



UNIVERSITY OF LEEDS

This is a repository copy of *Statistical quantification of sub-sampling representativeness and uncertainty for waste-derived solid recovered fuel (SRF): Comparison with theory of sampling (ToS)*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/155472/>

Version: Accepted Version

Article:

Gerassimidou, S, Velis, CA orcid.org/0000-0002-1906-726X, Bourne, RA orcid.org/0000-0001-7107-6297 et al. (3 more authors) (2020) Statistical quantification of sub-sampling representativeness and uncertainty for waste-derived solid recovered fuel (SRF): Comparison with theory of sampling (ToS). *Journal of Hazardous Materials*, 388. 122013. ISSN 0304-3894

<https://doi.org/10.1016/j.jhazmat.2019.122013>

© 2019 Published by Elsevier B.V. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

1 **Statistical quantification of sub-sampling representativeness and**
2 **uncertainty for waste-derived solid recovered fuel (SRF):**
3 **Comparison with theory of sampling (ToS)**

4

5 Spyridoula Gerassimidou^a, Costas A. Velis^{a,*}, Richard A. Bourne^b, Dimitrios Komilis^c, Emilio
6 Garcia-Taengua^a, Paul Williams^b

7

8 ^a School of Civil Engineering, University of Leeds, Leeds, LS2 9JT, United Kingdom

9 ^b School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, United
10 Kingdom

11 ^c Department of Environmental Engineering, Democritus University of Thrace, Xanthi,
12 Greece

13

14 *Corresponding author: c.velis@leeds.ac.uk; Telephone: +44 (0) 113 3432327; Room 304,
15 School of Civil Engineering, University of Leeds, Leeds, LS2 9JT, United Kingdom

16

17 Declarations of interest: none

18 **Abstract**

19 The level of uncertainty during quantification of hazardous elements/properties of waste-derived products is
20 affected by sub-sampling. Understanding sources of variability in sub-sampling can lead to more accurate risk
21 quantification and effective compliance statistics. Here, we investigate a sub-sampling scheme for the
22 characterisation of solid recovered fuel (SRF) - an example of an inherently heterogeneous mixture containing
23 hazardous properties. We used statistically designed experiments (DoE) (nested balanced ANOVA) to quantify
24 uncertainty arising from material properties, sub-sampling plan and analysis. This was compared with the
25 theoretically estimated uncertainty via theory of sampling (ToS). The sub-sampling scheme derives
26 representative analytical results for relatively uniformly dispersed properties (moisture, ash, and calorific
27 content: $RSD \leq 6.1\%$). Much higher uncertainty was recorded for the less uniformly dispersed chlorine (Cl)
28 ($RSD: 18.2\%$), but not considerably affecting SRF classification. The ToS formula overestimates the uncertainty
29 from sub-sampling stages without shredding, possibly due to considering uncertainty being proportional to the
30 cube of particle size ($FE \propto d^3$), which may not always apply e.g. for flat waste fragments. The relative
31 contribution of sub-sampling stages to the overall uncertainty differs by property, contrary to what ToS
32 stipulates. Therefore, the ToS approach needs adaptation for quantitative application in sub-sampling of waste-
33 derived materials.

34

35

36 **Keywords**

37 Solid recovered fuel (SRF); Sub-sampling; Uncertainty; Theory of sampling (ToS); Waste

38

39

40

41

42

43

List of abbreviations

ANOVA	Analysis of variance	M_L	Mass of the <i>lot</i> [g]
BS	British standard	M_s	Mass of the sample [g]
BC	Bomb combustion	MSW	Municipal solid waste
C	Sampling constant [g cm ⁻³]	MT	Mechanical treatment
CEN	European Committee Standardization	NCV	Net calorific value [MJ kg ⁻¹]
C&I	Commercial and industrial	PET	Polyethylene terephthalate
Cl	Chlorine	PP	Polypropylene
d	Particle size of sample [cm]	PTE	Potentially toxic elements
d.f.	Degrees of freedom	PVC	Polyvinyl chloride
DoE	Design of experiment	PVdC	Polyvinylidene chloride
EC	European Commission	RSD	Relative standard deviation
ERFO	European Recovered Fuel Organisation	s_{FE}^2	Variance of fundamental error
FE	Fundamental error	SRF	Solid recovered fuel
HDPE	High density polyethylene	ToS	Theory of sampling
HOV	Homogeneity of variance	Total [Cl]	Concentration of total chlorine content
KOH	Potassium hydroxide	WDF	Waste derived fuel
MBT	Mechanical-biological treatment	% w/w _d	Weight concentration on dry basis
MC	Moisture content [% w/w]		

44

45

46

47

48

49

50

51 **Contents**

52 1. Introduction..... 5

53 2. Materials and Methods 7

54 2.1. Materials..... 7

55 2.2. Methodology..... 7

56 2.2.1 Analytical techniques..... 8

57 2.2.2 Balanced nested design for sub-sampling process simulation 8

58 2.2.3 Statistical quantification of sub-sampling uncertainty: nested ANOVA 10

59 2.2.4 Theoretical estimation of sub-sampling uncertainty: ToS-based formula 11

60 3. Results 12

61 3.1. Overall sub-sampling uncertainty 12

62 3.1.1. Descriptive statistics: Classification of SRF sample 12

63 3.1.2. Overall uncertainty: Nested design vs ToS 13

64 3.2. Uncertainty arising from sub-sampling stages 14

65 3.2.1. Statistical significance: Representative sub-samples?..... 14

66 3.2.2. Uncertainty arising from each sub-sampling stage: Nested design vs ToS 16

67 4. Discussion..... 20

68 5. Conclusions..... 23

69 Acknowledgements..... 25

70 Funding 25

71 References 25

72

73

74 **1. Introduction**

75 Major environmental and human health risks are posed by materials present in consumer products in relation to
76 their waste, after-use phase and the processing/ disposal method they may undergo [1]. One of the very core
77 difficulties in controlling / mitigating exposure to such hazards relates to the reliable characterisation of solid
78 waste, because they consist of an extremely heterogeneous mixture [2]. Such inherent heterogeneous composition
79 is attributed to the geographical variation of products most commonly consumed, consumers' buying behaviour
80 and seasonality [3, 4] resulting in difficulty to accurately characterize the material properties [5, 6]. Accurate and
81 precise quantification of risks at a fitness for purpose level, typically involves considerable sample preparation and
82 sub-sampling efforts in the laboratory to reduce the mass of the initial sample (e.g. 1 kg) to just g or mg sample
83 size required for analytical determination [7, 8]. Sub-sampling can be a major source of variability, which needs
84 to be minimised in a way that maintains representativeness, minimises potential bias and enables precise
85 quantification of risk posed by hazardous elements/ substances and related properties [5, 9]. Despite that need, the
86 role of sub-sampling of highly heterogeneous waste is poorly understood in the quantification of uncertainty of
87 analytical measurement [9].

88 Indicatively, solid recovered fuel (SRF) is a waste derived fuel (WDF) typically manufactured from solid waste
89 that meets national and EU specifications for co-combustion applications [10]. Hazardous waste items may be
90 present in the flows processed into SRF (e.g. batteries, paints and small e-waste, such as mobile phones in
91 municipal solid waste) and hazardous chemical elements/ compounds are present or can be released during
92 thermal processing [11], such as dioxins formed from the chlorine (Cl) content of SRF [12].

93 The classification scheme laid down in the BS 15359 [13] specifies SRF properties against three quality criteria,
94 following specific compliance statistics criteria for assessing the risk posed: calorific value, the key economic
95 attribute; Cl content, the key technical attribute; and mercury content, the key environmental performance
96 attribute. Other properties of SRF, such as moisture, ash, biogenic content and potentially toxic elements (PTE)
97 can be included in classification schemes determining WFDs quality [14]. SRF application provides numerous
98 environmental and financial benefits [15-18], although a wider uptake of SRF in industrial applications
99 necessitates predefined SRF quality accurately and precisely determined so that to ascertain efficient utilization
100 [19].

101 However, quantifying and understanding the variability of solid waste properties remains a major challenge for
102 turning waste into secondary resources, and therefore impedes the transition to a circular economy [20].

103 Mechanical processing applied in mechanical-biological treatment (MBT) plants for SRF production results in a
104 relatively homogenized of SRF output [21, 22] in comparison with the plant input, but still SRF remains a highly
105 inherent heterogeneous material, still bearing hazardous components or substances [10, 23]. Thus, any individual
106 set of values obtained from the waste analysis can give invalid results and, therefore, misinformed conclusions
107 [5, 24, 25]. For example, a fragment of polyvinylchloride (PVC) contained in a particular SRF sample received
108 for analysis could lead to the overestimation of Cl content in SRF [19].

109 In order for MBT plants to produce a quality-assured SRF able to meet the needs of end-users, quality assurance
110 measures including sampling, sample treatment, analytical method, choice of quality parameters and data
111 interpretation should be considered [19]. Adherence to a strict sampling protocol is a prerequisite for keeping
112 the uncertainty of analytical results at reasonable levels, especially in waste management [26]. The uncertainty
113 associated with the sampling can exceed the uncertainty associated with the analytical method by an order of
114 magnitude or more in heterogeneous materials, such as SRF [5].

115 Pierre Gy's theory often referred to as 'Theory of Sampling' (ToS), which was developed mainly in the mining
116 industry, provides guiding principles for representative sampling [27]. During representative sampling, the
117 sample collected from a larger body (known as lot) exhibits the average properties of the lot [28, 29]. ToS
118 provides a mathematical formula that estimates the sampling uncertainty only due to the material constitutional
119 heterogeneity [27]. This formula does not consider additional uncertainties related to analytical method, sample
120 preparation and performance of sampling methods. According to Gerlach and Nocerino [5], this formula can be
121 also applied in the sub-sampling process: a repetition of the sampling process applied in the laboratory where a
122 sub-sample is drawn from the sample. However, there is little evidence that verifies this formula experimentally
123 despite the current interest of environmental studies [5].

124 The main operations during sampling and sub-sampling are mixing, mass reduction (extract a small mass from a
125 larger mass) and shredding [30]. Based on ToS, the sample mass should be obtained through composite sampling
126 creating a sample composed of individual material segments [31]. Shredding and mixing are processes for
127 sample homogenization [30]. In addition shredding helps the liberation of the analyte (e.g. Cl in SRF) [5].

