

An improved method for soil aggregate stability measurement using laser granulometry applied at regional scale

B. G. RAWLINS, G. TURNER, J. WRAGG, P. McLACHLAN & R. M. LARK

British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

Running heading: *Soil aggregate stability at regional scale*

Correspondence: B. G. Rawlins. E-mail: bgr@bgs.ac.uk

Summary

Laboratory-based aggregate stability (AS) tests should be applied to material wetted to a moisture content comparable with field soil. We have improved our original laser granulometer (LG)-based AS test published in this journal (Rawlins *et al.*, 2013) by including a pre-wetting stage. Our method estimates disaggregation reduction (DR; μm) for a soil sample (1–2 mm diameter aggregates). Soils with more stable aggregates have larger DR values. We apply the new technique to soils from 60 cultivated sites across eastern England, with ten samples from each of six different parent material (PM) types encompassing a wide range of soil organic carbon (SOC) concentrations (1.2–7.0%). There are large differences between the median DR values (rescaled to $<500 \mu\text{m}$) for soils over the PM types, which when used as a predictor (in combination with SOC concentration) accounted for 53% of the variation in DR. There was no evidence for including an interaction term between PM class and SOC concentration for the prediction of DR. After applying the aggregate stability tests using the sixty regional soil samples, they were stored for nine months and the tests were repeated resulting in a small but statistically significant increase in DR for samples from some, but not all, PM types. We show how a palaeosol excavated from a site in southern England can be used as an aggregate reference material (RM) to monitor the reproducibility of our technique. It has been suggested that soil quality, measured by critical soil physical properties, may decline if the organic carbon concentration is below a critical threshold. Our results show that, for aggregate stability, any such thresholds are specific to the PM.

23 Introduction

24 Aggregate stability (AS) is an important physical indicator of soil quality because it
25 influences a range of soil functions and degradation threats including water availability
26 to plant roots and runoff through its influence on pore size distribution (Zhao *et al.*,
27 2007), erodibility (Farres, 1987; Barthes & Roose, 2002) and reductions in oxygen
28 diffusion related to surface crusts formed by slaking (Rathore *et al.*, 1982). Aggregate
29 stability is of particular interest in cultivated soils. There is evidence that aggregates
30 are less stable under conventional tillage than under no-tillage systems (Haynes &
31 Knight, 1989) and this has been shown to lead to differences in the distribution of
32 organic matter in the soil profile. Because of its implications for soil management
33 aggregate stability may be a useful indicator for monitoring soil quality, although
34 there are several candidate indicators (Rickson *et al.*, 2013). To be applied at national
35 scales, methods for measuring the stability of soil aggregates must be relatively rapid
36 and reproducible.

37 In a recent study, Rawlins *et al.* (2013) presented and applied a novel method
38 to measure AS using a laser granulometer (LG) which overcomes some of the limita-
39 tions of earlier approaches based on wet-sieving. The advantages of the LG method
40 include: i) normalising the measurement of AS based on the distribution of particle
41 size of each sample; this avoids the confusion caused by large individual soil particles
42 which are retained on sieve meshes and which resemble stable aggregates, ii) having a
43 greater number of aggregate size measurements, and iii) being a less labour-intensive,
44 and potentially faster, procedure than approaches based on wet-sieving. However, the
45 published method (Rawlins *et al.*, 2013) did not re-create the typically moist condi-
46 tion of field soils (in temperate regions) because it relied on the addition of air-dried
47 material to low ionic strength water. This limitation had been addressed previously
48 by Le Bissonnais (1996) in which one test included an optional pre-wetting stage in
49 the measurement of AS. Based on this approach, Le Bissonnais considered the two
50 dominant processes influencing aggregate fragmentation: i) large disaggregation inten-

sities caused by slaking, and ii) more limited effects due to clay mineral expansion. Accurate measurements of AS using sample aliquots of both air-dried and pre-wetted aggregates may help to establish the relative importance of these two processes in governing fragmentation. For example, recent experiments undertaken on degassed aggregates (reduced air pressures) showed that aggregate disruption cannot be explained by compression of pore air alone, but must be due to other processes (Czachor *et al.*, *in press*).

