Representative sampling and use of HHXRF to characterize lot and sample quality of quartzite at a pyrometallurgical ferrosilicon plant

D. Desroches[a,b], L.P. Bédard[a,*], S. Lemieux[b], K.H. Esbensen[c,d,e]

[a] Sciences de la Terre, LabMaTer, Université du Québec à Chicoutimi, Chicoutimi, Québec G7H 2B1, Canada
[b] Elkem Métal Canada Inc., 2020 Chemin de la Réserve, Chicoutimi, Québec G7J 0E1, Canada
[c] Aalborg University (AAU), Denmark
[d] Geological Survey of Denmark and Greenland (GEUS), Denmark
[e] Sciences de la Terre, Université du Québec à Chicoutimi (UQAC), Québec, Canada

ARTICLE INFO

ABSTRACT

Material sampling is a critical component in mining and mineral processing industries. Nonetheless, sampling is often considered to be a simple matter and, as such, non-rigorous sampling protocols are often applied. The use of inappropriate methods produces inferior, non-representative estimates of sampling target composition. To address weaknesses in sampling protocols and evaluate the representativeness of collected samples, we performed a feasibility study of the ability of handheld X-ray fluorescence (HHXRF) to achieve a satisfactory characterization of a raw material lot at a pyrometallurgical ferrosilicon plant. Using composite and grab samples, we determined the various sampling error manifestations stemming from the fundamental sampling error, grouping and segregation error, as well as increment delimitation, increment extraction, and increment preparation errors), and performed a first foray determination of optimal sample mass, and estimated the heterogeneity within the sampling target. HHXRF results were compared with the results obtained using laboratory XRF. A first estimate of optimized sample mass for HHXRF was 10 kg, given the large size of crushed quartz blocks used in ferrosilicon plants—roughly cubic, 10 cm per side; accuracy improved with increased sample mass (18% error with a 10 kg sample versus 35% error when using a 1 kg sample). A 10 kg sample is also the mass a technician can realistically transport from the sampling site to the preparation facilities. The main contribution to the global estimation error is from primary sampling. Variographic analysis illustrated a sill equal to the nugget effect, indicating that two adjacent samples are no more similar than two samples separated by larger distance; this suggests equal spatial heterogeneity at all scales larger than the increment mass in the sampling target. Analytically, the HHXRF and desktop XRF results compared very well. Overall, the error associated with our first attempt at field composite sampling was half of that obtained via grab sampling for both the HHXRF and desktop XRF protocols. Relative to conventional analysis based on grab sampling and analysis via desktop XRF, the use of handheld XRF coupled with composite sampling would appear to be a feasible approach for an improved sampling protocol for obtaining fit-for-purpose characterizations of industrial quartzite.

1. Introduction

The theory of sampling (TOS) addresses heterogeneity, one of the most important phenomena in the characterization of materials. Heterogeneity, regardless of scale, complicates sampling of a larger sampling target, i.e. the lot, as there is no guarantee that the composition of smaller samples is identical to that of the lot to be processed if sampling is based on a non-representative approach.

Despite this fundamental uncertainty, it is surprising to find that sampling strategies and the verification of sample accuracy are often trivialized. The primary objective of sampling is to provide reliable information about the entire lot with a known and acceptable level of total uncertainty (sampling + analysis variability). Nonetheless, few industry professionals change from established, albeit potentially erroneous, sampling strategies to a more representative approach without serious reasons. Thus, the collection of a single sample, or a few haphazardly selected specimens, i.e. grab sampling, continues to dominate sampling practices in many instances.

Grab sampling does not, however, provide representative samples except by chance (DS 3077, 2013; Gy, 1998; Pitard, 2009; Esbensen and

---

Julius, 2009; Minkkinen and Esbensen, 2009; Esbensen and Wagner, 2014; Esbensen et al., 2018). Sampling is almost always carried out as a single operation, as this is the quickest, least expensive, and least effort-demanding approach. Only in rare instances, where the sampled lot material is of relatively uniform composition, is grab sampling potentially acceptable and relevant. Nonetheless, the representativeness of these selected cases still requires verification, as is the case in any other case with significant heterogeneity.

Most laboratories provide a valid estimate of the total analytical error (TAE); however, few laboratories include sampling error contributions (total sampling error, TSE) despite there often being 10–100× greater than TAE (Gy, 1998; Taylor et al., 2007; Pitard, 2009; Esbensen and Julius, 2009; Esbensen and Wagner, 2014; Esbensen et al., 2018). Typically > 80% of the global estimation error (GEE = TAE + TSE) can be due to sampling (primary sampling + mass-reduction sub-sampling in the laboratory) and < 20% stems from analysis (Esbensen and Petersen, 2005; Esbensen and Wagner, 2014). Any decrease in GEE improves analytical relevance. Unfortunately, industry largely continues to avoid addressing the contribution from sampling errors.

To improve the validity of sample measurements, standard sampling protocols must be revised to reduce or eliminate the most significant and dominant errors. In the mining and mineral processing sectors, increasing profitability often requires an increase in production capacity with a concomitant need to process greater amounts of unprocessed material. This scenario implies larger lots to manage, monitor, and characterize, and this is often combined with increased processing rates. A likely consequence is that it becomes more difficult to obtain representative results—from samples typically on the order of grams—when the target lot is on the order of several hundred metric tonnes. For such large mass-reduction situations, it is particularly important that the sampling practice be guided by TOS (DS 3077, 2013; Esbensen and Wagner, 2014; Esbensen and Julius, 2013; Esbensen and Petersen, 2005).

The use of handheld X-ray fluorescence (HHXRF) tools can address some of the key issues surrounding the need to obtain rapid, in-situ chemical characterization with minimal sample preparation (Taylor et al., 2005). HHXRF may improve the characterization of large lots due to its transportability and its ability to sample at a higher throughput (Ramsey and Boon, 2012; Ramsey et al., 2013). However, few studies have demonstrated the potential of HHXRF to provide fit-for-purpose representative results in an industrial context (Bédard and Barnes, 2011; Thompson and Ellison 2006; Thompson and Feam, 1996).

In an earlier study, the specific analytical performance of a dedicated HHXRF showed good potential in evaluating and quantifying quartzite geochemistry (Desroches et al., 2018). Furthermore, this study confirmed previous observations that correcting initial results with reference materials improves significantly the analytical accuracy of final estimates (e.g. Cohen et al., 2017; Gazley and Fisher 2014; Fisher et al., 2014; Hall et al., 2014; Lemière, 2018; Quye-Sawyer et al., 2015; Ross et al., 2014; de Winter et al., 2017). High purity quartzite, despite its chemically simple matrix, has a highly complex spatial heterogeneity of low concentration levels of intrinsic impurities. Using powdered samples, Desroches et al. (2018) nevertheless showed that an HHXRF (a Niton XL3t GOLDD+) can provide a reliable quantitative analysis even at such low concentrations as are of importance in the present feasibility study. Here, we assess the performance of HHXRF in its ability to achieve satisfactory representativeness within the context of an industrial lot at a ferrosilicon plant.