128 The first step for the characterization of commercially produced SRF is the sample collection from the MBT
129 plant following the sampling protocol specified in the corresponding CEN standard [32]. The typical mass of
130 SRF sample received for analysis fluctuates between 0.8 and 159 kg depending on the grain size and bulk
131 density, whereas the maximum mass of the lot can reach up to 1500 tonnes [32]. Unless the total mass of the

132 SRF sample as received in the laboratory can be directly analysed, sub-sampling is the next step following the
133 CEN standard related to the sample preparation and sampling methods in the laboratory for SRF [33]. During
134 sub-sampling, the sample is split into sub-samples consecutive times to obtain the test sub-sample, which is
135 ready for analysis [26]. From the test sub-sample, a small mass is drawn for analysis, which is called the test
136 portion [26]. The test portion, with a mass ranging from milligrams to grams, must be sufficiently representative
137 of the heterogeneous SRF lot based on the needs of the application area [7]. The suggested value for a reasonable
138 sub-sampling uncertainty could be considered $< 15\%$ [5], although the absence of relevant comparative evidence
139 does not let us to pose any acceptable limit.

140 Despite the preoccupation of researchers and industry with the variability of SRF properties [10, 34, 35], the
141 uncertainty arising from sampling and sub-sampling processes has not been quantitatively determined yet. Here,
142 we quantified through statistically designed experiments (Design of Experiments: DoE) the relative level and
143 sources of the uncertainty arising from inherent sample properties, sub-sampling scheme and operations, and
144 analytical techniques for SRF characterization and obtain insights on the applicability of ToS in sub-sampling of
145 waste-derived materials.

146

147 **2. Materials and Methods**

148 **2.1. Materials**

149 An SRF sample (ca. 1 kg) typically produced from residual MSW (30% w/w) and commercial and industrial
150 (C&I) waste (70% w/w) in a mechanical treatment (MT) plant in the UK was obtained for the analysis - an
151 example of an inherently heterogeneous mixture containing hazardous properties. Also, we used reagents, such
152 as liquid nitrogen for the cryogenic shredding of the sample, solution of 0.2 N KOH for Cl absorption during
153 combustion in the bomb calorimeter and Palintest acidifying and silver nitrate tablets for the determination of Cl
154 content.

155 **2.2. Methodology**

156 Typical properties of SRF related to economic and technical attributes were selected for the quality
157 determination of SRF. The economic attributes of SRF affect the financial value of the fuel (e.g. increased
158 moisture decreases the heating value of the fuel), whereas technical attributes affect the performance of the

159 combustion facility [14] (e.g. high Cl content can cause corrosion, chlorinated emissions, build-ups and ring
160 formation in cement kilns or high ash content can cause particulate emissions [36-38]).

161 We developed statistically designed experiments to simulate a sub-sampling scheme, in which multiple
162 consecutive steps of sub-sampling and shredding took place for the obtainment and chemical analysis of test
163 sub-samples. Following this sub-sampling plan, analysis of variance (ANOVA) for a balanced nested design
164 approach was used for the quantification of: the overall uncertainty emerging from inherent material
165 heterogeneity, sample preparation (e.g. shredding), sub-sampling practices (e.g. human mistakes and
166 performance of riffle splitting), and analysis (systematic and random error); and, the relative contribution of each
167 step of sub-sampling scheme to the overall uncertainty. This empirically measured sub-sampling uncertainty was
168 compared and contrasted with the theoretical estimation of sub-sampling uncertainty as calculated following the
169 ToS stipulations.

170 **2.2.1 Analytical techniques**

171 We analysed the selected properties following the British standards (BS) stipulations for SRF characterization.
172 Specifically, analytical techniques for the determination of: moisture content ('MC'), expressed in % w/w on a
173 wet basis according to BS 15414-3 [39]; ash content ('Ash'), expressed in % w/w on a dry basis (% w/w_d)
174 according to BS 15403 [8]; net calorific value ('NCV') via bomb combustion (BC), expressed in MJ kg⁻¹ on a
175 dry basis (MJ kg⁻¹_d) according to BS 15400 [40]; and total Cl concentration ('Total [Cl]') through a combined
176 BC and Palintest Chloridol test, expressed in % w/w on a dry basis (% w/w_d) [41, 42].

177 Chlorine in SRF, which is the most critical SRF quality assurance parameter with commercial relevance, is
178 predicted to vary most based on the theoretical calculations of ToS. Hence, we paid special attention to this
179 parameter by measuring the recovery rate of Total [Cl] to exclude the analytical systematic error (bias) from the
180 variance associated with the uncertainty due to sub-sampling [26]. We prepared synthetic mixtures resembling
181 SRF composition with known Total [Cl], which consisted of reference materials such as cellulose, xylans, lignin,
182 HDPE, PP, PET and PVC [43, 44]. The recovery of Total [Cl] during analysis was calculated at 98.3% and
183 consequently the bias of analysis at 1.7%.

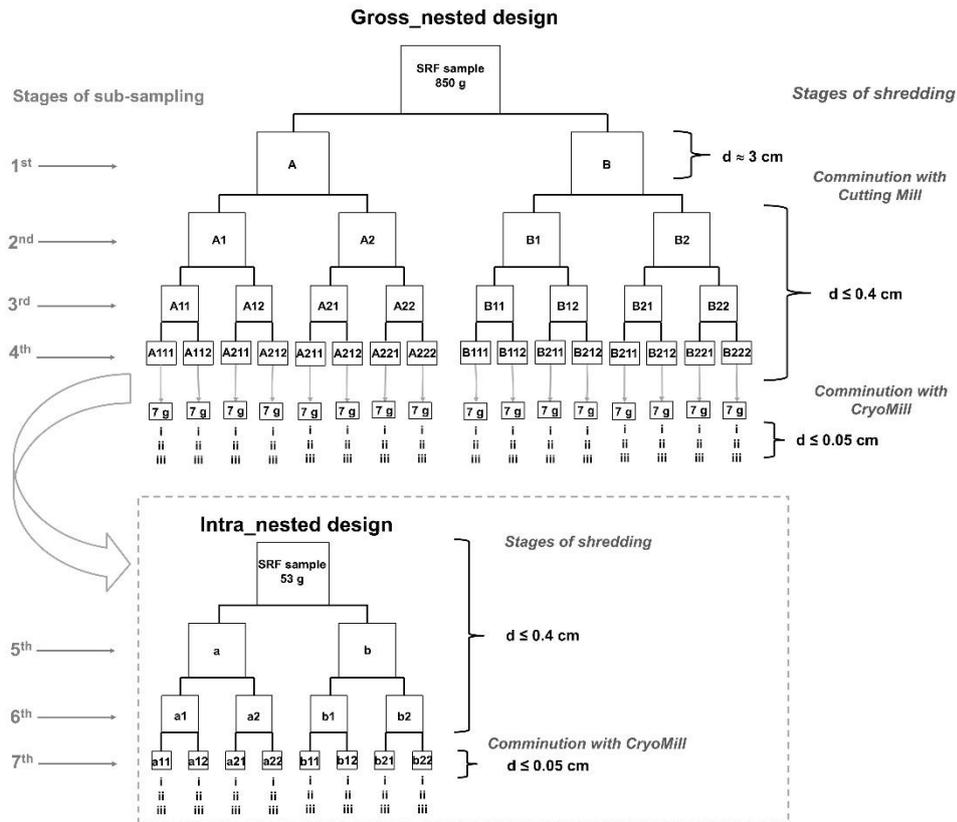
184 **2.2.2 Balanced nested design for sub-sampling process simulation**

185 The sub-sampling process followed for SRF characterization consisted of multiple consecutive stages of riffle
186 splitting (mass reduction) and two stages of shredding: with Cutting Mill (SM 300, Retsch, Germany) and

187 CryoMill (Retsch, Germany). After pre-drying of the SRF sample at 40 °C for 24 h to remove the moisture that
188 could interfere with the shredding process [39], the sample mass was reduced to ~~at~~ ca. 900 g. The pre-dried
189 sample with a particle size (d) ca. 3 cm was split into two equal sub-samples with riffle splitters. These sub-
190 samples were shredded with a Cutting Mill to $d \leq 4$ mm and split again many times ~~so that~~ to obtain the test sub-
191 samples. The test sub-samples were shredded by CryoMill to $d_{90} = 0.15$ mm only for the determination of Total
192 [CI] and NCV. Finally, three test portions from each test sub-sample were taken for analysis. The balanced
193 nested design simulates the stages of the sub-sampling process. Each stage of sub-sampling indicates the
194 operation of riffle splitting, which creates nested sub-samples within the sample. For example, in the 1st stage of
195 sub-sampling, the sample was split into sub-samples A and B (nested within the sample), in the 2nd stage sub-
196 samples A and B were shredded and split into sub-samples A1 and A2 (nested within A) and B1 and B2 (nested
197 within B), respectively, etc. Sub-samples are nested because they depend only on the sample (or higher level
198 sub-sample) from which they came from [45].

199 Specifically, we developed two balanced nested designs: *Gross_nested* to simulate the sub-sampling process
200 from the 1st stage to the 4th stage of sub-sampling (four-level balanced nested design); and *Intra_nested* to
201 simulate the sub-sampling process from the 5th stage to the 7th stage of sub-sampling (three-level balanced nested
202 design) (**Figure 1.**). The pre-dried sample was thoroughly mixed and a sub-sample of ca. 50 g extracted from
203 the sample of ca. 900 g. In *Gross_nested* design, the SRF sample with a mass 850 g was split 4 times to obtain
204 16 sub-samples of 53 g. Thereafter a test sub-sample of 6 – 7 g was taken from each sub-sample of 53 g with
205 riffle splitters and 3 replicates were taken for analysis. *Gross_nested* design examined the ability of a sub-sample
206 of 53 g to represent the initial SRF sample of 850 g (sub-sampling uncertainty from 1st to 4th stage). Similarly, in
207 *Intra_nested* design, the SRF sub-sample with a mass 53 g was split 3 times to obtain 8 test sub-samples of 6 – 7
208 g. *Intra_nested* design examined the ability of the test sub-sample of 6 – 7 g to represent the SRF sub-sample of
209 53g (sub-sampling uncertainty from 5th to 7th stage).

