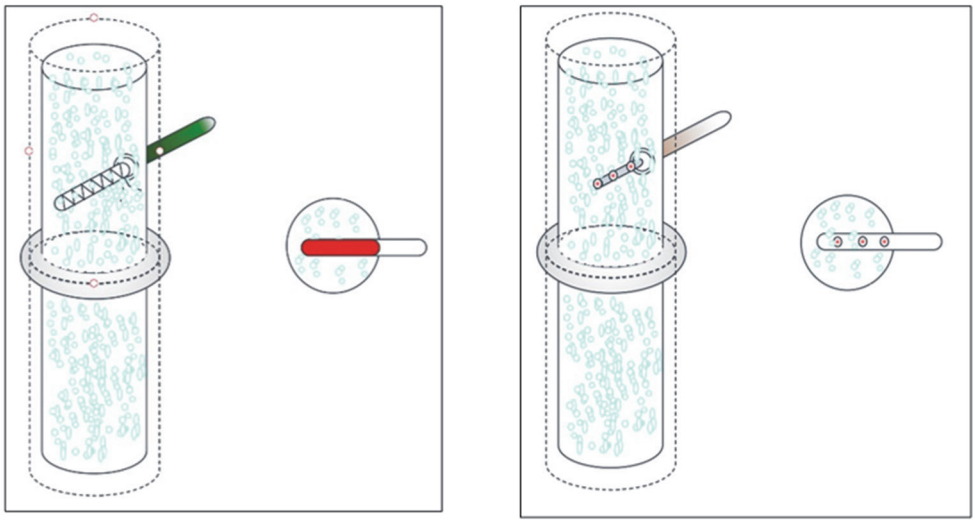

Fig. 9. Principal sketch of process sampling interfaces purporting to deliver physical increments that are representative of the streaming matter in conveyed flows, ducted pipes or from reactors. The realized analytical support volumes, depicted in red in cross-section views of the flow, do not correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Inserting probes and screw samplers makes no difference, despite many claims to the contrary in the marketplace. The deliverable is a far cry away from a correct, full slice of the stream, cfr. Fig. 2. Illustration copyright KHE Consulting teaching collection, used with permission.

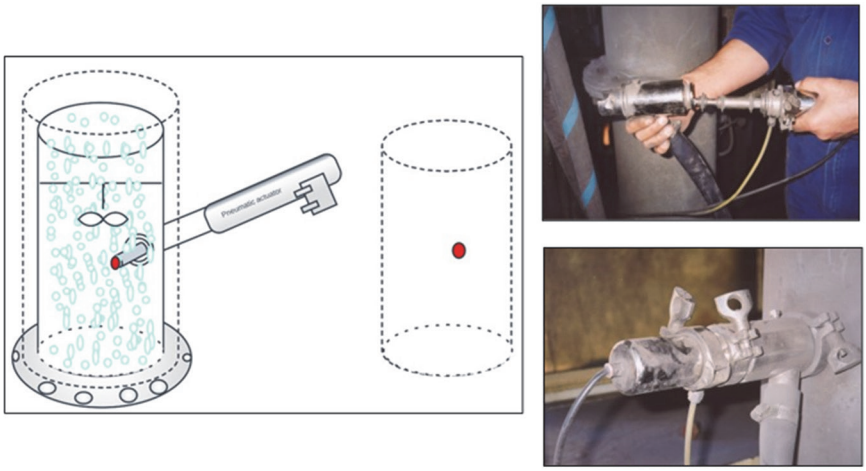

Fig. 10. Principal sketch of process sampling interface purporting to deliver physical increments that are representative of the streaming matter in conveyed flows, ducted pipes or from reactors. The realized analytical support volume, depicted in red in cross-section views of the flow, do not correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Process control samplers, including pressure pipe and poppet samplers, make no difference, despite many claims to the contrary in the marketplace. Illustration copyright KHE Consulting teaching collection, used with permission.