It is well-established that soil mineralogy (carbonate and iron oxides), exchangeable sodium percentage (ESP) and organic matter content exert a strong influence over AS (Amezqueta, 1999). Aggregates with larger quantities of soil organic matter are generally more stable. Based on applying a water coherence test (Emerson, 1967) to 180 soil samples from England and Wales, Greenland *et al.* (1975) suggested that there was a ‘critical level of organic matter (2% organic carbon), below which soils were very liable to structural deterioration especially in the absence of calcium carbonate’. However, it not clear whether there are biotic and abiotic *interactions* that influence the stability of aggregates, the combined effect of which may be greater than the sum of their independent effects. It may be possible to determine the significance of any such interaction empirically based on soil sample selection from our existing knowledge of soil variation at regional scales. First, by selecting soil samples from cultivated fields over differing parent material (PM) types we can ensure that there are substantial differences in soil mineralogy. Second, if we use measurements from recent (within the previous 12 years) soil surveys, we can select sampling sites with a wide range of organic matter concentrations; using our knowledge of organic matter turnover we know that establishing a new equilibrium soil organic matter content typically requires around 40 years due to changes in the quantity of organic matter inputs through altering land management practice (Greenland, 1995). In the UK there are data from a systematic regional soil geochemical survey undertaken between 1994 and 1996 in which alternative 1-km² cells of the British National Grid were sampled in the Humber-Trent region

of eastern England. The sampled soil was analysed for many properties including the concentration of organic carbon (Rawlins *et al.*, 2009a). We identified six parent material classes within this region. Soils in arable use on these parent materials within the region constitute our domain of interest. We then used the organic carbon data to stratify the domain and by stratified random sampling selected sample sites to collect soil so that we could examine the effects of organic carbon content and PM on AS.

To evaluate analytical reproducibility of laboratory-based tests, repeated measurements are typically made using aliquots taken from a homogenised reference material (RM) (European Accreditation, 2003). Such approaches are common for soil geochemical properties, but after an internet-based search we could not find any RMs for aggregation properties of soil or similar materials. To evaluate the reproducibility of our LG and sonication-based method (Rawlins *et al.* 2013), it would be necessary to establish a RM for AS measurement. If such a RM had consistent disaggregation properties then it could be measured repeatedly to evaluate systematic differences between the results obtained by different laboratories or drift in the results over time at any one laboratory. This would be essential if the method were to be adopted for soil monitoring. It is also necessary to establish the effect on AS, measured by this technique, of the duration of air-dry storage of soil samples, as shown in previous research (Blake & Gilman, 1970). This is important because in large-scale soil monitoring there may be considerable variation in the length of time that material is stored before analysis. Furthermore, it is standard practice to archive material from soil sampling and to re-analyse this material some years later.

In this paper we present an improved method for measuring the stability of aggregates (1–2 mm diameter); in this case applied to cultivated soils. We show how aliquots from a large quantity of homogenised palaeosol material can be used to monitor the reproducibility of the AS method. We investigate whether there are significant differences in AS within and between soils developed from different PM types, the influence of soil organic matter concentrations, and any interaction between them. We apply

our technique to both air-dried and pre-wetted aggregates and discuss the findings in terms of the likely fragmentation processes. We investigate whether air-dry storage of the soil samples influences AS, and whether this is the same for soils over all the PM types. We discuss the implications of our findings, in particular for establishing SOC concentration thresholds, below which soils may be prone to structural deterioration.

Materials and Methods

Selection of study sites, soil sampling and preparation

We wished to determine the stability of aggregates from cultivated soils with a wide range of SOC concentrations and a range of PM types. We used a database of SOC measurements from the Geochemical Baseline Survey of the Environment project (G-BASE) of the Humber-Trent region of eastern England (Rawlins *et al.*, 2009b) undertaken between 1994 and 1996. We identified the soil PM at each of these locations based on the PM map of the UK (British Geological Survey, 2006). We selected six soil PM types which occur in the region (Figure 1) to ensure that there would be marked variations among samples with respect to mineralogy. We selected two PMs with substantial calcium carbonate content, two PMs which weather to soil with a large clay content and one PM which is a sandstone. These five PMs are all solid bedrock, and we identified sites where no superficial Quaternary deposits had been mapped. These five bedrock types and the dominant World Reference Base (IUSS Working Group WRB, 2006) soil classes which form from them (in parenthesis) are: Chalk (Leptosol), Sherwood Sandstone (Arenosol), Lias clay (Gleysol), Lincolnshire Limestone (Leptosols) and Mercia Mudstone (Luvisol). The sixth PM was a superficial marine alluvium (Gleysol) which is present along the coastline of the study region. For purposes of statistical analysis we consider a classification of these parent materials into a nested set of categories, these are illustrated in Figure 2. The first division is between superficial and bedrock materials. Further divisions are based on lithological differences which may have a bearing on the aggregation properties of soils derived from these materials.