210 We broke the nested design into two complementary designs due to the small capacity of CryoMill (50 ml),
211 which cannot fill more than 6 – 7 g of SRF. It would be considerably expensive and time-consuming to
212 accurately divide 900 g into all possible test sub-samples of 6 – 7 g. This would correspond to more than 135 test
213 sub-samples. The decision to break the nested design at the 4th stage was taken after checking the minimum
214 number of test sub-samples (see **Supporting information (SI) 1**).



215

216 **Figure 1.** Four-level balanced nested design for the calculation of the uncertainty arising from the 1st to the 4th
 217 sub-sampling stage (*Gross_nested*: from the sample mass of nearly 850 g to the sub-sample of 53 g) and three-
 218 level balanced nested design for the calculation of the uncertainty arising from the 5th to the 7th sub-sampling
 219 stage (*Intra_nested*: from the sub-sample of 53 g to the sub-sample of 6 – 7 g). In every sub-sampling stage, sub-
 220 samples depend only on the lower level of nested sub-samples. *Gross_nested* and *Intra_nested* design consists of
 221 16 and 8 test sub-samples, respectively. Three replicates for each test sub-sample were taken to capture the
 222 uncertainty arising from analysis.

223

224 2.2.3 Statistical quantification of sub-sampling uncertainty: nested ANOVA

225 We statistically quantified the overall uncertainty emerging from sample constitutional heterogeneity, sub-
 226 sampling process and analysis and the uncertainty arising from each sub-sampling stage. Nested ANOVA at a
 227 significance level 0.05 was conducted in TIBCO Statistica™ 13.3.0 software for both designs: four-level nested
 228 ANOVA for *Gross_nested* and three-level nested ANOVA for *Intra_nested* design. Each level of nested-
 229 ANOVA referred to each stage of sub-sampling and compared the mean values between sub-samples nested
 230 within the sample or higher level of sub-sample. Significance results of nested-ANOVA detected the statistical
 231 difference at each stage and post hoc tests were carried out to detect the sources of the difference (pairs in which

232 nested sub-samples are significantly different to each other) [46]. Key assumptions required to use ANOVA
233 were examined (see **SI.2**). Randomization to reduce bias was applied by analysing sub-samples by chance rather
234 than by choice [26].

235 In every stage of sub-sampling, the mean values of sub-samples derived from the mean values of the lower level
236 of nested sub-samples capturing the variance introduced at each level were determined, known as variance
237 components. The relative standard deviation (RSD) of these variances provided evidence for the uncertainty
238 arising from each stage of sub-sampling process. The RSD of the sum of the variance components corresponded
239 to the overall uncertainty. The variance components given by nested ANOVA exhibited the contribution of each
240 sub-sampling stage to the overall uncertainty [47], which was calculated by dividing the RSD from each stage with
241 the overall RSD. This information gave insights on which stages affected mostly the overall sub-sampling
242 uncertainty.

243 In addition, we used descriptive statistics to summarize the analytical results of the SRF properties obtained
244 through the established sub-sampling process (arithmetic mean, median, standard deviation, 95% confidence
245 interval and range). Descriptive statistics derived from the average values of 17 sub-samples: 16 sub-samples
246 analyzed in *Gross_nested* design and 1 sub-sample analyzed in *Intra_nested*. We did not obtain the individual
247 measurements as the assumption of independence would be violated. Replicate measurements from a test sub-
248 sample are more related to each other than the measurements from different test sub-samples. At this stage, we
249 also included the NCV on wet basis so that to characterize the SRF sample according to the classification system
250 [13]. However, the sub-sampling uncertainty was not quantified for the properties on a wet basis as it would
251 include the uncertainty for moisture, which was obtained separately.

252 **2.2.4 Theoretical estimation of sub-sampling uncertainty: ToS-based formula**

253 ToS provides a mathematical formula (**Eq. 1**) that calculates the sub-sampling uncertainty only due to the
254 constitutional heterogeneity of the material, known as *fundamental error* (FE) [27]. This formula can be used
255 before the sub-sampling process to gain insights on the minimum sub-sampling uncertainty as it does not include
256 uncertainties related to analysis, sample preparation and sub-sampling methods [5].

$$s_{FE}^2 = \left(\frac{1}{M_S} - \frac{1}{M_L} \right) * C * d^3 \quad \text{Eq. 1}$$

257 where M_s is the sample mass [g], M_L is the mass of the *lot* [g], C is the sampling constant related to the
258 characteristics of the sample [g cm^{-3}] and d is the nominal size of the particles [cm]. In case of sub-sampling, the

259 M_s becomes the mass of the new sub-sample and M_L becomes the mass of sample or higher level of sub-sample
260 [5].

261 According to ToS, the square root of **Eq. 1** gives the RSD of FE. We used this formula to calculate the minimum
262 sub-sampling uncertainty for the selected SRF properties based on ToS. However, results only for MC, Ash and
263 Total [CI] are included since ToS considers the analyte (e.g SRF property) as a contaminant, whereas NCV is
264 not. Specifically, the concentration and density of the analyte are required for the calculation of sampling
265 constant, C , while NCV cannot be expressed neither as a fraction in the *lot* nor as an ingredient with density.

266 The calculation of C relied on typical values of physical characteristics for a fluff type SRF given by BS 15442
267 [32]. However, the value of C changes after shredding. The precise quantification of C after shredding was not
268 obtained as a significant amount of information for the target material, that is not available here, is required [7].
269 We estimated these values (see **SI.3**) based on the optical observation of the physical characteristics of sub-
270 samples after each shredding stage and on typical values given by ToS depending on the physical characteristics
271 of the sample [27].

272

273 **3. Results**

274 **3.1. Overall sub-sampling uncertainty**

275 **3.1.1. Descriptive statistics: Classification of SRF sample**

276 Descriptive statistics that summarizes the analytical results of SRF properties were obtained (**Table 1.**).

277 **Table 1.** Central tendency, spread and confidence intervals of SRF properties derived from 17 sub-samples of 53
278 g obtained through the established sub-sampling process.

Descriptive statistics*	\bar{x}	M	s.d.	95% CI	Range
MC (% w/w)	16.62	16.59	0.26	(16.48, 16.76)	[16.05, 17.07]
Ash (% w/w _d)	13.45	13.54	0.39	(13.26, 13.65)	[12.81, 14.19]
Total [CI] (% w/w _d)**	1.09	1.07	0.15	(1.02, 1.18)	[0.90, 1.40]
NCV (MJ kg _d ⁻¹)	24.05	24.03	0.65	(23.72, 24.38)	[22.59, 25.22]
NCV (MJ kg _d ⁻¹)***	20.05	20.08	0.51	(19.79, 20.32)	[18.90, 21.00]

*Descriptive statistics derived from 17 SRF sub-samples including: arithmetic mean, median, standard deviation, 95% confidence interval and range between minimum and maximum values, respectively;
** Descriptive statistics only for Total [Cl] were calculated after the removal of systematic analytical error (1.7%); ***NCV expressed on wet basis (MJ kg^{-1}) in order to compare the values with the classification system [13]

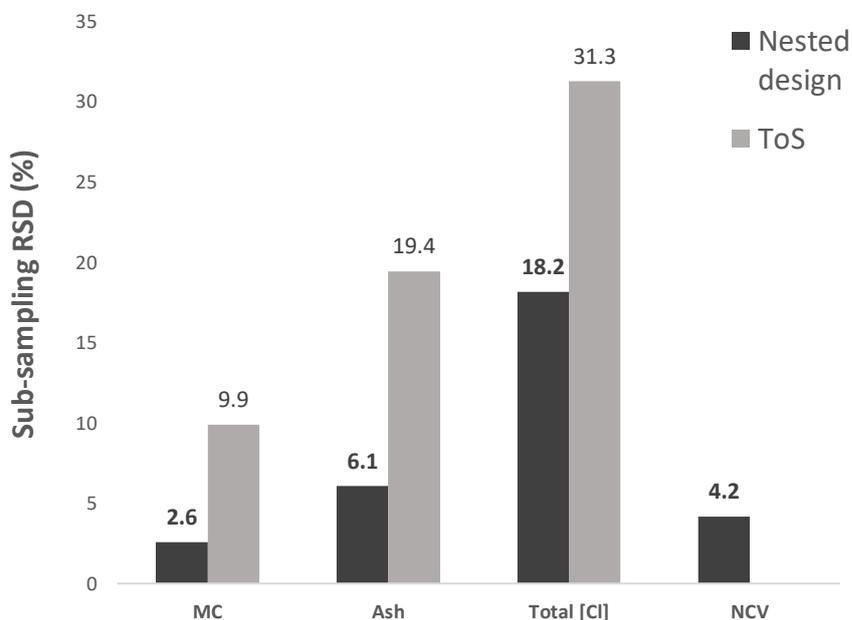
279 Total [Cl] lies within a range of 1.02 - 1.18% w/w_d with 95% confidence, which designates the commercially
280 produced sample as class code 4 for Total [Cl] (1 – 1.5% w/w_d) based on the specification requirements laid
281 down in the BS 15359 [13] standard. The NCV lies within 19.79 and 20.32 MJ kg^{-1} with 95% confidence
282 interval. This range contains the borderline between class code 3 ($\geq 15 \text{ MJ kg}^{-1}$) and class code 2 ($\geq 20 \text{ MJ kg}^{-1}$),
283 but the average value (20.05 MJ kg^{-1}) designates the SRF sample as class code 2 for NCV. Based on the
284 classification scheme for the quality of WDF, MC (economic quality parameter) specifies the sample as class
285 code 3 ($\leq 20\%$ w/w), whereas Ash (technical quality parameter) defines the sample as class code 2 ($\leq 20\%$ w/w_d)
286 [14].