2. Materials and methods

2.1. Elkem Métal Canada ferrosilicon plant

Elkem Métal Canada Inc. is a producer of ferrosilicon alloy. Production relies on two main pathways characterized respectively by a base of 50% or 75% silicon, and it has several customer-specified additive options, e.g. foundry and steel plants. For this type of production, the plant reduces quartz (SiO_2) by a combination of melting in a blast furnace and Söderberg electrolysis (Collins et al. 1974; Lemieux 2009). For this process to be successful, the composition of quartz (more precisely, quartzite) must be carefully monitored and controlled. Contaminants, such as aluminum, can dilute the silica phase. They become less available for reduction and accumulate as a viscous melt in the furnace (Tuset 1992; Tomé Torquemada 2019). Although quartz may appear to be a simple matrix, the low concentrations of impurities—in the form of discrete minute grains of ilmenite ($FeTiO_3$) or magnetite (Fe_3O_4), for example—increase the difficulty of accurate analysis. These impurities exhibit an extreme degree of spatial heterogeneity and thus make proper sampling highly challenging.

As the matrix is essentially SiO_2 (> 98% m/m), Elkem Métal does not analyze individual shipments of quartz but instead relies on traditional grab samples over contractually specified volume/time periods. This protocol is a major potential weakness for proper QC/QA and represents one of the main reasons driving this study.

2.2. Theory of sampling overview

Although grab sampling may, by chance, result in a sample having a broadly similar composition as the larger lot, it is never possible to identify when this situation occurs in practice. TOS can provide the necessary framework to ensure representative sampling of a lot. TOS is also the proper basis upon which specific HHXRF evaluations should be carried out.

2.2.1. Fundamental sampling principle

According to TOS, all units of the lot, *i.e.* minerals and/or increments, must be fully accessible for sampling; this is the fundamental sampling principle (FSP). A "lot" designates the sampling target which has a specific scale; *i.e.* a stockpile, a barrel, a truck or railroad load, or a geological outcrop. Its precise meaning depends on the specific situation. All lots are comprised of a specific material, occupying a specific geometrical volume, and they have a specific lot mass, density, etc. The common adverse characteristic of all lots in regard to the possibility of proper sampling is heterogeneity; therefore, a focused and unified approach is required. In the mining and mineral processing sectors, surface samples of a stationary lot are often assumed to be sufficient for characterization of a lot; however, such protocols do not follow TOS.

TOS also mandates the use of composite sampling with a heterogeneity-related number of increments (Q). An increment designates a single extracted unit, which is often defined by the volume of the sampling tool. When combined with other such units, these 'partial lot units' make up a composite sample. The only free parameter in composite sampling is the required number of increments Q to ensure the results are fit-for-purpose. Combining spatially dispersed Q that cover the entire lot is thus the only sampling process that will be acceptable; this protocol should also be able to indicate the total sample weight (Q × the increment weight).

2.2.2. Sampling errors

TOS identifies five types of error that may occur during the sampling of stationary lots (Esbensen and Julius, 2013; Minkkinen and Esbensen, 2009; Pitard, 2009); these errors are listed in Table 1, organized according to their logical order in the sampling process and not necessarily in order of their quantitative importance. Two additional specific errors occur when sampling moving or dynamic lots (Esbensen and Julius, 2013; Minkkinen and Esbensen, 2009; Pitard, 2009). The fundamental sampling error (FSE) is material specific; for example, it can only be reduced via physically modifying the sampled material by decreasing the size of the largest particles via crushing, technically comminution. FSE is inversely proportional to the realized sample mass; however, increasing sample mass by itself will not significantly

Table 1
Sampling errors and their sources.

Error type		Error source
Fundamental sampling error (FSE)		Associated with lot heterogeneity; FSE can only be reduced by comminution (crushing)
Grouping and segregation error (GSE)		Associated with a heterogeneous spatial distribution; GSE can be reduced by composite sampling and by thorough mixing
Incorrect sampling error (ISE)	Increment delimitation error (IDE)	Difficulties in reproducing the exact delimitation geometry of the increment being sampled; Can be reduced, or eliminated by reproducing the delimitation exactly
	Increment extraction error (IEE)	Particles do not have the same probability of being part of the increment; IEE can be reduced by using properly designed, maintained and monitored sampling tools
	Increment preparation error (IPE)	Errors reflecting non-controlled adverse handling, mixing, and comminution of samples or sub.samples; IPE can be reduced by careful sample preparation procedures and due diligence, *i.e.* good laboratory practices (GLP)
Total	Total sampling error (TSE)	TSE = FSE + GSE + IDE + IEE + IPE

improve the representativeness of the sampling process as all grab samples represent only a minute fraction of the entire industrial lot.

The grouping and segregation error (GSE) increases as spatial heterogeneity increases. Heterogeneity is most often due to the presence of local "hot spots," *i.e.* markedly higher or lower analyte values than the average concentration due to grouping, stratification, and segregation. GSE also occurs due to temporal heterogeneity whenever transportation plays a role, *i.e.* smaller fragments and denser fragments will segregate downward during transport-induced shaking. This error can be reduced and even almost eliminated on occasion by effective composite sampling and, for smaller lots, forceful mixing.

The increment delimitation error (IDE) occurs when the geometric delineations of the increments being sampled are unable to be reproduced precisely. The increment extraction error (IEE) relates to the extraction of increments with a specific sampling tool. Sampling becomes more accurate depending on that all particles can be extracted with equal probability by the sampling tool employed; increment must only reflect the delineated inctement volume (no contamination etc.). IEE can be reduced by careful equipment design and careful monitoring of the sampling process. The increment preparation error (IPE) encapsulates all errors produced during post-sampling handling, mixing, comminution, transportation, or other processes in the laboratory. It is related to the potential loss of matter—via dust, moisture, improper cleaning, etc. IPE is also a reflection of potential chemical or physical alteration of the sample, and any negligence or non-adherence to good laboratory practices by technicians. These three sampling errors, IDE, IEE, and IPE, are collectively termed the "incorrect sampling errors" (ISE). These are critical sampling errors because, if not eliminated from the sampling process, they generate critical biases, which cannot be corrected by any means (Esbensen and Wagner, 2014). The sum of all these five identified errors represent the total sampling error (TSE). The increment weighing error (IWE) could be added to the TSE. However, IWE is not directly involved in HHXRF sampling, which typically reflects a comparable analytical volume. In addition, acceptable variation for increment weights during conventional process sampling can be up to ± 20% (Esbensen and Petersen, 2005).

2.2.3. Replication experiment

Variation between samples that have been sampled by the exact same procedures can be quantified by analyzing replicate samples. The precise number of replicates needs to be defined after careful consideration (see Esbensen and Wagner, 2014; DS 3077, 2013). Replication experiments (RE) can quantify the effects of variation in total sampling, processing, and analytical error (Juran and Godfrey, 1998; DS 3077, 2013; Esbensen and Julius 2009; Esbensen and Wagner, 2014). RE can determine whether the HHXRF procedure is fit-for-purpose, when applied 'from the top' *i.e.* when replication starts with the primary sampling - and, when applied hierarchically, can also identify those subsequent steps that contribute the most to the overall

variability. Thus, RE can improve understanding of the entire "lot-to-aliquot" pathway. RE can be applied to any new or existing sampling procedure, where RE must always begin at the primary sampling stage to ensure that the all sampling errors are included in the estimate of the total measurement uncertainty.