We collated the data on SOC concentrations for the samples from the original survey for each of the six PM types. We then sorted the samples by their SOC concentrations and randomly selected two samples from each decile of this distribution, providing 120 sample identifiers and associated location information. We used a set of recent, national air photos to check that all the fields where samples had been collected were still in arable production. We then sought permission to collect soil samples from one of this pair of sample sites; if this was not granted for the first, then we asked for permission to sample at the second site. Based on this procedure we resurveyed 60 of the original sampling locations during April 2012; conventional cultivation techniques appeared to have been used at all sites in the last growing season. At each site we adopted the same sampling procedure as the original survey. At each sampling site, five incremental soil samples were collected using a Dutch auger at the corners and centre of a square with a side of length 20 m and combined to form a composite sample of approximately 0.5 kg in a Kraft paper bag. At each of these five points, any surface litter was removed and the soil sampled to a depth of 15 cm into the exposed mineral soil. On return to the laboratory the soil samples were air-dried for 48 hours at room temperature, then sieved to pass 2 mm.

In addition, we collected two intact cylindrical soil cores for the determination of soil bulk density (BD), a fundamental soil physical property. The cores measured 50 mm in length and 53 mm diameter; they were collected from randomly selected opposite corners of the sampling square and placed these in separate plastic bags. On return to the laboratory the material was removed from the cores and oven-dried (105 °C for 24 hours), sieved to pass 2 mm and the resulting dry fine-fraction material was weighed. The coarse, intact material retained by the 2 mm sieve was weighed and its volume was measured by displacement. Bulk density of the fine-fraction was then computed as the oven-dry mass of the fine-fraction divided by the volume of the fine-earth fraction in the field. This latter volume was calculated by subtracting the volume of the material that did not pass the sieve from the volume of the section. The

resulting bulk density is that of the fine fraction (BDF; Hall *et al.*, 1977).

Total Organic Carbon

Soil organic carbon was estimated in each sample using loss-on-ignition analysis by heating a sub-sample to 450 °C for eight hours and multiplying the mass difference by 0.58 (Broadbent, 1953). The coefficient of variation for this method for 174 replicate analyses of a sample standard was 3.6 %.

Improved method for measuring aggregate stability

The complete description of the improved AS test is provided in Appendix 1; we also made a short film of the laboratory test which is available at the following internet url (http://youtu.be/7Y3qd_bqAXg). Here we describe the three changes we made to the original version of the AS test published in Rawlins *et al.* (2013) and our reasons for making these changes based on the results of testing various modifications:

1. *Aggregate pre-wetting*: We included a procedure to wet aggregates before they are used in the stability test; we felt this was more realistic of soils under field conditions of most temperate regions. However, it is still possible to apply the test using air-dry aggregates (no pre-wetting). We used both air-dry and pre-wetted aggregates on a series of samples in this study to determine the effects of the pre-wetting step on disaggregation reduction. However, for the main tests applied to all 60 samples, and the post-storage (repeat) analyses, we adopted the pre-wetting procedure.
2. *Initial aggregate dispersal and shorter period for particle size measurement*: In our original method, our first measurement of psd lasted for 60 seconds and began immediately after we had increased the pump speed to rapidly circulate aggregates through the LG. In the improved method we use a slower pump speed, but allow the aggregates to circulate for 30 seconds before undertaking an analysis of psd (over the next 30 seconds). We found that this shorter period of analysis (30 seconds) gave more reproducible mean weight diameter (MWD) values by

comparison to the original method (60 seconds psd analysis period).

3. *Change of sonication energy and duration:* In our original paper, to disrupt aggregates we used a sonicator (with its probe immersed in the water of the LG aqueous vessel) at maximum power (18 W) for ten minutes. In our new method we apply a more powerful sonicator (100 W) for 5 minutes. Tests showed that this greater power applied over the shorter period caused the aggregates to fragment to the same extent as in the original method (data not shown) and has the advantage of substantially reducing the overall duration of the test.