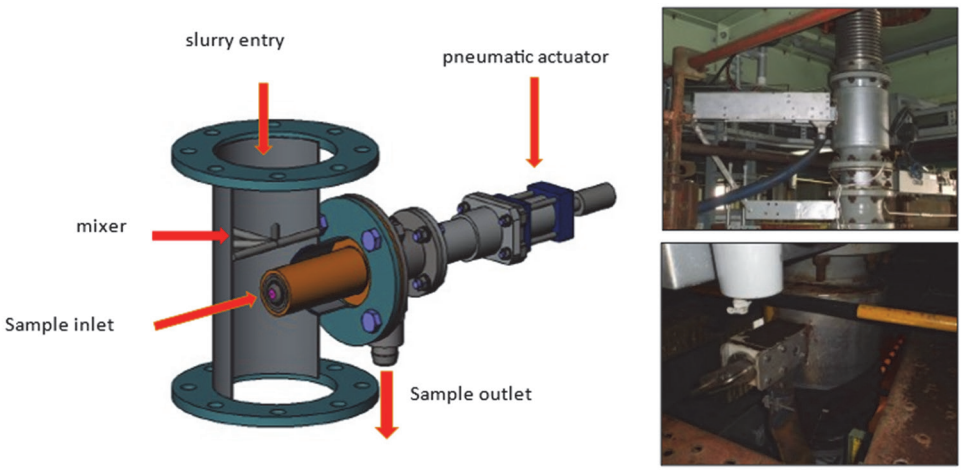

Fig. 11. Principal sketch of process sampling interfaces purporting to deliver physical increments that are representative of the streaming matter in conveyed flows, ducted pipes or from reactors. The realized analytical support volume ("sample inlet") does not correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Poppet samplers makes no difference, despite many claims to the contrary in the marketplace. Illustration copyright KHE Consulting teaching collection, used with permission.

Fig. 12. Principal sketch of process sampling interface purporting to deliver physical increments that are representative of the streaming matter in conveyed flows and ducted pipes. The realized analytical support volume, here depicted in green in cross-section view of the flow, do not correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. This illustration shows the famous 'shark fin' sampler. It fails to extract a complete slice of the moving matter but delivers only a narrow central 'part of the whole stream all of the time'. The cross-sectional areas indicated in red are structurally never available for sampling and will therefore never be able to comply with TOS' Fundamental Sampling Principle (FSP) resulting in a highly significant sampling bias. Despite many claims in the marketplace, it is obvious that the shark fin sampler is not representative. Illustration copyright KHE Consulting teaching collection, used with permission.

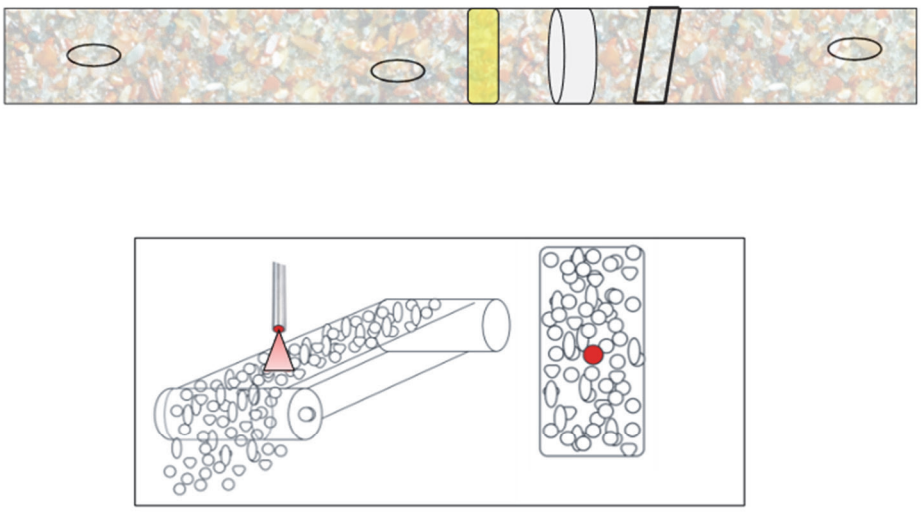

Fig. 13. Principal sketch of a process sampling interface purporting to deliver sensor signals that are representative of the streaming matter in conveyed flows. The realized analytical support volume (red) does not in any way correspond to the imperative demand for a full cross-sectional slice, cfr. Fig. 2. Many hopeful suggestions involve sensors, or cameras, capturing views (narrow-angle, or broad-angle) of moving steams of matters, which fall under the category process sampling. It is obvious however, that such solutions will always fail w.r.t. the imperative slice rule, at best delivering sensor grab samples. The duality with traditionally extracted physical grab samples from the top of the conveyed matter only is striking. Grab sampling is never a solution, neither in TOS' domain, nor in the PAT sensor domain. Illustration copyright KHE Consulting teaching collection, used with permission.

Underlying assumption of cross-sectional homogeneity

Upon reflection, all the non-representative exemplifications presented above, *would* to some degree be able to function – if only all instantaneous flow cross-sections *were* homogenous. This underlying *assumption* is widespread in very many process technology contexts where TOS literacy has not been obtained.