The RE approach can be applied to primary sampling or in a hierarchical fashion to each sampling, processing, or analytical stage. The latter quantifies the effective errors that are produced at each stage. We selected this approach in the present study as our main feasibility objective was to quantify the effective error for each step of the sampling procedure. The evaluation of each step requires a minimum of ten replicate sampling operations (Esbensen and Julius, 2009). With five identified stages, this translated into fifty analytical samples: beginning with ten primary samples (PSE), we then prepared, in an identical manner, ten new samples from one of the primary samples, selected at random. We repeated this protocol—ten subsamples from a randomly selected sample—at each step, cascading hierarchically down through the sample preparation steps (Fig. 1). From this protocol, we could identify and quantify the procedure, stage, or equipment that contributed the most to the empirical error. This is an enfolding experiment design (elaborately explained in the TOS literature referred to above, keeping the total number of "runs" at a manageable level. Note that this approach is different from a traditional Design of Experiment approach with the intention of being variance decomposed by ANOVA. One is in no way able to control all factors involved in the present, practical design which instead follows the lot-to-aliquot pathway in a straight realistic fashion.

2.2.4. Sampling and sample preparation

The first step in mass reduction is crushing. Although reducing sample mass appears simple and straightforward, mistakes do occur, and highly significant sampling errors can be introduced during mass reduction (Petersen et al., 2004). By minimizing the largest particle sizes, the total lot heterogeneity can be reduced most effectively by subsequent mixing, and the efficiency of mixing is also improved for smaller particles. This reduces both GSE and FSE, leaving the comminuted particles a markedly improved chance of being sampled evenly. If mixing occurs with maximum effect, GSE can in many cases be almost eliminated, especially for operations that can be controlled at the laboratory scale.

Composite sampling involves collecting many increments (Q) to form an aggregate, composite sample. The greater the number of increments used, the greater the likelihood that the resulting aggregate sample will be a better representative of the full lot. However, the spatial pattern of the increment locations must comply with the FSP. By adhering to the latter (FSP), this approach has a profound effect on TSE, which will decrease as the number of increments increases (DS 3077, 2013); Esbensen and Wagner, 2014). FSE is inversely proportional to the mass sampled, but this is only of interest for composite samples, not
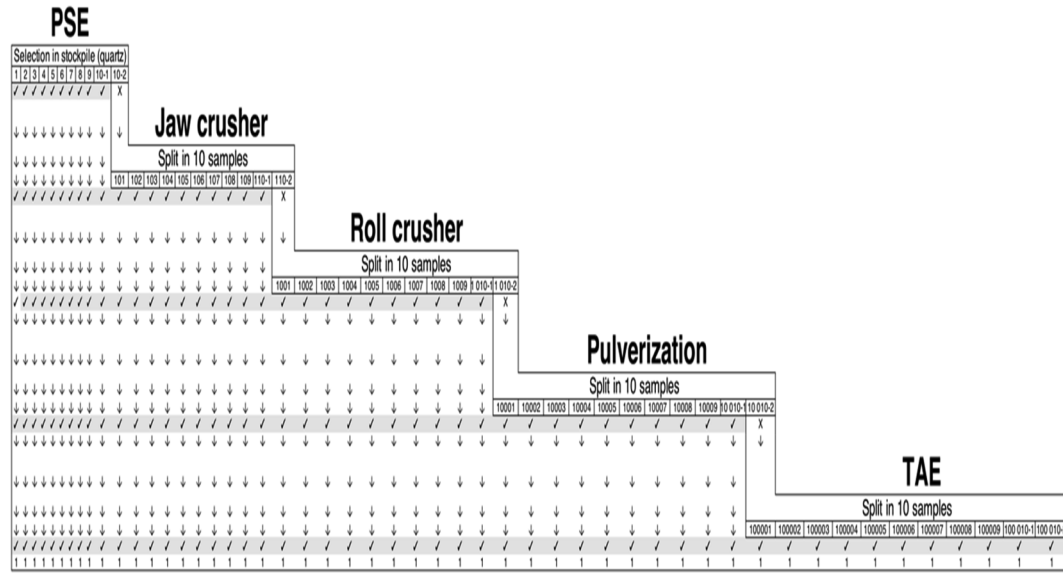
Fig. 1. The hierarchical replication experiment. At each stage, a randomly selected sample is divided into ten subsamples and processed in the following step (shown here as the tenth sample, but this positioning is for illustration purposes only). PSE: primary sampling stage; TAE: total analytical error.

for individual grab (bulk) samples. Q increments must be deployed to represent the full geometry (the full volume) of the lot; this is the only means by which composite sampling reduces GSE, and this is very often the only approach available at the primary sampling stage. This insight was also the basis for designing the handheld XRF procedures investigated in this study.

There is a natural, minimum state of heterogeneity for all aggregate materials; in fact, homogeneity is an ideal state that does not exist for all naturally occurring and technological materials (Pitard, 2009; Gy, 1998). In contrast, 'too much' mixing can potentially increase segregation temporarily, and will always end up producing a steady state mixing–demixing situation (Esbensen and Julius, 2013).

2.3. Variogram

Variographic analysis provides a superior characterization of the combination of a specific sampling process and the heterogeneity of the lot material. Thus, this approach is always preferred for characterization (DS 3077, 2013; Pitard, 2009; Esbensen et al., 2007; Esbensen and Wagner, 2014). A variogram describes simultaneously, at all scales, the empirical variance of a given one-dimensional data series. The variogram presents variation as a function of inter-sample distance, called the lag (h). This data series may stem from process data collection or, as in this study, from a transect along which samples have been collected. Minnitt and Esbensen (2017) explain variograms in full operative detail, and their application to the mineral industry is presented by Engström and Esbensen (2017a, 2017b). The variogram outlines the degree of temporal or spatial autocorrelation (the variogram range). Its most useful feature for our study is its ability to highlight the nugget effect. The nugget effect quantifies total empirical sampling, preparation, and analytical error relative to the total observable process or transect variability, called the sill of the variogram. The derived feature, the nugget effect/sill ratio, decomposes the variation within the raw serial data into the noise of the total measurement system versus the true (decomposed) transect variability (DS 3077, 2013; Engström and Esbensen, 2017a, 2017b; Esbensen et al., 2007; Minnitt and Esbensen, 2017, Pitard, 2009). On this basis, the possibility of observing a more-or-less reduced transect variability is not seldomly surprising. Variographics is the only know approach that is able to decompose the total "measurement uncertainty" (MU) from the underlying transect variability reflecting the spatial heterogeneity in the lot.

2.4. Experimental approach

2.4.1. Determining optimal sample mass

In industrial settings, such as mineral processing, critical tasks are outlined in formal guidelines, standards, and procedures. However, such documents do not always specify the optimized or required sample mass, but they often recommend a "constant mass". Worse still, such fixed sample masses that are often derived from historical and non-TOS practices are unable to deal with the variable heterogeneity within lots. At Elkem Métal Canada Inc., the primary mass of a typical sample is ca. 10 kg, this being the weight limit that allows a technician to transport the collected material easily from sampling points to the laboratory. Note that this mass historically was determined without any preliminary study or prior knowledge of TOS for that matter.

To determine whether a 10 kg mass is adequate for fit-for-purpose sampling in the present context:

1. For each experiment detailed below, we collected 55 kg of quartzite grab samples (crushed blocks of quartzite measuring about 10 cm on a side) from the stockpile (Fig. 2);
2. We spread the samples on a table;
3. From these samples, we randomly selected quartz rocks to form ten subsamples of increasing mass. The first subsample was 1 kg, the second subsample 2 kg, the third was 3 kg, etc. up to 55 kg of rock sample in total (1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 kg = 55 kg);
4. We then prepared each subsample (1, 2, 3 … 10 kg)—crushed, pulverized, and mass reduced (split) to an analytical mass (50 g)—and then analyzed as a regular sample (Fig. 3, yet lacking the mixing stage).