We compute the difference in the mean weight diameter (MWD) of the < 500 μm fraction of the two particle size distributions (aggregates and fundamental particles) and subtract the latter from the former which provides a measure we refer to as disaggregation reduction (units μm). Soils with more stable aggregates have larger values of DR. As in our original procedure, we undertake an AS test for two aliquots taken from each soil sample because we found that the mean value provides a more robust estimate of DR than a test on a single aliquot.

Aggregate reference material

Topsoils which are subject to seasonal cycles of organic matter input from litter and plant roots plus wetting and drying cycles, are unlikely to have sufficiently consistent disaggregation properties for use as an aggregate RM (Blackman, 1992). Palaeosols are likely to be less responsive to seasonal cycles because they may occur at greater depth in the soil profile (e.g. >1 m) subject to smaller changes in moisture and biogeochemical cycles, and we considered they may disaggregate more consistently over time (following appropriate treatment and storage). We collected a large (40 cm by 40 cm by 30 cm) block of palaeosol (brickearth) material from a site at Ospringe in Kent (southern England) at a depth of between 1.6 and 2 m. Further details of its mineralogy are described in detail by Clarke *et al.* (2007). Based on optical stimulated luminescence dating, the material at the base of the block has an age of around 18 700 yrs BP (\pm

217 2290 years).

218 We extracted a small volume (10 cm \times 5 cm \times 5 cm) from the block of Ospringe
219 palaeosol and dried it in an oven overnight at 40 ° C. We then sieved this material
220 into three aggregate size fractions: 250–500 μ m, 500–1000 μ m and 1000–2000 μ m and
221 discarded the finer material. We applied our modified LG aggregate test using samples
222 from each of the three size fractions, without the pre-wetting step (see Appendix 1).
223 We found that aggregates between 250 and 500 μ m gave the most consistent results
224 (not shown). We then broke up the large block of palaeosol material into smaller
225 blocks and followed the same drying and sieving procedure. We then combined all
226 the material in a large container and rotated this gently to homogenise the material,
227 but avoiding disruption to the aggregates. We then placed aliquots of this material
228 into around 200 labelled plastic containers with secure lids, and placed these in cold
229 storage (4 °C) to minimise any microbially mediated changes in the properties of the
230 palaeosol. We randomly selected five of these containers and determined DR values for
231 eleven aliquots from each container. We determined the quantity of SOC in one aliquot
232 of sample material taken from each from each of these containers using the method
233 described above.

234 There are several potential sources of error variance in the determination of DR
235 values for the aggregate RM including that introduced by: i) subsampling the RM,
236 ii) variation in energy imparted to the aggregates by the circulating water, iii) errors
237 from the two particle size measurements used to compute DR, and iv) variation in
238 the energy imparted by the sonicator forming the fundamental particles. We might
239 therefore expect the coefficient of variation of DR values computed from analyses of
240 aliquots of the RM to be somewhat larger than for traditional reference materials used
241 in soil science.

242 *Aggregate stability measurements*

243 A summary of the time periods between soil sample collection and the two main sets
244 of AS analyses is provided in Table 1. We spent nine months optimising the AS test,

altering several features of the test and also the pre-wetting procedure (Appendix 1). After each analysis we rescaled the size distributions to estimate DR for particle diameters $<500\text{ }\mu\text{m}$. We undertook several sets of AS tests based on groups of soil samples and the palaeosol RM:

1. We undertook 10 repeat analyses of the new procedure on aliquots from a single soil specimen to determine its reproducibility.
2. Throughout our analyses we made regular DR measurements of aliquots of the palaeosol RM; in total we measured the DR value 55 times (eleven aliquots from each of five selected containers). Two standard particle size materials (supplied by Beckman Coulter; mean diameter 32 and 500 μm) were used throughout the series of analysis to check for accuracy and precision of the psd measurements by the LG instrument (results not shown).
3. We determined the DR value for two aliquots from each of the sixty soil samples and computed the mean DR value (activity 2 in Table 1).
4. We selected one soil specimen from each PM and measured DR values using two aliquots of both dry and pre-wetted aggregates and computed the mean DR value based on these two analyses.
5. We undertook AS tests on two samples developed over the marine alluvium using a weak saline solution ($0.292\text{ g NaCl l}^{-1}$) to investigate whether the concentration of dissolved Na^+ ions influenced AS through its role in dispersing clays (Shainberg *et al.*, 1980). In these tests we used this saline solution in both the aggregate pre-wetting stage and also in place of the RO water circulating in the LG instrument.
6. After storing the samples for a total of 9 months we repeated the original analyses (see 3 above) on all sixty samples.