A warning, also part of the process sampling interface discussion, concerns the *by-pass valve*, often used to try to mass-reduce (sub-sample) a streaming flow of matter on the fly. While this objective is highly desirable, it is also very difficult to realise in practice, because of a fundamental misunderstanding. Not any by-pass valve follows TOS' demands for correct increment cutting. In fact, of the five alternatives shown in Fig. 14, only one is TOS compliant! Check out your next blueprint showing 'by-pass streams'. Exactly **how** is this by-pass diversion effectuated? There are many examples in which use is made of one of the non-compliant stream diversion options shown in Fig. 14, the effect of which can be of varying importance, but never neglectable. For full description, see[2,5,13].

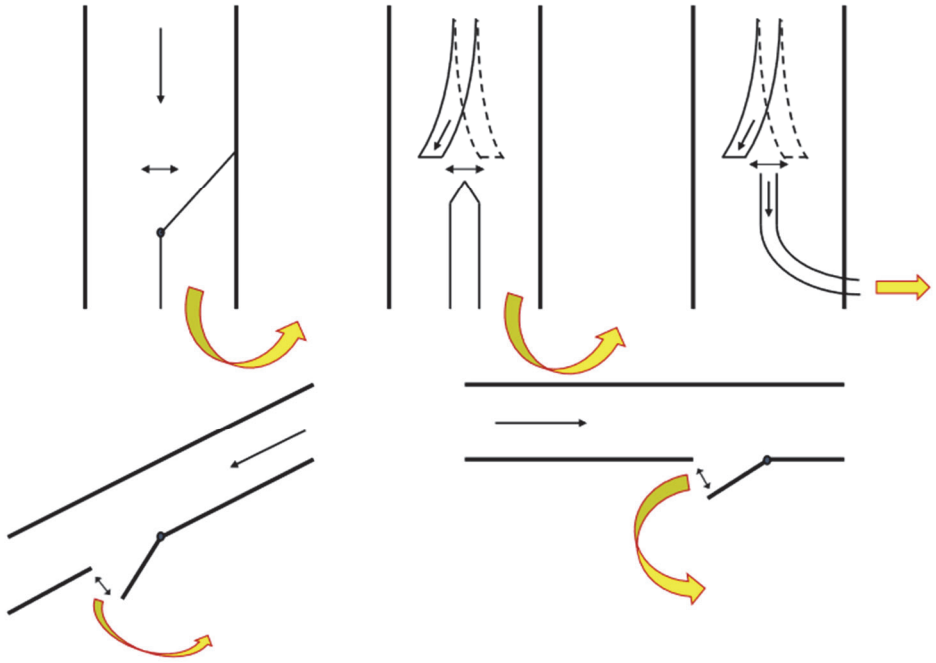

Fig. 14. The by-pass valve enigma. How to divert a moving stream in a representative fashion? Of the five alternative designs for diverting a mass- and flow-reducing by-pass stream, only one is TOS-compliant (top rightmost example). Illustration from KHE Consulting teaching collection, used with permission.

A (very small) compositional caveat

To be fair, many of the process sampling system solutions depicted above actually do work according to their fundamental objectives of securing representative process samples – but this **only** applies to one-phase material systems, e.g. infinitely diluted solutions ('fluids') or mono-disperse particulate materials (uniform materials) a.o. in which there by definition are no compositional heterogeneity, no untoward grain-size distributions variability, and no grain size-grade heterogeneity issues.

Upon reflections, these latter constitute the overwhelmingly dominating proportion of the kind of materials treated with process sampling. Thus, the sampling difficulties illustrated in this paper concerns what in TOS parlance is known as 'significantly heterogeneous materials', which comprise the by far overwhelming proportion of materials targeted for process sampling and analysis in science, technology, industry, trade, commerce, environment … and society.

An ongoing concern is that solutions devised for 'easy' chemical or physical systems have led to a various process sampling interfaces that do a fair job *within* their quite specific restricting limits, but only here. Great problems ensue when such system solutions unwittingly are being carried over to other types of materials of the significantly heterogeneous type – because here they are *bound* to fail, as the intrinsic heterogeneity of such systems (compositional heterogeneity (CH), distributional heterogeneity (DH), grain-size distribution heterogeneity and grain-size-grade heterogeneity obviously deviates from easy-to-sample 'uniform' material systems.

So, unfortunately, as a caveat, this one is very, very small and it cannot be used as justification for carrying-over of any process sampling solution from the 'easy' realm.