The weighted average of the samples was determined by:

$$\text{Weighted Average} = \sum_i \lfloor \frac{\%\beta_i \times i}{55} \rfloor$$

where i represents the masses of 1 to 10 kg, β represents % element (such as Fe_2O_3), and 55 represents the total mass of the lot (55 kg).

For determining the accuracy of each mass, we followed:

$$\text{Accuracy } \beta_i = \lfloor \frac{\%\beta_i - \text{weighted average}}{\text{weighted average}} \rfloor$$

This calculation was carried out for each element (analyte), see

ARTICLE IN PRESS

D. Desroches, et al.   Minerals Engineering xxx (xxxx) xxxx

Fig. 2. Sample collection from the quartzite stockpile. The stockpile dimensions are typical for this type of industrial setting. Note that Elkem Métal currently extracts surface samples only. The layout of the sample bags shows the transect profile lines for the variographic analysis.
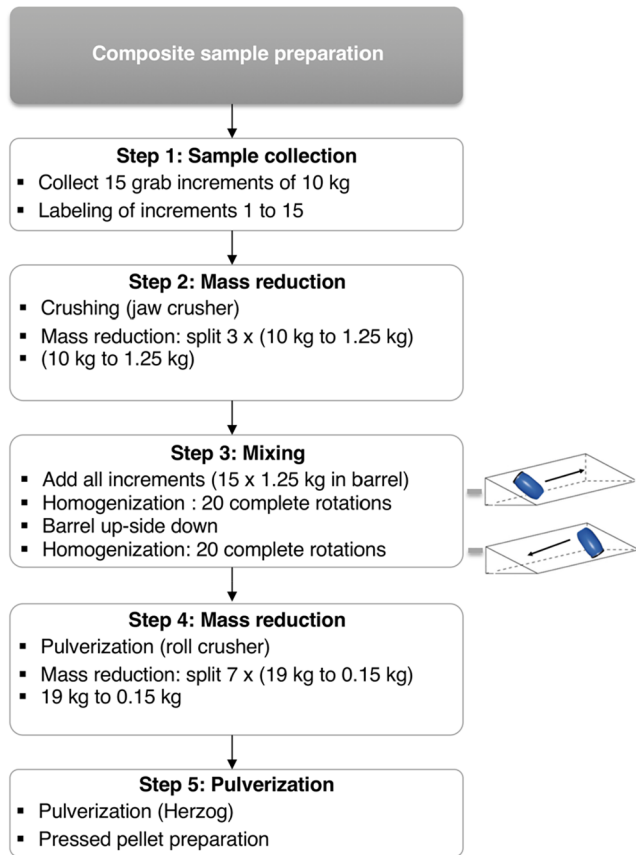
Fig. 3. Protocol for preparing composite samples as used in the present study.

further below. To produce an overall representative estimate of accuracy, we repeated these measurements eight times. We then calculated the mean accuracy corresponding to each mass.

We determined the mean accuracy via the following equation:

$$mean \ Accuracy_i = \lfloor \frac{\sum Accuracy \ \beta_i + \sum Accuracy \ \alpha_i + \sum Accuracy \ \mu_i}{24} \rfloor$$

where i represents the mass (1 to 10 kg), β represents %Fe_2O_3, α represents %Al_2O_3, μ is %TiO_2, and 24 represents the number of elements multiplied by the number of realized repetitions (3 × 8).

2.5. Replication experiment

To determine the stage that most contributed to the GEE, we

performed a replication experiment in the form of repeated sampling of the primary sampling stage along the complete 'from-field-to-analysis' pathway (Esbensen and Wagner, 2016). We subsequently also ran a hierarchical RE (Fig. 1), allowing us to identify the primary sampling error (field sampling), primary crushing error (jaw crusher), secondary crushing error (roll crusher), pulverization and homogenization error (Herzog Pulverizers, HP–MA), whereas the total analytical error (TAE) is known from specific analytical validations.

2.6. Sampling methods

All samples, whether grab, composite, or variogram samples, were collected on the same day from the same primary lot. The estimated total lot mass (from the calibrated truckloads) was ca. 5000 t. We also prepared ten primary composite samples, each having a mass of 150 kg from this lot, as follows:

1. Each composite sample comprised 15 individual grab increments (of 10 kg each), collected randomly over the entire surface of the stockpile (Fig. 3), for a total of 150 kg;
2. Each 10 kg increment was crushed (grain-size reduction) and the mass reduced (split) down to 1.25 kg. We performed mass reduction using a riffle splitter following a TOS-compliant approach (Petersen et al., 2004). We prepared each increment in identical fashion;
3. We then aggregated and homogenized the 15 × 1.25 kg samples by mixing them in a plastic barrel (half-filled with 15 × 1.25 kg ≅ 19 kg) that we rolled on an inclined plane (approx. 30° incline) to ensure effective mixing (Fig. 3);
4. We emptied the barrel's contents (ca. 19 kg) slowly into the roll crusher and reduced it to a ca. 100 mesh; we then mass reduced the contents seven times down to 0.15 kg, again using riffle slitting;
5. For the XRF analyses, we pulverized this 0.15 kg fraction to a 400 mesh (37 μm);
6. We repeated all the steps for each of the ten composite samples.

For the variogram analysis, we characterized the typical stockpile of quartz (Fig. 2; ca. 5000 t) based on 65 samples collected along a randomly selected directional transect over the surface of the stockpile. The lag distance between increment locations was 30 cm.

To compare the alternative sampling protocols, we followed a strict procedure for all subsequent crushing and splitting steps (Fig. 3).

2.7. Desktop XRF and HHXRF equipment

We selected a Niton XL3t GOLDD + (8 mm beam footprint diameter) developed by Thermo Fisher Scientific (Tewksbury, Massachusetts, United States) as the HHXRF unit for this study. The instrument has an

Rh tube with a Silicon Drift Detector (SDD). We ran all analyses in the "TestAllGeo" mode with an analysis time of 60 s (Desroches et al., 2018). HHXRF can be used with minimal sample preparation to obtain a fast turnaround (direct measurements on mineral surfaces). As such, we tested this instrument within an industrial setting using unprepared quartzite; thus, conditions differed markedly from those surrounding the laboratory-based desktop XRF run on prepared samples.

We used the XRF laboratory facilities at Elkem Métal and UQAC were used to compare analytical results from the field (HHXRF, UQAC) and laboratory (desktop XRF, Elkem Métal). The Elkem Métal laboratory is equipped with a WDX Primini (Rigaku, Tokyo, Japan), which uses a sequential-type X-ray fluorescence spectrometer. This model can analyze the elements fluoride (Z = 9) through uranium (Z = 92) (Interactive Corporation, 2014). The X-ray chamber operates in an atmosphere controlled by P10 gas (argon/methane) having a pressure system during analysis down to one Pa. The Pd X-ray tube has a voltage of 40 kV and 1.25 mA.

We applied the Eurachem approach to establish the specific HHXRF detection limits for this quartzite (Currie, 1995). This approach relies on determining the acceptable standard deviations from repeated measurements on samples having similar matrices but with diminishing concentrations. Desroches et al. (2018) established detection limits of 28 ppm for TiO_2, 1100 ppm for Al_2O_3, 66 ppm for Fe_2O_3, 3400 ppm for MgO, and 66 ppm for CaO, limits that are acceptable for the needs of this study. We recorded five measurements along a 10-cm-side block of quartzite to improve estimate accuracy, e.g., from 21% to 12% for TiO_2, with accuracy improving even further with additional measurements, although this decreases efficiency (Desroches et al., 2018). The quality of analytical results is negatively affected when surfaces are not flat, as this produces variation in the distance between the detector and the analyzed surface. This distance is critical in XRF analysis, and more particularly for in-situ determinations, as variation in this distance alters concentration estimates (Potts et al., 1997; Potts and West 2008; among others). Highly erratic results (2–3× above/below average) obtained due to an irregular surface were removed from our analysis (duely noted where/when applied).