287 3.1.2. Overall uncertainty: Nested design vs ToS

288 **Figure 2** presents the overall sub-sampling uncertainty for key SRF properties as statistically calculated (nested
289 design) and theoretically estimated (ToS) (see **SI.4**). Both approaches, nested design and ToS, are in agreement
290 regarding the dependence of sub-sampling uncertainty on SRF property. Based on ToS, the lower the
291 concentration of analyte, the higher the sampling uncertainty [27]. The highest uncertainty is introduced for the
292 determination of Total [Cl] due to its lower fraction in the sample (average 1.09% w/w_d) compared to the other
293 SRF properties, whereas the lowest uncertainty found in MC determination (average 16.62% w/w). The
294 uncertainty for NCV was calculated based only on nested design as ToS considers that analytes are contaminants
295 (see section **2.2.4**)

296 The statistical approach was expected to give higher RSD than the theoretical as ToS-based formula calculates
297 the RSD only due to constitutional heterogeneity of the material [27], whereas the RSD from nested-ANOVA
298 includes also all the related factors that may introduce uncertainty, such as shredding, experimenter skills,
299 performance of sub-sampling methods and analytical errors. However, **Figure 2** is opposed to this expectation.
300 For example, RSD estimated by ToS is more than 3 times higher compared to nested design for the
301 determination of MC (2.6%) and Ash (6.1%). This difference was less than twice for Total [Cl] as the RSD was
302 statistically determined at 18.2% and theoretically estimated at 31.3%. However, the precise quantification of
303 RSD based on ToS was not obtained due to the insufficient amount of information required for the precise

304 calculation of sampling constant, C (see section 2.2.4). Thus, this difference might be attributed to the
 305 overestimation of C, unless the relative contribution of sub-sampling stages to the uncertainty is considerably
 306 different between nested design and ToS (see section 0).



307
 308 **Figure 2.** Overall uncertainty, expressed as RSD (%), arising from the established sub-sampling plan for the
 309 determination of SRF properties statistically (nested design) and theoretically (ToS) determined: NCV was
 310 calculated only statistically due to the consideration of ToS that analytes are contaminants.

311

312 3.2. Uncertainty arising from sub-sampling stages

313 3.2.1. Statistical significance: Representative sub-samples?

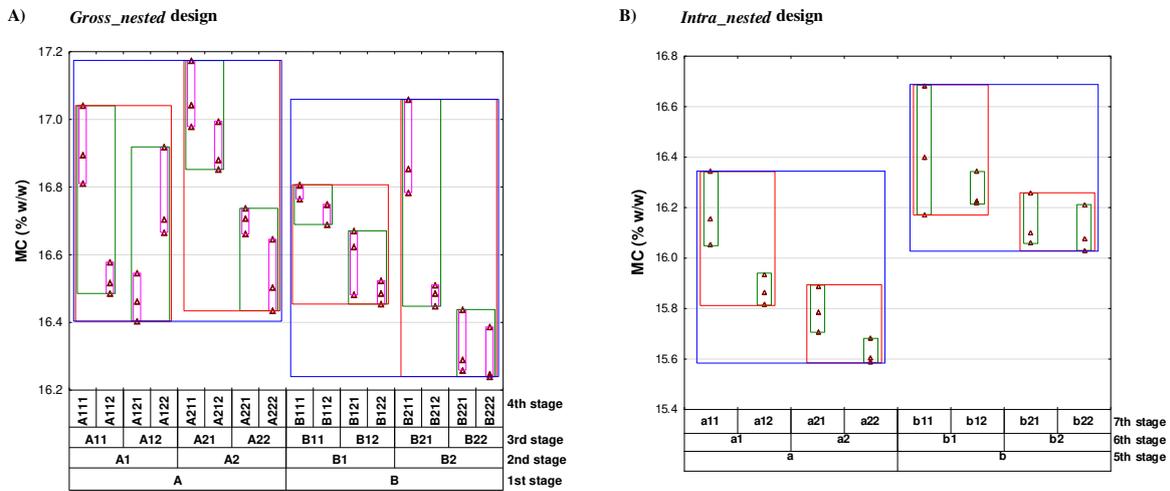
314 The significance test of nested-ANOVA showed that there is at least one pair of nested sub-samples, which are
 315 statistically different to each other only in the last stages of both designs: 4th (*Gross_nested*) and 7th
 316 (*Intra_nested*) (see SI.5). This could evidence that the lower sample mass, the higher possibility of not getting
 317 representative sub-samples, which is also confirmed by ToS (Eq. 1). Despite that, the number of pairwise
 318 comparisons between sub-samples nested within the sample is larger in the last stages than in the first stages
 319 increasing the sensitivity to reject the null hypothesis. For example, the null hypothesis at the 5th stage, which is
 320 the first stage of *Intra_nested* design, considered that the mean value of analyte in sub-sample *a* is equal with
 321 that in sub-sample *b*. Still, the null hypothesis at the 7th stage (or final stage of *Intra_nested* design) was that: the
 322 mean value in sub-sample *a11* is equal with the mean value in sub-sample *a12*; and *a21* equal with *a22*; and *b11*

323 equal with $b12$; and $b21$ equal with $b22$ (**Figure 1.**). Post hoc tests showed that there is statistical difference
 324 between sub-samples nested within samples for almost every pair not only in the 4th and 7th sub-sampling stages
 325 (see **SI.5**). However, the disagreement between ANOVA significance test and post hoc test might presage a false
 326 alarm of significant difference [46].

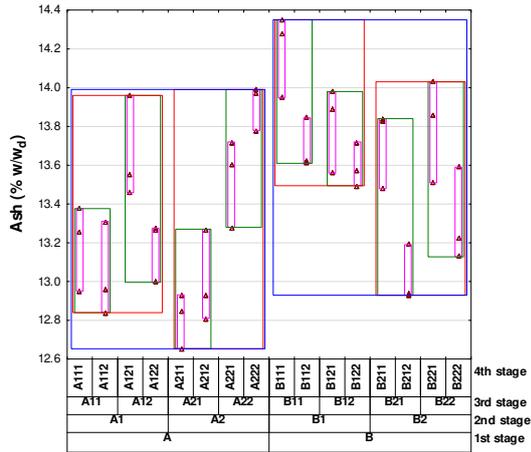
327 In order to get better insights in to the difference between sub-samples, variability plots demonstrating the spread
 328 of values between nested sub-samples were obtained (Figure 3.). The blue parallelograms represent the sub-
 329 samples of the first stage of riffle splitting for both designs, the red parallelograms nested within the blue ones
 330 indicate the sub-samples of the second stage, etc. In most pairs of nested sub-samples, the spread of values looks
 331 quite alike. In the 2nd stage, the sub-samples (red parallelograms of *Gross_nested* design) look more uniform
 332 pairwise compared to the other stages due to shredding with the Cutting Mill applied before the 2nd stage. In the
 333 7th stage, the nested sub-samples present a higher spread of values (green parallelograms of *Intra_nested* design)
 334 due to their low mass.

335

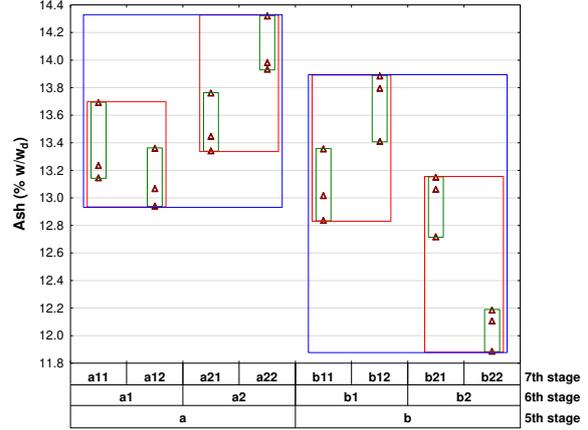
336



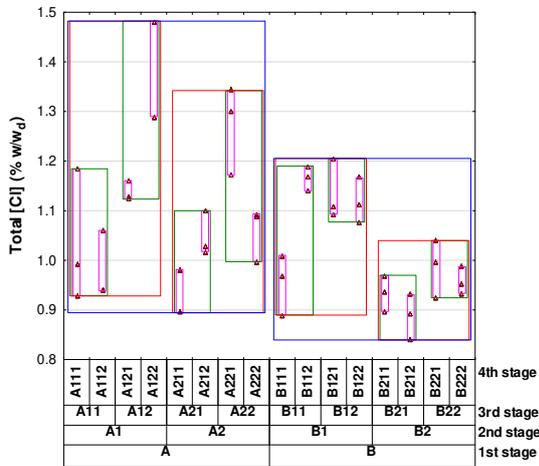
C) *Gross_nested* design



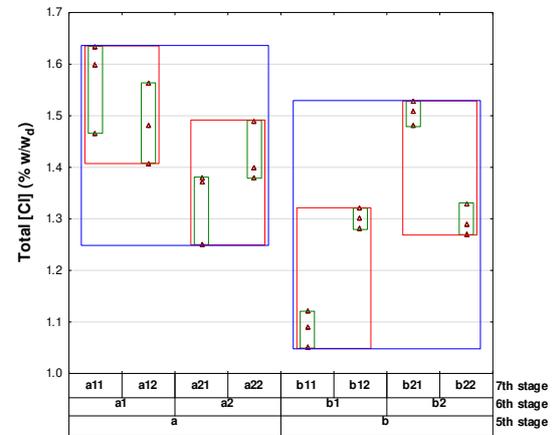
D) *Intra_nested* design



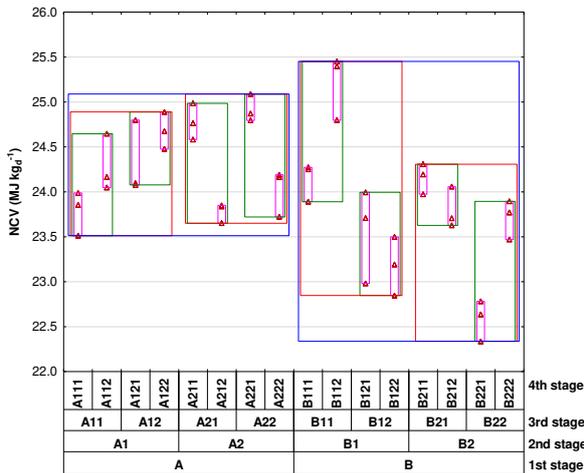
E) *Gross_nested* design



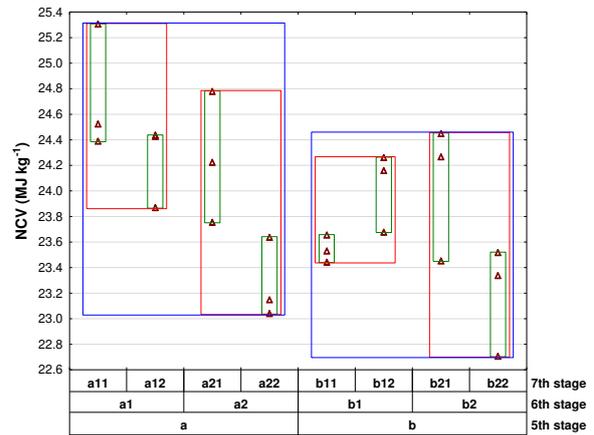
F) *Intra_nested* design



G) *Gross_nested* design



H) *Intra_nested* design



337 **Figure 3.** Variability plots of SRF properties between sub-samples nested within sample at each stage of sub-
 338 sampling process. A: Variability plot of MC from the 1st to the 4th sampling stage (*Gross_nested*); B: Variability
 339 plot of MC from the 5th to the 7th sampling stage (*Intra_nested*); C: Variability plot of Ash from the 1st to the 4th
 340 sampling stage (*Gross_nested*); D: Variability plot of Ash from the 5th to the 7th sampling stage (*Intra_nested*);
 341 E: Variability plot of Total [Cl] from the 1st to the 4th sampling stage (*Gross_nested*); F: Variability plot of Total