Statistical analyses

Our samples were selected by a random procedure from a large survey so for the purposes of statistical analyses we can consider them as independent samples. We used the basic linear model framework with mean DR (mean value from analyses of two aliquots) as the independent variable and SOC concentration and PM class as dependent variables for the 60 soil samples in our regional survey. We tested the effects of these factors (including an interaction term between them) by an analysis of variance (ANOVA).

In the resulting ANOVA the differences between PM groups are fixed effects because the sample design entailed the selection of a fixed number of samples from each class. In the ANOVA the PM effect has five degrees of freedom, and the associated *F*-test tests the null hypothesis of no difference among the PM means. It is informative to partition the PM effect into a set of pre-planned orthogonal contrasts, each with one degree of freedom. We based these contrasts on the hierarchical classification (Figure 2) of the six PM groups. The contrasts (C1-C5) were as follows.

1. Superficial versus Bedrock (C1)
2. Carbonate versus Siliciclastic-dominated (Bedrock) (C2)
3. Chalk versus Limestone (C3)
4. Mudrock versus Sandstone (C4)
5. Mudstone versus Clay (C5)

The effect captured by C1 in this particular setting is expected to be substantial because soils formed over the alluvium PM group may have poor aggregation properties because of a relatively large sodium absorption ratio inherited from its marine origin. Contrast 2 (C2) is of interest because soils formed over calcareous parent materials are expected to have good aggregation properties because of a large exchangeable calcium content. Contrast 4 (C4) compares soils expected to differ with respect to clay content because of their PM class. Contrasts 3 (C3) and 5 (C5) examine any variation within the carbonate and mudrock PM groups respectively.

We also undertook an ANOVA to investigate the variability of DR values from samples of the RM, which is necessary to assess the uniformity of the material and its suitability as a reference. Eleven DR measurements were made on material from each of five randomly selected containers. Because the containers were selected at random the container effect in the ANOVA is a random effect, and the between-container variance component can be estimated. This is of interest because it may indicate the effects of any biases in the subsampling process to prepare aliquots of the RM, as may be done when producing a sample to send to a laboratory as a reference. In the random effects ANOVA with $n=11$ measurements per container the variance among the container means is V_b :

$$V_b = s_b^2 + \frac{s_w^2}{n} \quad (1)$$

where s_b^2 is the between-pot variance component and s_w^2 is the within-container variance component. The residual mean-square (MS_r) in the ANOVA table estimates s_w^2 . The between-container variance component (s_b^2) can be estimated by:

$$s_b^2 = \frac{MS_c - MS_r}{n} \quad (2)$$

where MS_c is the container mean-square from the ANOVA table.

Results and their interpretation

Resurvey soil properties

Summary statistics of the SOC concentrations and fine earth bulk density (Bdf) are summarised by PM class in Table 2. We produced a scatterplot of the SOC concentrations between the original and resurvey samples (Figure 3) showing the PM of each sample. The concordance correlation (Lin, 1989), which assesses how strongly the values of two variates are clustered around the 1:1 line, had a value of 0.68 highlighting the strong similarity between the two surveys. We consider that our approach of resampling selected sites based on previous survey data was successful in ensuring a range of SOC concentrations in our resurvey. The SOC concentrations from the 2012 survey are within the typical range of cultivated topsoils of England (Verheijin *et al.*,

2005). The median SOC concentrations for five of the PM groups are quite similar (2.4–3.2 %) with a smaller value of 1.6% for the Sherwood Sandstone, likely accounted for by its coarser texture that results in less mineral-associated stabilisation of organic matter. With the exception of the soils over the Chalk and Lincolnshire limestone, the bulk density values (BDf) for the soils from four of the PM types are representative of those across England (Hall *et al.*, 1977). The minimum BDf values over the Chalk (0.3 g cm⁻³) and limestone (0.4 g cm⁻³) are substantially smaller than most mineral soils across England, with low to moderate SOC concentrations, and these may be accounted for by the occurrence of large quantities of porous, coccolithophores (Kerry *et al.*, 2009) which account for a large proportion of the mineral component of these soils.