The analytical quality will be affected by small chemical heterogeneities. The presence of sub-microscopic minerals having trace elements as their major constituent, such as Fe-Ti oxides (ilmenite, $FeTiO_3$), can alter concentration estimates significantly. The presence of minute minerals under the XRF beam will certainly influence analysis (Desroches et al., 2018); for example, a single grain of ilmenite ($FeTiO_3$) of less than 0.5 mm can increase concentrations by 0.05% TiO_2, a value close to the measured concentrations in the studied material.

With the HHXRF, we analyzed each quartzite block within a 10 kg sample bag. We ran more than 500 in-situ analyses with the HHXRF. We applied the grand mean of all analytical results, and we submit that this approach provides a relevant and reliable approximation of the composition of the full stockpile, constituting a useful datum with which to assess all variabilities.

3. Results

3.1. Estimating the concentrations of trace impurities in the lot

Given that there is no access to a single, true value of the various elemental concentrations of a lot, the best estimates are obtained from the above analytical average of the 500 multiple measurements. For industrial projects, the aim is to achieve an acceptable 'best estimate' at the lowest cost; therefore, a compromise is very likely regarding the sampling strategy. The adequate number of samples, increments, measurements and the derived minimum mass, will depend on the needs of the given industrial process and context and, in particular, the degree of inherent and observed heterogeneity.

The combination of the different sample types in this study, *i.e.* composite, grab, and transect samples for variographic analyses,

Table 2
Analytical determinations of trace element chemistry in quartzite as determined by desktop XRF and HHXRF. Both instruments analyzed pressed pellets.

Desktop XRF					
	TiO$_2$ (wt%)	Al$_2$O$_3$ (wt%)	Fe$_2$O$_3$ (wt%)	MgO (wt%)	CaO (wt%)
Mean (n = 230)	0.055	0.302	0.059	0.024	0.002
RSD (%)	23.4	60.0	57.5	99.7	162.4
HHXRF					
Mean (n = 230)	0.044	0.369	0.042	0.684	0.005
RSD (%)	25.2	36.1	39.6	23.7	40.4
LOD	0.0028	0.11	0.0066	0.34	0.0066

Note: For CaO, many results are below LOD; see Desroches et al. (2018). wt % = weight % or % m/m.

produced a total of 230 bags of quartz from the stockpile. We used these samples to calculate the 'best' estimate of the overall trace impurity concentrations in the entire lot and determine the associated sampling error. We determined the relative standard deviation (RSD) of the impurity concentrations using both the HHXRF and desktop XRF (Table 2).

Desktop XRF and HHXRF produced similar ranges and comparable results, except for MgO. RSD levels for HHXRF were high as all the remaining elements are found in very low concentrations, and Mg, Ca, and Al are analytically 'light elements.' RSD values for MgO and CaO are extremely high for the desktop XRF due to the higher sensitivity of the desktop instrument combined with the very low absolute concentrations of these two elements. A full comparison of the results obtained by both instruments in terms of their estimates of the composition of a quartz matrix is presented in the parallel study described in Desroches et al. (2018).

3.2. Determination of optimal sample mass

To determine a first optimized sample mass under the constraining requirements described above, we replicated chemical analyses on one sample bag using the desktop Primini XRF. Accuracy has been computed as the obtained result for a mass minus the weighted average divided by average of all results in a test (Table 3). Accuracy was found to decrease from ca. 35% to 20% as sample mass increased from 1 kg to 10 kg (Fig. 4). Considering the sample weight that can be transported by a technician, sample preparation time, and quartz block size, 10 kg is the de facto maximum mass in the present industrial context. Also, for health and security reasons, a > 10 kg sample should not be used according to industrial security norms.

3.3. Replication experiment

For sample preparation to be adequate for XRF analysis, we must identify those steps that most contribute to the overall uncertainty of the sampling preparation process. We use Fe as an example, as it presents typical relative sample variability values (RSV) for the entire process; this pattern suggests that sample preparation is well controlled (Fig. 5). The largest RSV values were produced by primary sampling, followed by the jaw crusher stage. Typically, primary sampling produces the largest error, as this error is related to the lot heterogeneity of the sampled material (Juran and Godfrey, 1998; DS 3077, 2013; Esbensen and Julius, 2013). During the second step, portions of the steel jaw and roll crushers that contact the quartz contribute to Fe contamination and thereby increase the RSV.

3.4. Variogram

Variograms are sensitive to extreme values and outliers (Esbensen et al., 2007). For CaO, almost half of the HHXRF values fell below detection limits, thereby rendering these results meaningless; for this

Table 3
Total measurement accuracy for different sample masses.

Mass / Test	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg
		1				2				3				4				5		
1	48%	53%	12%		26%	45%	1%		26%	24%	154%		15%	54%	20%		7%	72%	25%	
2	65%	55%	29%		6%	3%	11%		29%	29%	3%		18%	12%	15%		6%	8%	9%	
3	25%	6%	6%		16%	15%	21%		29%	69%	14%		131%	2%	70%		138%	10%	14%	
4	41%	21%	31%		7%	15%	12%		24%	12%	1%		27%	31%	9%		15%	36%	34%	
5	73%	46%	12%	35%	15%	162%	35%	33%	80%	44%	38%	34%	73%	3%	0%	31%	19%	9%	53%	33%
6	49%	28%	0%		77%	80%	36%		0%	13%	23%		30%	12%	7%		2%	26%	31%	
7	55%	28%	4%		1%	21%	4%		39%	26%	30%		50%	22%	13%		60%	6%	16%	
8	76%	30%	36%		69%	89%	18%		16%	77%	21%		44%	64%	24%		82%	153%	61%	

| Mass / Test	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg	Al_2O_3	Fe_2O_3	TiO_2	Avg
		6				7				8				9				10		
1	23%	34%	20%		29%	10%	27%		20%	21%	23%		12%	19%	19%		29%	44%	20%	
2	41%	21%	6%		19%	47%	17%		18%	240%	39%		58%	47%	26%		11%	91%	6%	
3	27%	64%	0%		54%	19%	8%		32%	16%	21%		1%	21%	21%		27%	10%	6%	
4	21%	6%	17%		46%	8%	11%		57%	28%	17%		16%	17%	8%		35%	73%	35%	
5	42%	40%	12%	26%	86%	52%	3%	22%	36%	18%	3%	24%	32%	21%	3%	16%	36%	36%	34%	19%
6	26%	47%	0%		7%	14%	49%		1%	94%	0%		8%	68%	13%		10%	4%	5%	
7	49%	18%	13%		157%	5%	13%		48%	99%	42%		44%	18%	7%		55%	52%	22%	
8	37%	38%	17%		34%	77%	26%		37%	16%	1%		3%	103%	5%		23%	56%	3%	

Note: avg = average.

reason, we did not assess CaO. The remaining variograms (Fig. 6) showed a random spatial variation for all major impurities, including Fe_2O_3, Al_2O_3, MgO, and TiO_2, for both the HHXRF and the desktop XRF. The variograms were produced using a lag of 30 cm. All variograms were of the "flat variogram"-type (known as "pure nugget effect" variograms), demonstrating that there was no spatial autocorrelation even at this low lag scale (Esbensen and Wagner, 2014; DS 3077, 2013).