342 [CI] from the 5th to the 7th sampling stage (*Intra_nested*); G: Variability plot of NCV from the 1st to the 4th
 343 sampling stage (*Gross_nested*); H: Variability plot of NCV from the 5th to the 7th sampling stage (*Intra_nested*).

344 3.2.2. Uncertainty arising from each sub-sampling stage: Nested design vs ToS

345 The table of the variance components (see **SI.6**) provided by nested-ANOVA used for the calculation of RSD
 346 arising from each sub-sampling stage are shown in **Table 2**. For all fuel properties, the RSD presents an upward
 347 trend as the sub-sampling process unfolds, which verifies the statement of ToS that the lower the sample mass,
 348 the higher the uncertainty. However, the RSD arising from the 2nd stage of sub-sampling is zero for all properties
 349 except for Total [CI] that remains almost constant (2.2%). This drop is attributed to the shredding process with
 350 the Cutting Mill applied after the 1st and before the 2nd stage of riffle splitting revealing the beneficial role of
 351 shredding in the reduction of the uncertainty.

352 Likewise, the RSD arising from the 4th stage is higher than the 5th stage for most SRF properties (except to MC),
 353 which is opposed to the tendency of RSD to be increased as the sample mass is reduced. This is attributed to the
 354 breakdown of nested design in *Gross_nested* (1st – 4th stage) and *Intra_nested* (5th – 7th stage). From the
 355 statistical point of view, balanced nested designs have higher statistical power to the later levels than preceding
 356 levels due to more degrees of freedom (d.f.) [48]. Here, the 5th stage which corresponds to the first level of
 357 *Intra_nested* has only one d.f., whereas the 4th stage which corresponds to the final stage of *Gross_nested* has
 358 eight d.f. (see **SI.6**).

359 **Table 2.** Sub-sampling uncertainty, expressed as RSD (%), arising from each stage of the established sub-
 360 sampling plan for the determination of key SRF properties statistically calculated (nested design).

Sub-sampling stages		RSD (%) [*]			
		MC	Ash	Total [CI]	NCV
1 st	From 850 g to 425 g	0.5	1.6	2.3	1.4
2 nd	From 425 g to 212 g	0.0	0.0	2.2	0.0
3 rd	From 212 g to 106 g	1.0	1.8	7.8	1.8
4 th	From 160 g to 53 g	1.0	2.1	8.5	2.2
5 th	From 53 g to 26 g	1.4	1.3	4.5	0.6
6 th	From 26 g to 13 g	1.0	3.1	7.2	0.2
7 th	From 13 g to 6.5 g	0.7	3.2	7.9	2.0
‘Error’ ^{**}	From 6.5 g to test portion	0.7	1.6	5.0	1.4

^{*}Relative standard deviation derived from the ratio of variance components at each sub-sampling stage given by nested-ANOVA to the arithmetic mean of SRF properties in 16 test sub-samples for *Gross_nested* and 8 test sub-samples for *Intra_nested* design; ^{**} RSD of ‘Error’ stage derives from

the variance of replicates including the analytical error and the sub-sampling uncertainty arising from test sub-sample to test portion

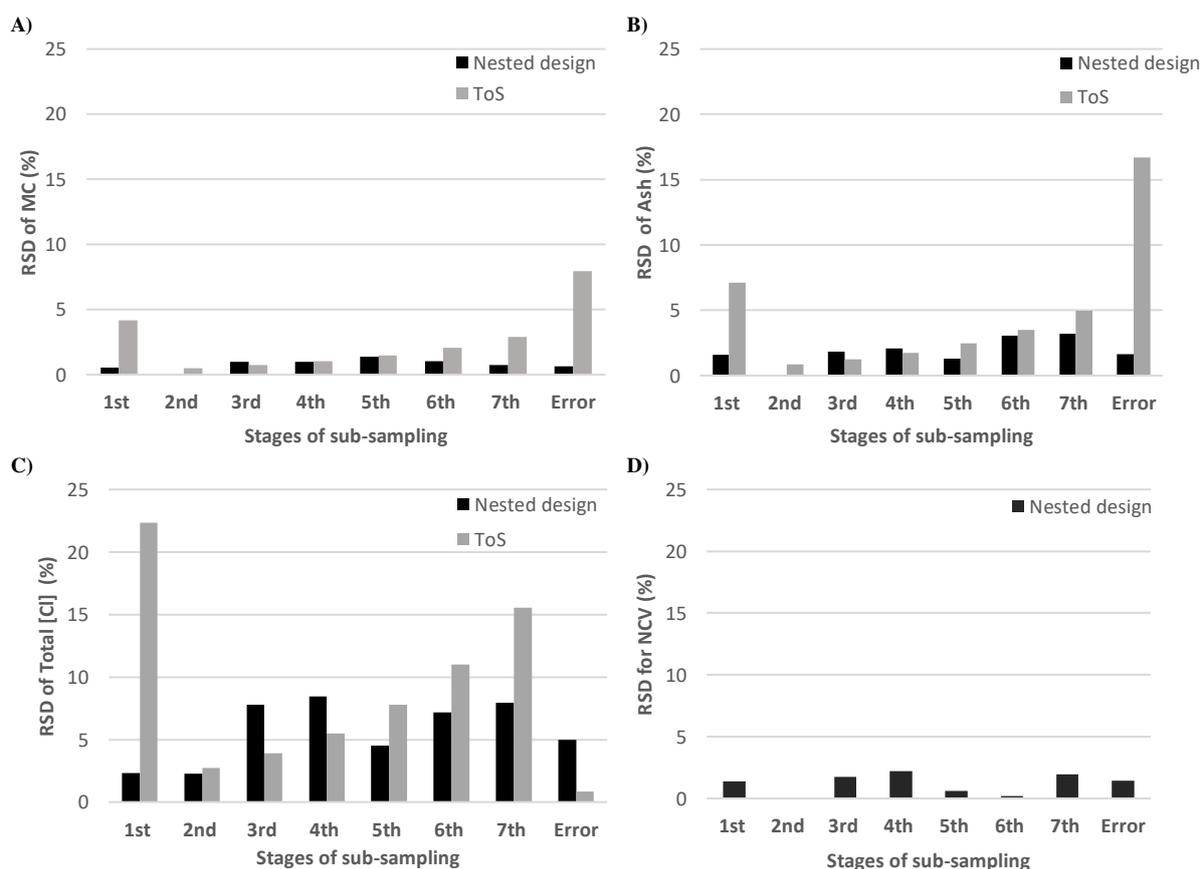
361 Despite the lower sensitivity of nested-ANOVA to detect variance in upper levels, the comparison of RSD
362 between the sub-sampling stages with equal d.f. (e.g. 1st with 5th, 2nd with 6th, and 3rd to 7th) demonstrates the
363 upward trend of RSD as the sample mass decreases. An exception is observed in the case of NCV as the RSD
364 introduced during the 1st stage (1.4%) is more than twice higher than the 5th stage (0.6%). The trend of RSD
365 during the sub-sampling is not always predictable. For example, in case of MC, the RSD tends to decrease from
366 the 5th to the 7th stage. This trend was also observed for the determination of NCV, the RSD arising from the 6th
367 stage (0.2%) is 3 times lower compared to the 5th stage (0.6%).

368 The RSD of ‘Error’ stage derived from the variance of replicate measurements is attributed to: sub-sampling
369 from the test sub-sample to test portion; and analytical technique. The RSD arising from the ‘Error’ stage ranges
370 at acceptable levels (< 5%). In most cases, it is lower than the RSD arising from the sub-sampling stages, which
371 confirms the statement that the sub-sampling uncertainty may considerably exceeds the analytical error in highly
372 heterogeneous materials [5].

373 We compared the statistically calculated RSD (nested design) with the theoretically estimated RSD (ToS) arising
374 from each sub-sampling stage (**Figure 4.**). ToS provides considerably higher RSD compared to nested design in
375 stages that shredding was not applied: in the 1st stage for all SRF properties (mass reduction with initial $d \approx 3$ cm);
376 and in the ‘Error’ stage for MC and Ash (collection of test portions without cryogenic shredding). In the
377 intermediate stages, from the 2nd to the 6th, the results between nested design and ToS seem more compatible.
378 The difference of RSD between the 1st and 2nd stage is higher in the theoretical approach rather than the
379 statistical approach. For example, the RSD arising from the 1st stage is at least 10 times higher than the 2nd stage
380 of sub-sampling for the determination of Total [Cl] based on ToS, while nested design demonstrates that the
381 RSD arising from the 1st (2.3%) and 2nd stage (2.2%) are similar. Shredding with the Cutting Mill reduced the
382 uncertainty as the RSD did not change for lower sample mass, but this reduction is not as high as the estimated
383 one by ToS.