How to gauge the performance of a process sampling solution?

All process system solutions, in which the sampling interface interacting with a specific moving material has not been previously characterized, must be subjected to a complete system's performance validation. It is necessary to validate all sampling systems for its status w.r.t. an a priori decided TSE variance threshold, which would qualify the sampling system as 'fit-for-purpose' representative, or not. There is no space here to go into detailed descriptions of the well-known quality assessment approaches just listed below; but full details can be found in very many references within the TOS realm. Suffice here to mention[1-5].

Full system sampling performance facilities:

- Interleaved sampling (IS)
- Replication Experiment (RE)
- Variographic characterization (VC)

So, what to do, then?

TOS calls for compliance validation of all existing designs not previously characterised and for innovative compliance in all future designs of process sampling interfaces. Validation in the view of the above presentations is of critical importance regarding representativity, especially concerning already installed systems that are "too expensive to fail". For all such cases, the key issue is simple: does the existing, or the newly designed sampling solution, capture the *entire* cross-section of flowing streams of matter in an *TOS-appropriate* manner, or not. It cannot be stated simpler …

The hopeful quest for a universal process sampling interface will probably go on – regarding both sensor sampling and physical sample extraction. But is there a universal interface design? Whatever the answer to this question, TOS must serve as the guiding framework for all future developments with Fig. 2 outlining the principal, extremely simple, yet apparently so difficult demand: "Obey the slice rule – or representativity is lost". However, there is no other way!

References

1. K.H. Esbensen, *Introduction to the Theory and Practice of Sampling*, IMPOpen, Chichester, UK. (2020).
2. F. F. Pitard, *Theory of Sampling and Sampling Practice* 3rd ed., CRC Press, Taylor & Francis Group, Boca Raton, FL. (2019).
3. G. J. Lyman, *Theory and Practice of Particulate Sampling an Engineering Approach*, Materials Sampling & Consultancy PTY LTD., Southport, Queensland, AUS. (2019).
4. *TOS Forum*. 10, https://www.impopen.com/tosf-toc/20_10.IMPOpen, 15-19, IMPOpen, Chichester, UK. (2020).
5. K.A. Bakeev (Ed.) *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries* 2nd ed., John Wiley & Sons, LTD., West Sussex, UK (2010).
6. K.H. Esbensen and P. Minkkinen (Eds.) Proceedings of the First World Conference and Sampling and Blending, Special Issue: 50 Years of Pierre Gy's Theory of Sampling, *Chemometrics and Intelligent Laboratory Systems*, **74:1**, Elsevier B.V., Oxford, UK. (2004)
7. K.H. Esbensen and C. Wagner (Eds.) *Proceedings for the 7th World Conference on Sampling and Blending*, in TOS forum, vol **5**, IMPOpen, Chichester, UK (2015).
8. S.C. Dominy and K.H. Esbensen (Eds). *Proceedings of the 8th World Conference on Sampling and Blending*, AusIMM, Carlton Victoria, AUS. (2017).
9. BGRIMM Group (Eds.) *Proceedings for the 9th World Conference on Sampling and Blending*, BGRIMM Group, Beijing, China (2019).
10. K.H. Esbensen and B. Swarbrick, *Multivariate Data Analysis: An introduction to Multivariate Analysis, Process Analytical Technology and Quality by Design*, 6th edition, Camo Software A/S, Oslo, Norway (2018).
11. A.R. Novaes and B.M. Hidding, 'Choosing metallurgical samplers or static cutters for process control in slurry: when or why to avoid the increment delimitation error, in *Proceedings of the 10th World Conference on Sampling and Blending*. IMP Open, Chichester, UK, pp. 171–175 (2022).
12. D. Mittrup, O. Dominguez and M. Haest, "A new approach to implement quality assurance and quality control to technological innovations—quality of spectral data capture and processes in the minerals industry" in *Proceedings of the 8th World Conference on Sampling and Blending*, Ed. by S.C. Dominy and K.H. Esbensen, AusIMM, Carlton Victoria, AUS. (2017).
13. A.P. Ferreira, J.C. Menezes and M. Tobyn (Eds.) *Multivariate Analysis in The Pharmaceutical Industry*, Elsevier, Academic Press, London, UK (2018).
14. R.J. Romañach, A. Joubert Castro and K.H. Esbensen, "WHAT are sampling errors—and WHAT can we do about them? Part 1", *Spectrosc. Europe* **32(2),** 36–42 (2021). https://doi.org/10.1255/sew.2021.a11