HHXRF and the desktop XRF results for TiO_2 were similar with a sill of 0.08. Al_2O_3, Fe_2O_3, and MgO also produced similar flat variograms although the HHXRF sills were smaller: 0.15 versus 0.4 for Al_2O_3 and Fe_2O_3, 0.1 versus 0.7 for MgO. The lower sensitivity of the HHXRF

explains the lower sill values as does the lack of crushing that results in less contamination.

A flat variogram may also reflect a situation in which the lag chosen for characterizing the variogram is mistakenly too large for the encountered small-scale heterogeneity (Minnitt and Esbensen, 2017; Engström and Esbensen, 2017a, 2017b). However, we selected a lag of 30 cm, which is the smallest scale of practical interest for this particular industrial situation characterized by lots of ~5000 t and mineral blocks having a mean width of 10 cm. We submit that 30 cm is a highly relevant scale for the practical primary sampling.
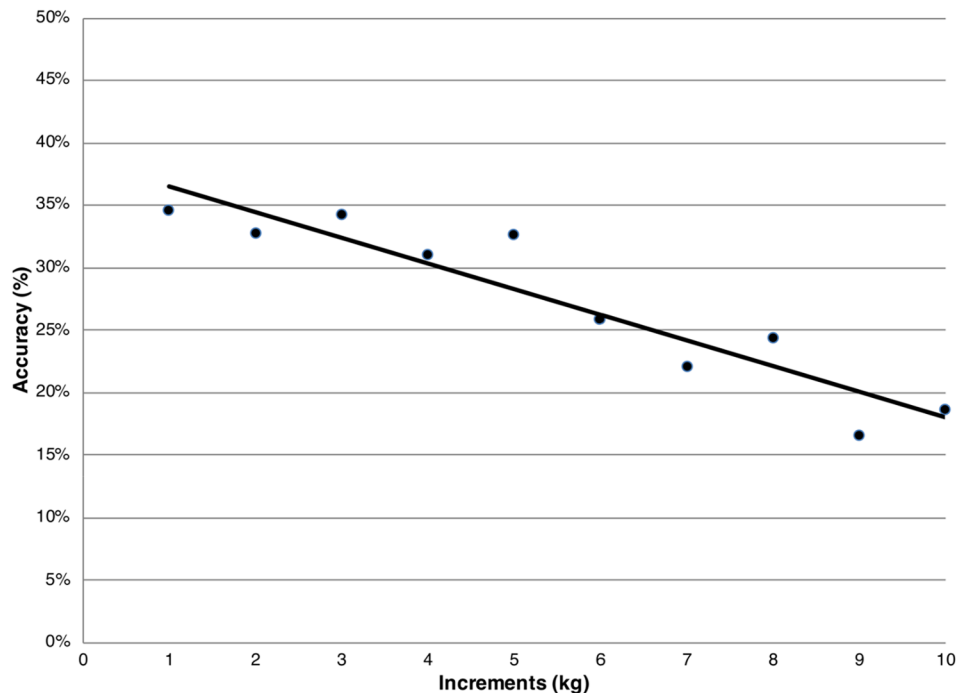
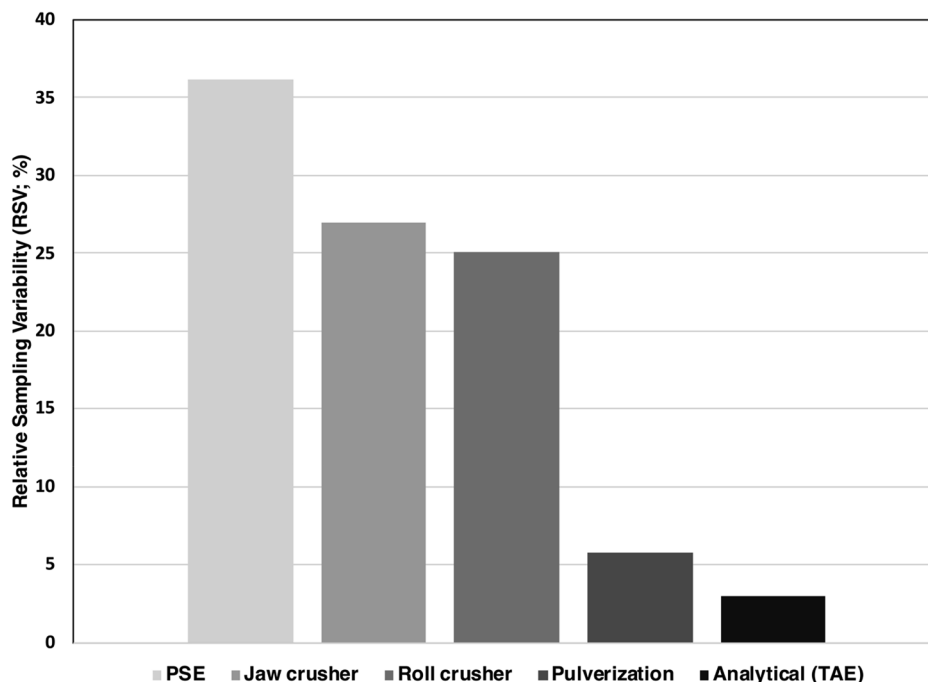Fig. 4. The relationship between accuracy and sample mass.

Fig. 5. Results of the hierarchical replication experiment, used to identify error contributions from each step, from primary sampling to analysis. The vertical scale is the relative standard deviation (RSV; %, standard deviation/average of replicated results × 100 for Fe_2O_3) as proposed by Esbensen and Wagner (2016).