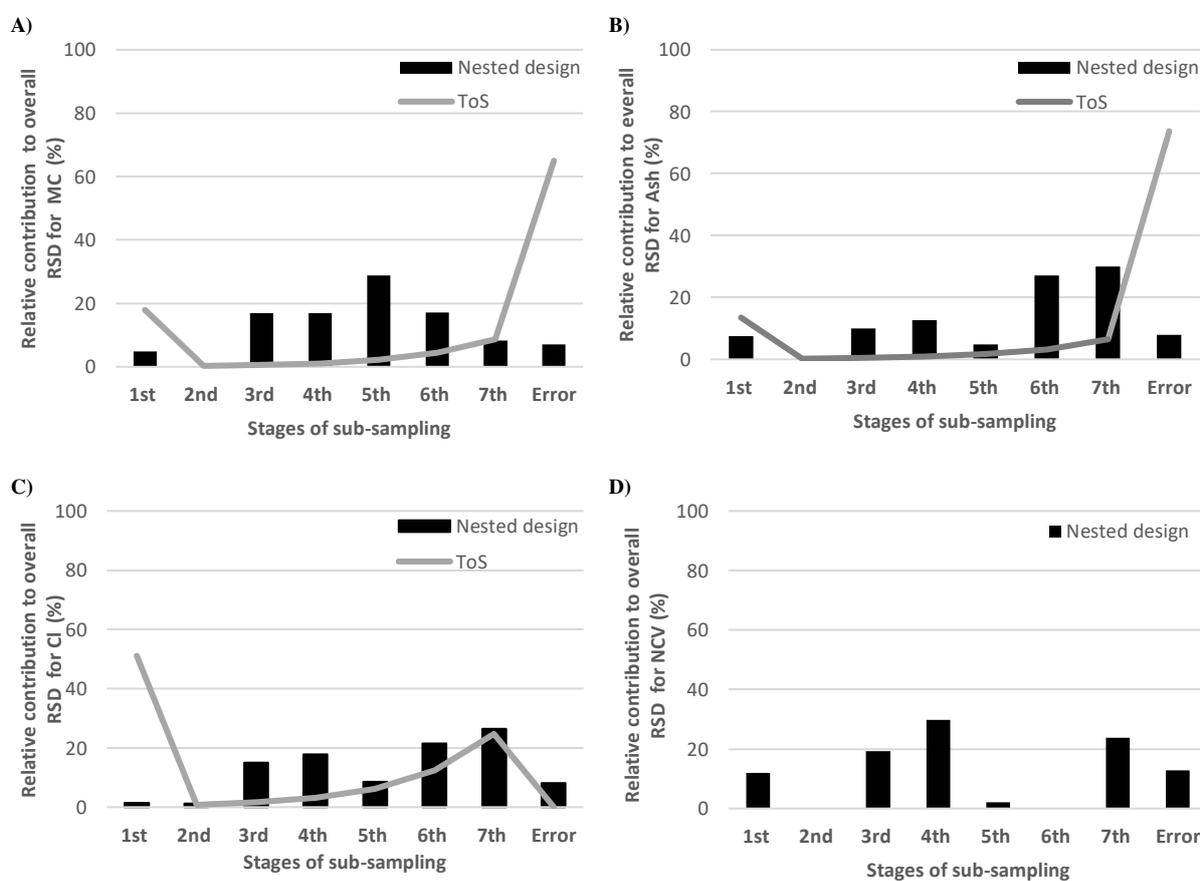
384 In the stage of ‘Error’ (from test sub-sample to test portion) ToS finds that the RSD for MC and Ash is more
385 than 10 times higher in comparison with the nested design. Here, the d.f. of nested design are sufficient in order
386 to concern about the statistical power of the design. Furthermore, from the perspective of ToS the sampling
387 constant, C, was not changed in the final stages (no shredding) for MC and Ash, whilst the difference between
388 the two approaches became considerably wider compared to the previous stage (7th). Thus, the high difference

389 between the two approaches in the ‘Error’ stage for MC and Ash is not related neither to the lower sensitivity of
 390 balanced nested design in the upper stages nor to the inaccurate calculation of sampling constant. The role of
 391 particle size in the sub-sampling uncertainty seems less important than ToS formula performs. This is also
 392 confirmed by the fact that only for Total [Cl] the RSD arising from the ‘Error’ stage, where cryogenic shredding
 393 was applied, calculated by nested design is more than 6 times higher than the RSD based on ToS. It is the only
 394 stage that the RSD theoretically estimated (ToS) is considerably lower than the statistically calculated RSD.
 395 The relative contribution of each stage of sub-sampling to the overall RSD based on both approaches is
 396 presented (**Figure 5.**). According to ToS, for the determination of Total [Cl] the relative contribution of the sub-
 397 sampling stages to the overall uncertainty is formed in an order $1^{st} > 7^{th} > 6^{th} > 5^{th} > 4^{th} > 3^{rd} > 2^{nd} > \text{‘Error’}$. For
 398 MC and Ash the order is slightly different due to absence of shredding the test sub-samples with CryoMill
 399 increasing the uncertainty emerging from the final ‘Error’ stage: $\text{‘Error’} > 1^{st} > 7^{th} > 6^{th} > 5^{th} > 4^{th} > 3^{rd} > 2^{nd}$.



400 **Figure 4.** Sub-sampling uncertainty, expressed as RSD (%), arising from each sub-sampling stage for the
 401 determination of key SRF properties calculated statistically (nested design) and theoretically (ToS): NCV was
 402 calculated only statistically due to the consideration of ToS that analytes are contaminants.

403 According to nested design, the relative contribution of each stage to the overall RSD is different for each SRF
 404 property. For example, in case of MC the order is 5th > 6th > 3rd > 4th > 7th > "Error" > 1st > 2nd, for Ash is 7th > 6th
 405 > 4th > 3rd > 'Error' > 1st > 5th > 2nd, for Total [Cl] is 7th > 6th > 4th > 3rd > 5th > 'Error' > 1st > 2nd and for NCV is 4th
 406 > 7th > 3rd > 'Error' > 1st > 5th > 6th > 5th > 2nd. Based ToS, the order of the relative contribution of the sub-
 407 sampling stages to the overall uncertainty should be the same for all analytes under an identical sub-sampling
 408 plan, but nested design showed that the relative contribution depends on fuel property. However, the 2nd stage, in
 409 which sub-samples have the highest sub-sample mass (circa 450 g) with the lowest particle size (< 0.4 cm)
 410 compared to the other sub-sampling stages, provides the lowest relative contribution for all SRF property based
 411 on both approaches.



A) MC; B) Ash; C) Cl; D) NCV only based on nested design; The stage of 'Error' refers to the RSD of replicates

412 **Figure 5.** Relative contribution of sub-sampling stages to the overall sub-sampling uncertainty (RSD) for the
 413 determination of SRF properties based on the statistical (nested design) and theoretical approach (ToS): NCV
 414 was calculated only statistically due to the consideration of ToS that analytes are contaminants.

415

416 4. Discussion

417 The results have shown the merits and demerits of both approaches applied for the quantification of sub-
418 sampling uncertainty, ToS and nested design. The ToS formula considers the physical and chemical
419 characteristics of the sample as constant factors into the equation, but these factors might be considerably
420 different between sub-samples depending on the size and composition of every particle contained in the sub-
421 samples [49]. For example, particles with variable sizes in the sample may have a different chemical
422 composition or particles with different composition may belong to a specific particle size and/or the analyte
423 might be abundant in some particles and confined in others [7]. Thus, the target material tends to behave more
424 unpredictably rather than predictably as ToS implies [7]. This behaviour was observed by the results of nested
425 design: the RSD for MC was slightly increased as the sub-sample mass reduced in *Gross_nested* design, but the
426 reverse behaviour observed in *Intra_nested*; and, the relative contribution of the sub-sampling stages to the
427 overall uncertainty was different for each SRF property despite the identical sub-sampling process applied in
428 SRF properties. Unless the target material has a predictable physical and chemical constitution, such as a narrow
429 range of particle size and uniform dispersion of analyte into the sample, ToS formula can be inaccurate [7, 49].
430 Besides, the effort needed to accurately quantify the determinants of sampling factor would exceed the effort to
431 quantify empirically the sampling uncertainty due to the significant amount of information required for the target
432 material [7].

433 The results of nested design revealed the lower sensitivity of a balanced nested design to detect the variance at
434 upper levels, where the d.f. are fewer. But, the construction of an unbalanced nested design in order to create a
435 better balance of d.f. amongst the stages would reduce the power of ANOVA test [50]. Balanced nested designs
436 are orthogonal and the estimators of the variance components are independent resulting in higher precision of
437 estimates than unbalanced designs, which are more sensitive to the assumptions of homoscedasticity and
438 normality [51, 52]. Unless the entire sample mass is analysed, we cannot quantify accurately and precisely the
439 sub-sampling uncertainty. Here, the variance components derived from 72 analytical measurements (48 in
440 *Gross_nested* and 24 in *Intra_nested* design). In the case of NCV and Total [Cl], this amount of measurements
441 includes the analytical determination of 2.4% w/w of total sample mass (test portion 0.3 g x 72 measurements =
442 21.6 g out of 900 g) and in the case of MC and Ash this percentage is 8% w/w (test portion 1 g x 72
443 measurements = 72 g out of 900 g). However, a prudent use of ANOVA under a strict sub-sampling protocol can
444 give valuable results [19, 53]. The prerequisites for using ANOVA were thoroughly checked (see **SI.2**),
445 experiment randomization was applied and advanced sub-sampling practices and equipment were used in order
446 to produce trustworthy results [5, 47, 53]. In the process of turning waste materials into secondary resources,

447 statistically designed experiments (DoE) are valuable tools able to quantify the variability, although there is
448 limited application in the waste processing sector [20].

449 In comparison with the nested design, the ToS approach gave a substantially higher sub-sampling uncertainty
450 (overall RSD) and higher relative contribution in the sub-sampling stages, where either the particle size was
451 large (1st stage) or the test sub-samples were not cryogenically shredded ('Error' stage for MC and Ash). These
452 findings indicate that the particle size of the sample might affect less the sub-sampling uncertainty than ToS
453 stipulates. Most of the particles in solid waste tend to be either flat (i.e. paper and textiles) or hollow (e.g.
454 containers), thus their thickness does not change with shredding so much as in granular materials, in which ToS
455 is mainly applied. ToS-based formula needs re-evaluation and possibly adjustment regarding the proportional
456 relationship between FE and the cube of particle size ($d^3 \propto FE$) before applied in the sub-sampling for waste-
457 derived materials. Therefore, the role of ToS is to provide guidance on correct sampling practices and equipment
458 so that to adopt a sampling plan able to minimize the uncertainty [5, 7].

459 Based on the results of nested design, the overall uncertainty for the determination of relatively uniformly
460 dispersed analytes in solid waste, such as MC, Ash and NCV can be considered acceptable ($RSD \leq 6.1\%$).
461 However, the uncertainty for the less uniformly dispersed Total [Cl] ($RSD: 18.2\%$) exceeded the 15% limit
462 suggested by Gerlach and Nocerino [5]. Chlorine in SRF varies widely between SRF components and its
463 variability has preoccupied researchers more than any other SRF property [35, 36]. For example, specific plastic
464 polymers, such as PVC and PVdC are highly chlorinated materials with Total [Cl] ranging from 46 – 73%,
465 whereas Cl is absent in PP and HDPE.

466 The uncertainty states the difference between the experimentally identified estimate (measured value) and the
467 'true' value (also known as real population value). Here, RSD indicates that the Total [Cl] in a test sub-sample
468 can be up to 0.20% w/w_d below or above the average Total [Cl] (1.09% w/w_d), which represents the 'true' value,
469 still an unknown quantity, in the SRF sample. Insights on the difference between the 'true' value and the
470 estimate of average of Total [Cl] derived from 17 sub-samples (16 in *Gross_nested* design and 1 in *Intra_nested*)
471 were obtained by the margin of error, which was 0.08% w/w_d at 95% confidence level. Putting this into
472 perspective, the overall RSD did not exceed the intervals ranges used in the class codes of classification CEN
473 SRF scheme, which can be from 0.4% w/w_d to 1.5% w/w_d (class code 1 $\leq 0.2\%$ w/w_d; class code 2 $\leq 0.6\%$ w/w_d;
474 class code 3 $\leq 1\%$ w/w_d; class code 4 $\leq 1.5\%$ w/w_d; and, class code 5 $\leq 3\%$ w/w_d).

475 Higher levels of uncertainty can in fact be tolerable and fit for purpose, depending on the user needs, inherent
476 material heterogeneity and the variability of analyte in the sample (e.g uniformly dispersed analyte such as the
477 chemical substance in a drug or less uniformly dispersed analyte such as PTE in solid waste) [7, 54], because
478 representative sampling might be never fully achieved [55]. Taking into account the highly inherent
479 heterogeneous composition of SRF and the considerable CI variability in waste component categories, a sub-
480 sampling plan able to obtain values that lie within a permissible range for the classification of SRF can fulfil the
481 fitness for purpose requirements. Besides, end-users have to acknowledge that there is always the possibility of
482 incorrect classification of SRF associated with CI even under a thorough sub-sampling protocol, but this
483 possibility can be confined and controlled with the use of appropriate practices and equipment. However, in case
484 of less uniformly dispersed hazardous properties (e.g PTE) the analysis of duplicate samples for a single lot
485 might be needed [33].

486 The analytical determination of analytes in SRF and generally in solid waste must include the sampling
487 uncertainty [19, 55]. Specifically, compliance evaluation with existing quality management specifications
488 requires the incorporation of the uncertainty level of the measurand at a selected set limit leading to the creation
489 of acceptance and rejection zones [56]. In most studies the properties of solid waste are expressed as individual
490 values and the standard deviation of replicates is given. However, the variability of replicates, known as
491 repeatability, refers to the precision (closeness of measurements to each other) and not to the accuracy
492 (proximity to the true value) of analytical results [26]. Besides, we found that the RSD arising from the 'Error'
493 stage which describes the variability of replicates, constitutes only a small part of the sub-sampling uncertainty
494 ranging from 7 to 14% of total RSD for the selected SRF properties. The properties of solid waste need to be
495 specified by a level of uncertainty using: quality control practices, such as reference materials for the calculation
496 of systematic errors [26], replicate measurements [26], randomization [47], correct sampling practices and
497 equipment [5]; and statistical tools [10, 20].

498

499 **5. Conclusions**

500 Through statistically designed experiments, we tested here for the representativeness of sub-sampling (laboratory
501 sampling, sample preparation and analytical determination) of solid waste samples, which are subject to great
502 inherent variability, not least due to their composition. Testing was performed under optimally practicable

503 correct sampling practices and equipment. Via quantification of sub-sampling uncertainty (RSD), we established
504 that it is overall feasible to obtain sufficiently representative analytical results for certain key fuel properties in
505 SRF (moisture, ash and calorific value: $RSD \leq 6.1\%$), which are more uniformly dispersed compared to PTE
506 such as cadmium and mercury based on BS 15443:2011.

507 However, for the determination of Total [Cl], which is a less uniformly dispersed property and the key limiting
508 factor of SRF end-uses, the sub-sampling uncertainty is far from negligible (RSD: 18.2%). Whereas this level of
509 uncertainty is three times higher than the other key properties, it is just above the 15% limit suggested as
510 acceptable in the literature – although an arbitrary limit in the absence of relevant comparative evidence. The
511 outcome for Cl indicates that we could possibly have to live with that level of uncertainty when we quantify Cl
512 in SRF and similar properties in waste-derived samples. This conclusion considers as fitness for purpose criteria
513 the fact that: (i) SRF is inherently a highly heterogeneous material; (ii) Cl varies widely between SRF
514 components (e.g. different types of plastic); and (iii) what is practicable for compliance statistics as established
515 in the relevant SRF classification standard. Indeed, an RSD at 18.2% indicates that the measured Total [Cl] in a
516 test sub-sample can range from 0.89 to 1.29% w/w_d, which corresponds to a concentration difference of 0.20%
517 w/w_d from the average value (1.09% w/w_d). Putting this into perspective, the intervals between the class codes of
518 the classification scheme of SRF for Cl are defined with ranges of 0.4 – 1.5% w/w_d. Arguably, lowering the sub-
519 sampling uncertainty of the Cl with the current technological state of the art would possibly require excessive
520 effort and cost.

521 We also provide here a quantification of the relative sources of the sub-sampling and analytical determination
522 process. Nested design confirmed the statement of ToS that sub-sampling uncertainty can significantly exceed
523 the uncertainty associated with the analytical method in highly heterogeneous materials. The uncertainty
524 introduced at the final stage ('Error' stage: from the test sub-sample to test portion) constituted only 7 – 14% of
525 total sub-sampling uncertainty. This is just around 1/10th of the overall uncertainty: most of the uncertainty is
526 introduced in the preceding sub-sampling stages.

527 Our work offers tangible insights on the applicability of the ToS in the context of waste samples. The nested
528 design (statistical approach) indeed confirmed the ToS (theoretical approach) with respect to the increase of
529 uncertainty as the sample mass decreases (higher RSD at final stages) and the dependence of sub-sampling
530 uncertainty on the concentration of analytes (highest RSD found in Total [Cl] with the lowest fraction in SRF -
531 lowest uncertainty found in MC with the highest fraction).

532 We also conclude here that the ToS-based formula needs to be re-evaluated and possibly adjusted for
533 applications in sub-sampling of waste-derived materials. First, the relative contribution of sub-sampling stages to
534 the overall uncertainty was found to differ depending on fuel property, contrary to what the ToS stipulates.
535 Second, comparison of nested design approach with the ToS-based formula reveals that the latter overestimates
536 the uncertainty emerging from stages without shredding (1st and 7th stage). We suggest that this overestimation
537 could be attributed to the ToS uncertainty formula being proportional to the cube of particle size ($FE \propto d^3$),
538 which may not be universally applicable to all waste item fragments. Shredding decreases the thickness of flat or
539 hollow particles mainly included in SRF to a lower degree compared to granular materials, for which the ToS
540 formula was developed. The work presented here can set the basis for the introduction of statistically informed
541 sub-sampling standards in SRF and wider solid waste samples, and enable the informed revision of existing
542 technical standards that apply ToS for waste-derived fuels, such as the BS 15442.

543

544

545 **Acknowledgements**

546 We are grateful to the lab technician personal at the Solid Waste Management and Sample Preparation
547 laboratory at the University of Leeds, for training and generic quality assurance support.

548

549 **Funding**

550 This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-
551 profit sectors. Doctoral researcher S. Gerassimidou was funded by the University of Leeds Scholarships
552 programme.

553

554 **References**

555 [1] Z.A. Coventry, R. Tize, A.T. Karunanithi, Comparative life cycle assessment of solid waste management
556 strategies, *Clean Technol. Environ. Policy*, 18 (2016) 1515-1524.