3.5. Analytical performance

Comparing the concentration estimates for the ten samples analyzed in situ using HHXRF with those obtained via desktop XRF on pressed pellets, the desktop XRF produced better results than HHXRF for Al_2O_3 and Fe_2O_3, whereas both instruments had similar results for TiO_2 (Fig. 7). We added a reference material (CRM STD Sand No. 8; Society of Glass Technology) to test the reproducibility of the analytical methods. In all cases, both the HHXRF and desktop XRF obtained a lower RSD for the reference material than for the measured lot samples. For Al_2O_3, concentrations were too close to the detection limits, and as the sensitivity of the HHXRF is insufficient at such low concentrations, we observed less variability between measurements. Moreover, Al_2O_3 is a light element that is highly influenced by the quality of the sample surface during in-situ measurements. Desktop XRF results for Fe_2O_3 were slightly more variable than those of HHXRF due to contamination during sample preparation for the desktop approach. The contamination pathways included crushing, pulverizing, and the use of handling
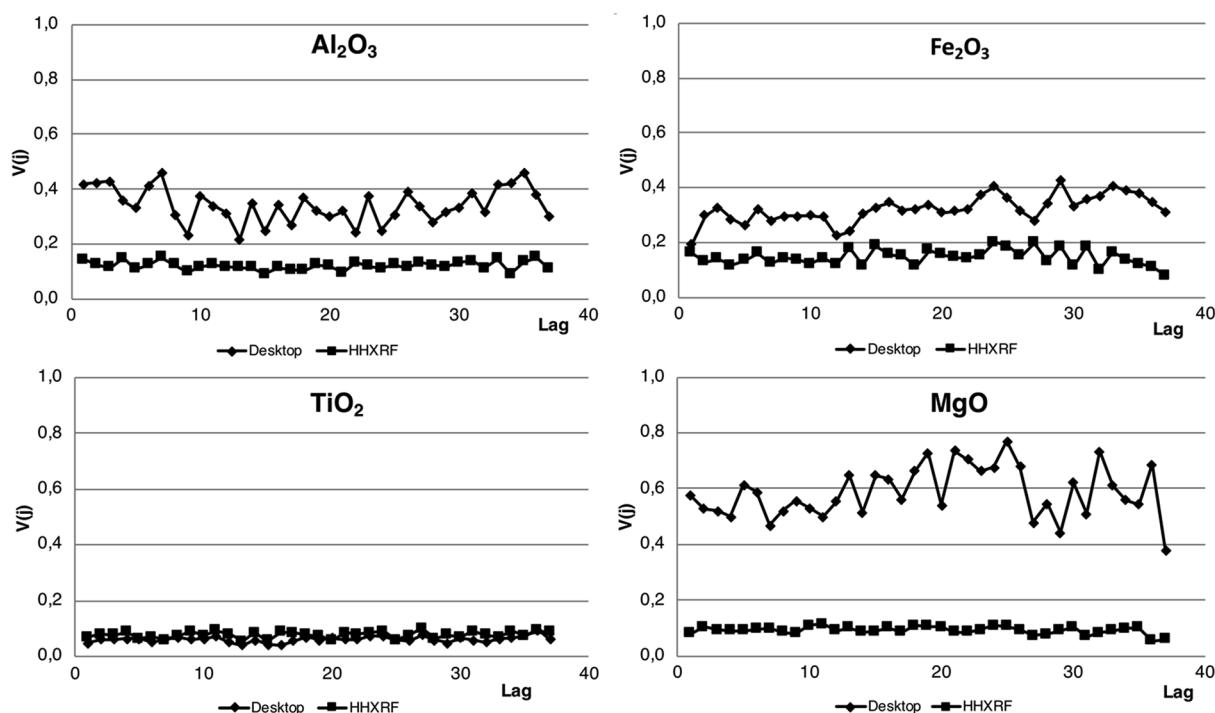

Fig. 6. Sampling variograms comparing desktop XRF and HHXRF results for a large quartzite lot. The lag is 30 cm. Note the lack of significant autocorrelation between samples (no range).
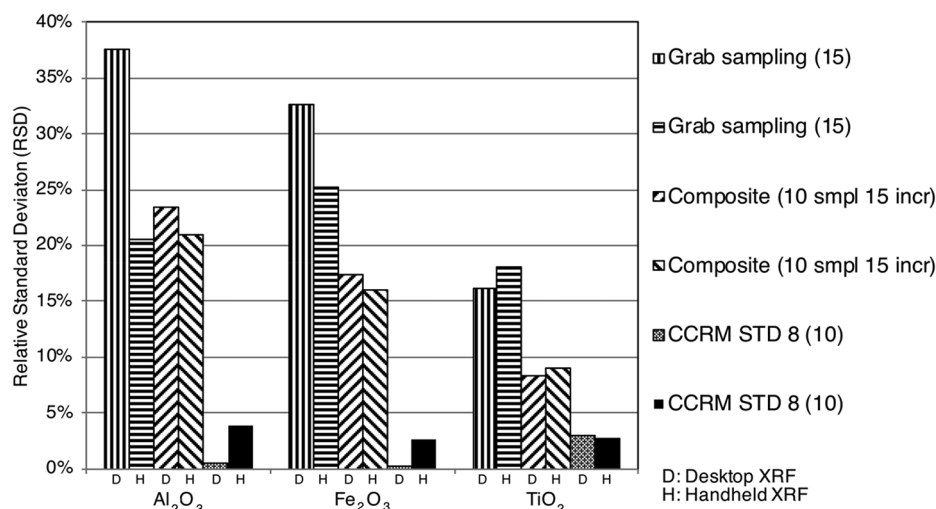
Fig. 7. Analytical performance comparison based on of grab, composite, and reference material samples using in-situ HHXRF (H) and (D) desktop XRF respectively. Desktop XRF was performed on pressed pellets. Both HHXRF and desktop XRF analyzed the reference material as pressed pellets.

iron instruments. At such low concentrations within the quartz, any Fe from an instrument involved in sample preparation can increase the measured concentration significantly. Both the HHXRF and desktop XRF provided comparable results for TiO_2.

3.6. Sampling performance

For Fe_2O_3 and TiO_2, the use of composite samples reduced sampling errors significantly. From an industrial perspective, the most important result is that the composite sample produces less variability than the grab samples. For elements such as Al_2O_3 and Fe_2O_3, the differences between desktop and HHXRF were greater between grab samples than between composite samples. For both analytical approaches, this represents an analytical issue—lower sensitivity for Al_2O_3 and a different sample preparation method for Fe_2O_3. Overall, the use of composite samples reduced the variation between samples by about half, and thus use of HHXRF was deemed adequate for the quality control of quartz within the context of ferrosilicon production.

4. Conclusions

A handheld XRF instrument (HHXRF) used on a basis of deliberate composite sampling provides more accurate sampling + in-situ analysis than the use of conventional grab samples as analyzed by desktop XRF. The lower error of the composite sampling is a consequence of successful reduction of GSE + FSE errors and due to the central limit theorem. Given that the composite sample is the sum of the contribution of each primary unit, measurements generally follow a normal distribution even in cases where the primary samples are not normally distributed. This occurs as the sum of the independent random variables having finite means and variance will gravitate toward a normal distribution (Rohlf et al., 1996).

HHXRF offers the advantage of minimal sample preparation and a fast analytical turnaround. Average lot concentrations can be estimated within an hour, and this approach improves the control over sampling error. However, the application of HHXRF is limited by its lower sensitivity. Analytical results are also easily influenced by traces of minerals found erratically under the XRF beam and by surface quality, this impact being more pronounced for light elements. Thus, a high number of HHXRF scans—five for each block making up a sample bag—are necessary and will, to a certain extent, compensate for these shortcomings. An acceptable HHXRF-based approach based on these empirical results is foreseeable; nonetheless frequent assessments of experimental lot heterogeneity remain necessary. Fortunately, these

assessments are now relatively easy using the HHXRF-based approach. The feasibility of HHXRF has been successfully demonstrated with its intrinsic limitations. The present study outlines a basis for comparing the pro's and con's in the case of a typical industrial lot size, given the specific analytical conditions for HHXRF.

Declaration of Competing Interest

The authors declare no conflict of interest.

References

Bedard, L.P., Barnes, S.-J., 2011. How fit are your data? Geostand. Geoanal. Res. 34, 275–280.

Cohen, D.R., Cohen, E.J., Graham, I.T., Soares, G.G., Hand, S.J., Archer, M., 2017. Geochemical exploration for vertebrate fossils using field portable XRF. J. Geochem. Explor. 181, 1–9.

Collins, B., Napier-Munn, T.J., Sciarone, M., 1974. The production, properties, and selection of ferrosilicon powders for heavy-medium separation. J. South Afr. Inst. Min. Metall. 75, 103–119.

Currie, L.A., 1995. Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations, 1995). Pure Appl. Chem. 67, 1699–1723.

Desroches, D., Bedard, L.P., Lemieux, S., Esbensen, K.H., 2018. Suitability of handheld XRF for quality control of a quartz at a ferrosilicon plant. Miner. Eng (accepted).

de Winter, N.J., Sinnesael, M., Makarona, C., Vansteenberge, S., Claeys, P., 2017. Trace element analyses of carbonates using portable and micro-X-ray fluorescence: performance and optimization of measurement parameters and strategies. J. Anal. Atomic Spectrometry 32, 1211–1223.

DS 3077, 2013. DS 3077. Representative sampling - Horizontal Standard. Danish Standards. www.ds.dk.

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). Chemomet. Intell. Laborat. Syst. 88, 41–59.

Esbensen, K.H., Julius, L.P., 2009. Representative sampling, data quality, validation – a necessary trinity in chemometrics. Comprehen. Chemomet. 4, 1–20.

Esbensen, K.H., Julius, L.P., 2013. DS 3077 Horizontal—a new standard for representative sampling. Design, history and acknowledgements. NIR News 24, 16–19.

Esbensen, K.H., Petersen, L., 2005. Sampling in practise: a TOS toolbox of unit operations. AAUE, Esbjerg.

Esbensen, K.H., Wagner, C., 2014. Theory of sampling (TOS) vs. measurement uncertainty (MU) – a call for integration. Trends Analytical Chemistry (TrAC) 57, 93–106.

Esbensen, K.H., Wagner, C., 2016. Sampling quality assessment: the replication experiment. Spectrosc. Eur. 28, 20–25.

Esbensen, K.H., Wagner, C., Michaels, I., 2018. Entry-level introduction to the theory and practice of sampling - Sampling columns. Spectroscopy Europe/Asia (special issue). Wiley/IMPublications, spectroscopyeurope.com/sampling 82 p.

Engström, K., Esbensen, K.H., 2017a. Sampling column no. 17: Industrial variographic analysis for continuous sampling system validation. Spectrosc. Eur. 29, 14–17.

Engström, K., Esbensen, K.H., 2017b. Evaluation of sampling systems in iron ore concentrating and pelletizing processes – Quantification of Total Sampling Error (TSE) vs. process variation. Miner. Eng. 116, 203–208.

Fisher, L.A., Gazley, M.F., Baensch, A., Barnes, S.J., Cleverley, J., Duclaux, G., 2014. Application of portable X-ray fluorescence analysis to characterize dolerite dykes at the Plutonic Gold Mine, Western Australia. Geochem.: Explor. Environ., Anal. 13, 223–231.

Gazley, M., Fisher, L., 2014. A review of the reliability and validity of portable X-ray fluorescence spectrometry (pXRF) data. Mineral Resource and Ore Reserve Estimation–The AusIMM Guide to Good Practice, pp. 69–82.

Gy, P., 1998. Sampling for analytical purposes. Wiley, New York.

Hall, G.E., Bonham-Carter, G.F., Buchar, A., 2014. Evaluation of portable X-ray fluorescence (pXRF) in exploration and mining: Phase 1, control reference materials. Geochem.: Explor. Environ., Anal. 14, 99–123.

Interactive Corporation, 2014. Brands. http://www.intactive.ru/en/brands/rigaku/Simultix12/?PHPSESSID=hl3m7umofbjhnfrjkl69f87mj6, (accessed 5 December 2018).

Juran, J.M., Godfrey, A., 1998. Juran's quality handbook. McGraw-Hill Professional, New York.

Lemière, B., 2018. A review of pXRF (field portable X-ray fluorescence) applications for applied geochemistry. J. Geochem. Explor. 188, 350–363.

Lemieux, S. 2009. Étude des phénomènes d'appauvrissement en magnésium et strontium dans l'élaboration du ferrosilicium. Unpublished master thesis, Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.

Minnitt, R.C.A., Esbensen, K.H., 2017. Pierre Gy's development of the Theory of Sampling: a retrospective summary with a didactic tutorial on quantitative sampling of one-dimensional lots. TOS Forum 7, 7–19.

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.

Petersen, L., Dahl, C., Esbensen, K.H., 2004. Representative mass reduction in sampling – a critical survey of techniques and hardware. Chemomet. Intell. Laborat. Syst. 74, 95–114.

Pitard, F., 2009. Pierre Gy's Theory of Sampling and C.O. Ingamells' Poisson process approach: pathways to representative sampling and appropriate industrial standards, Doctoral thesis. Aalborg University, Esbjerg Campus, Denmark, pp. 309 p..

Potts, P.J., Webb, P.C., Williams-Thorpe, O., 1997. Investigation of a correction procedure for surface irregularity effects based on scatter peak intensities in the field analysis of geological and archaeological rock samples by portable X-ray fluorescence spectrometry. J. Anal. Atomic Spectromet. 12, 769–776.

Potts, P.J., West, M., 2008. Portable X-ray fluorescence spectrometry: capabilities for in situ analysis. Royal Society of Chemistry, Cambridge, pp. 291.

Quye-Sawyer, J., Vandeginste, V., Johnston, K.J., 2015. Application of handheld energy-dispersive X-ray fluorescence spectrometry to carbonate studies: opportunities and challenges. J. Anal. Atomic Spectromet. 30, 1490–1499.

Ramsey, M.H., Boon, K.A., 2012. Can in situ geochemical measurements be more fit-for-purpose than those made ex situ? Appl. Geochem. 27, 969–976.

Ramsey, M.H., Solomon-Wisdom, G., Argyraki, A., 2013. Evaluation of in situ heterogeneity of elements in solids: implications for analytical geochemistry. Geostand. Geoanal. Res. 37, 379–391.

Rohlf, F.J., Akçakaya, H.R., Ferraro, S.P., 1996. Optimizing composite sampling protocols. Environ. Sci. Technol. 30, 2899–2905.

Ross, P.-S., Bourke, A., Fresia, B., 2014. Improving lithological discrimination in exploration drill-cores using portable X-ray fluorescence measurements: (2) applications to the Zn-Cu Matagami mining camp, Canada. Geochem.: Explor. Environ. Anal. 14, 187–196.

Taylor, P.D., Ramsey, M.H., Boon, K.A., 2007. Estimating and optimising measurement uncertainty in environmental monitoring: An example using six contrasting contaminated land investigations. Geostand. Geoanal. Res. 31, 237–249.

Taylor, P.D., Ramsey, M.H., Potts, P.J., 2005. Spatial contaminant heterogeneity quantification with scale of measurement at contrasting sites. J. Environ. Monit. 7, 1364–1370.

Thermo Fisher Scientific, https://www.thermofisher.com/fr/fr/home/industrial/spectroscopy-elemental-isotope-analysis/portable-analysis-material-id/portable-mining-exploration-solutions/portable-x-ray-fluorescence-xrf-tools-accessories.html, (accessed 7 December, 2018).

Thompson, M., Ellison, S.L., 2006. Fitness for purpose – the integrating theme of the revised harmonised protocol for proficiency testing in analytical chemistry laboratories. Accred. Qual. Assur. 11, 373–378.

Thompson, M., Feam, T., 1996. What exactly is fitness for purpose in analytical measurement? The Analyst 121, 275–278.

Tomé Torquemada, S., 2019. Studies of some thermodynamic aspects related to the production of silicon and ferrosilicon. Ph.D. Thesis, KTH Royal Institute of Technology.

Tuset, J.K.S., 1992. The refining of silicon and ferrosilicon. In: INFACON 6. Proceedings of the 6th International Ferroalloys Congress, Vol. 1, South African Institute of Mining and Metallurgy, Cape Town, pp. 193–199.