- 557 [2] C. Riber, I. Rodushkin, H. Spliid, T.H. Christensen, Method for fractional solid-waste sampling and chemical
558 analysis, *Int. J. of Environ. An. Ch.*, 87 (2007) 321-335.
- 559 [3] J. Gajdoš Kljusuric, J. Čačić, A. Misir, D. Čačić, Geographical region as a factor influencing consumers'
560 perception of functional food—case of Croatia, *Brit. Food J.*, 117 (2015) 1017-1031.
- 561 [4] M. Beckmann, M. Pohl, D. Bernhardt, K. Gebauer, Criteria for solid recovered fuels as a substitute for fossil
562 fuels—a review, *Waste Manage. Res.*, 30 (2012) 354-369.
- 563 [5] R.W. Gerlach, J.M. Nocerino, Guidance for obtaining representative laboratory analytical subsamples from
564 particulate laboratory samples, US Environmental Protection Agency, Washington, DC, 2003.
- 565 [6] M.E. Edjabou, M.B. Jensen, R. Götze, K. Pivnenko, C. Petersen, C. Scheutz, T.F. Astrup, Municipal solid
566 waste composition: Sampling methodology, statistical analyses, and case study evaluation, *Waste Manage.*, 36
567 (2015) 12-23.
- 568 [7] M.H. Ramsey, M. Thompson, Uncertainty from sampling, in the context of fitness for purpose, *Accred. Qual.*
569 *Assur.*, 12 (2007) 503-513.
- 570 [8] BS 15403, Solid recovered fuels - Determination of ash content, British Standard Institution, 2011.
- 571 [9] F. Risso, R. Narizzano, M. Ottonelli, A. Magherini, M. Giampieri, E. Carlini, Validation of subsampling
572 procedures and granulometry distribution uncertainty assessment in environmental soil analysis, *Analytical*
573 *Methods*, 4 (2012) 570-574.
- 574 [10] C. Velis, P.J. Longhurst, G.H. Drew, R. Smith, S.J. Pollard, Production and quality assurance of solid
575 recovered fuels using mechanical — biological treatment (MBT) of waste: a comprehensive assessment, *Crit.*
576 *Rev. Environ. Sci. Technol.*, 40 (2010) 979-1105.
- 577 [11] Environmental Protection Agency, European waste catalogue and hazardous waste list, Environmental
578 Protection Agency Ireland, County Wexford, Ireland, 2002.
- 579 [12] J.A. Conesa, L. Rey, S. Egea, M.D. Rey, Pollutant formation and emissions from cement kiln stack using a
580 solid recovered fuel from municipal solid waste, *Environ. Sci. Technol.*, 45 (2011) 5878-5884.
- 581 [13] BS 15359, Solid recovered fuels - Specifications and classes, British Standard Institution, 2011.
- 582 [14] WRAP, A classification scheme to define the quality of waste derived fuels, Waste & Resources Action
583 Programme, 2012.

- 584 [15] R. Sarc, K.E. Lorber, R. Pomberger, M. Rogetzer, E.M. Sipple, Design, quality, and quality assurance of
585 solid recovered fuels for the substitution of fossil feedstock in the cement industry, *Waste Manage. Res.*, 32
586 (2014) 565-585.
- 587 [16] A. Garg, R. Smith, D. Hill, N. Simms, S. Pollard, Wastes as co-fuels: the policy framework for solid
588 recovered fuel (SRF) in Europe, with UK implications, *Environ. Sci. Technol.*, 41 (2007) 4868-4874.
- 589 [17] A.A. Uson, A.M. López-Sabirón, G. Ferreira, E.L. Sastresa, Uses of alternative fuels and raw materials in
590 the cement industry as sustainable waste management options, *Renew Sust. Energ. Rev.*, 23 (2013) 242-260.
- 591 [18] E. Iacovidou, J. Hahladakis, I. Deans, C. Velis, P. Purnell, Technical properties of biomass and solid
592 recovered fuel (SRF) co-fired with coal: impact on multi-dimensional resource recovery value, *Waste Manage.*,
593 73 (2018) 535-545.
- 594 [19] S. Flamme, J. Ceiping, Quality assurance of solid recovered fuels (SRF), *ZKG Int.*, 67 (2014) 54-57.
- 595 [20] K.H. Esbensen, C. Velis, Transition to circular economy requires reliable statistical quantification and
596 control of uncertainty and variability in waste, *Waste Manage. Res.*, 34 (2016).
- 597 [21] C. Velis, S. Wagland, P. Longhurst, B. Robson, K. Sinfield, S. Wise, S. Pollard, Solid recovered fuel:
598 influence of waste stream composition and processing on chlorine content and fuel quality, *Environ. Sci.*
599 *Technol.*, 46 (2012) 1923-1931.
- 600 [22] Y.-H. Chang, W.C. Chen, N.-B. Chang, Comparative evaluation of RDF and MSW incineration, *J. Hazard.*
601 *Mater.*, 58 (1998) 33-45.
- 602 [23] M. Kallassy, B. Efremenko, M. Champel, Waste processing: the status of mechanical and biological
603 treatment, in: ISWA Beacon conference "The global challenge: optimising the cycle of biological treatment of
604 biowaste", Perugia, Italy, 2008.
- 605 [24] G.L. Rupp, R.R. Jones, *Characterizing Heterogenous Wastes: Methods and Recommendations*, United
606 States Department of Commerce, National Technical Information Services, 1992.
- 607 [25] C. Wagner, K.H. Esbensen, Theory of sampling: Four critical success factors before analysis, *J. AOAC Int.*,
608 98 (2015) 275-281.
- 609 [26] E. Prichard, V. Barwick, *Quality assurance in analytical chemistry*, John Wiley & Sons, 2007.

610 [27] F.F. Pitard, Pierre Gy's sampling theory and sampling practice: heterogeneity, sampling correctness, and
611 statistical process control, CRC press, 1993.

612 [28] P. Gy, Sampling of particulate materials theory and practice, 1st ed., Elsevier Scientific Publishing
613 Company, Netherlands 2012.

614 [29] B. Kratochvil, J.K. Taylor, Sampling for chemical analysis, *Anal. Chem.*, 53 (1981) 924A-938A.

615 [30] L. Petersen, P. Minkkinen, K.H. Esbensen, Representative sampling for reliable data analysis: theory of
616 sampling, *Chemom. Intell. Lab. Syst.*, 77 (2005) 261-277.

617 [31] H. Horn, J. Dibdiakova, R.S. Aanerød, A. Vestlund, K.H. Esbensen, Method for representative biomass
618 sampling to improve its qualitative parameters and to optimize combustion processes, in: *European Biomass
619 Conference and Exhibition Proceedings*, 2018, pp. 1588-1592.

620 [32] BS 15442, Solid recovered fuels - Methods for sampling, British Standards Institution, 2011.

621 [33] BS 15443, Solid recovered fuels - Methods for the preparation of the laboratory sample, British Standard
622 Institution, 2011.

623 [34] J. Cuperus, E. van Dijk, R. de Boer, Pre-normative research on SRF, TAUW, Deventer, Netherlands, 2005.

624 [35] I. Deans, I. Dimas, C.A. Velis, Modelling of solid recovered fuel (SRF) properties based on material
625 composition – chloride quality, *Waste Manage.*, 6 (2016) 389-398.

626 [36] W. Ma, G. Hoffmann, M. Schirmer, G. Chen, V.S. Rotter, Chlorine characterization and thermal behavior in
627 MSW and RDF, *J. Hazard. Mater.*, 178 (2010) 489-498.

628 [37] R. Jin, J. Zhan, G. Liu, Y. Zhao, M. Zheng, Variations and factors that influence the formation of
629 polychlorinated naphthalenes in cement kilns co-processing solid waste, *J. Hazard. Mater.*, 315 (2016) 117-125.

630 [38] M.d.M. Cortada Mut, L.K. Nørskov, F.J. Frandsen, P. Glarborg, K. Dam-Johansen, Circulation of inorganic
631 elements in combustion of alternative fuels in cement plants, *Energy & Fuels*, 29 (2015) 4076-4099.

632 [39] BS 15414-3, Solid recovered fuels - Determination of moisture content using the oven dry method.
633 Moisture in general analysis sample, British Standards Institution, 2011.

634 [40] BS 15400, Solid recovered fuels - Determination of calorific value, British Standards Institution, 2011.

635 [41] BS 15408, Solid recovered fuels - Methods for the determination of sulphur (S), chlorine (Cl), fluorine (F)
636 and bromine (Br) content, British Standards Institution, 2011.

637 [42] Palintest-Test, Chloride (Chloridol)-Photometer method, 2019.
638 https://www.palintest.com/application/files/6914/5510/6190/Chloride_Chloridol_INST_En.pdf (Accessed:
639 24.04.2019).

640 [43] J. Heikkinen, J.d. Hordijk, W. de Jong, H. Spliethoff, Thermogravimetry as a tool to classify waste
641 components to be used for energy generation, *J. Anal. Appl. Pyrol.*, 71 (2004) 883-900.

642 [44] J. Cuperus, E. van Dijk, R. de Boer, Pre-normative research on SRF, TAUW, Deventer, Netherlands, (2005)
643 128.

644 [45] M. Krzywinski, N. Altman, P. Blainey, Nested designs, *Nat. Methods*, 11 (2014) 977.

645 [46] C. Tian, X. Manfei, T. Justin, W. Hongyue, N. Xiaohui, Relationship between Omnibus and Post-hoc Tests:
646 An Investigation of performance of the F test in ANOVA, *Shanghai Arch. Psychiatry*, 30 (2018) 60.

647 [47] D.C. Montgomery, *Design and analysis of experiments*, John wiley & sons, 2017.

648 [48] Y. Ojima, Generalized staggered nested designs for variance components estimation, *J. Appl. Stat.*, 27
649 (2000) 541-553.

650 [49] D. Francois-Bongarcon, P. Gy, The most common error in applying Gy's Formula'in the theory of mineral
651 sampling, and the history of the liberation factor, *J. South. Afr. Inst. Min. Metall.*, 102 (2002) 475-479.

652 [50] R.G. Shaw, T. Mitchell-Olds, ANOVA for unbalanced data: an overview, *Ecology*, 74 (1993) 1638-1645.

653 [51] C. Fernandes, P. Ramos, J. Mexia, Crossing balanced and stair nested designs, *Electron. J. Linear. Al.*, 25
654 (2012) 22-47.

655 [52] H. Schielzeth, S. Nakagawa, Nested by design: model fitting and interpretation in a mixed model era,
656 *Methods Ecol. Evol.*, 4 (2013) 14-24.

657 [53] A.N. Analytical Methods Committee, Quality control of routine sampling in chemical analysis, *Anal.*
658 *Methods*, 4 (2012) 1482-1483.

659 [54] R. Bethem, J. Boison, J. Gale, D. Heller, S. Lehotay, J. Loo, S. Musser, P. Price, S. Stein, Establishing the
660 fitness for purpose of mass spectrometric methods, *J. Am. Soc. Mass. Spectrom.*, 14 (2003) 528-541.

661 [55] M.H. Ramsey, Appropriate rather than representative sampling, based on acceptable levels of uncertainty,
662 in: *Measurement Uncertainty in Chemical Analysis*, Springer, 2002, pp. 163-169.

663 [56] C. Velis, Solid recovered fuel production through the mechanical-biological treatment of wastes, in: School
664 of Applied Sciences, Cranfield University, United Kingdom, 2010.

665