Reproducibility of DR values

Repeated analyses (in which aggregates were pre-wetted) of ten aliquots taken from a soil specimen from the Mercia Mudstone PM gave a mean DR value (rescaled to <500 µm; see Appendix) of 64.7 µm, with a coefficient of variation (CV) of 11.6%. This CV is smaller than we reported for repeated analyses of soil aliquots from a single soil specimen (19%) using our original procedure (Rawlins *et al.*, 2013).

The mean concentration of SOC for a single aliquot taken from each of the five palaeosol RM containers was 1.4%, with a standard deviation of 0.22%. The overall mean DR value from 55 analyses of aliquots of the palaeosol RM (11 aliquots from each of 5 containers) was 21.2 µm (standard deviation=3.46 µm; Table 3) with an overall coefficient of variation of 16.3%. The boxplot of DR values for the aggregate RM by container number in Figure 4 shows that the majority of the variation is within containers. We used the output from the ANOVA analyses (Table 4) in Equation (2) to compute an estimate of between-container variance (s_b^2) of 0.02, which is not significantly different from zero ($P=0.41$). Although the coefficient of variation is larger than for repeated analyses of more common soil RMs (e.g. geochemical analyses), due in part to the sources of variation we described above, we consider our containers of

RM can provide a consistent baseline for monitoring the reproducibility of our method.

Influence of soil organic carbon and parent material type

In Figure 5 we present a scatterplot of SOC concentrations versus DR value for soils from each of the six PM types. Although soils with larger organic carbon concentrations tend to have larger DR values, the influence of PM group appears to be substantial with marked differences in AS (DR) for the soil samples collected over the six different PM types (Figure 6). Median DR values for four of the six parent materials (Chalk; CK, limestone; LL, clay; LI; mudstone; MMG) were markedly larger ($>80\text{ }\mu\text{m}$) than the median DR value ($<50\text{ }\mu\text{m}$) for the other two PM types (marine alluvium (ALV) and sandstone (SSG)). It is noteworthy that the soils with the smallest mean DR values, developed over the Sherwood Sandstone and Marine Alluvium, also had the smallest inter-quartile ranges (box lengths in Figure 4).

The summary output from analysis of variance (Table 5) shows that both PM class and SOC concentration are statistically significant effects ($P < 0.006$), but the interaction between them is not significant. This suggests that in a linear model for DR, the slopes of the regression lines would be the same for the different PM classes, but that these have different intercepts. The two predictors (PM class and SOC) accounted for 53% of the total variation in mean DR value for the 60 soil samples in our survey.

Of the five orthogonal contrasts that were included in the statistical analysis the following were statistically significant: Superficial (marine alluvium) versus Bedrock (C1), Carbonate versus Siliciclastic-dominated (C2) and Mudrock versus Sandstone (C4; Table 5). The marine alluvium (ALV) deposit is substantially more sodic than the other five PM types; its total Na concentrations ($> 0.4\%$) are amongst the largest across England and Wales (see Rawlins *et al.*, 2012, p. 168) and we might expect aggregates in these soils to be unstable due to Na-related clay mineral dispersion. The second contrast (C2: soils developed over Carbonate versus Siliciclastic bedrock) was also statistically significant which in part may be attributable to the effect of a larger

exchangeable calcium content of the former soil types.

Neither of the contrasts (C3 and C5) between soils derived from similar PM classes (carbonate and mudrocks) were statistically significant which we might expect because these soils have similarities in their dominant mineral types. The fourth contrast (C4), between sandstone and mudrock-derived soils, was statistically significant ($P=0.0003$). The findings from the analyses of these contrasts highlight the importance of large mineralogical differences between the PM groups in determining aggregate stability (DR values).

Soils developed over the Mercia Mudstone had the largest variation in DR value (31–112 μm) which may be attributable to the occurrence of regions of both clay and sand-rich sequences in the PM. Previous psd analyses of soils from sites over the Mercia Mudstone Group (Rawlins *et al.*, 2009b), where geological mapping suggests no superficial material occurs, showed that its texture varies widely (sand: 11–69%, clay: 11–46%, silt 19–59%).

Air-dry versus pre-wetted aggregates

The results of our AS analyses based on aliquots of dry and pre-wetted samples are presented in Table 6. For three of the PM types (ALV, MMG, SSG) the pre-wetted aggregates have substantially smaller (22–50 μm) DR values than the air-dry aggregates taken from the same specimen of soil. These three PM types also have smaller median DR values (Figure 5) than the other PM types. For the other three PM types (CK, LI, LL), the differences between DR values for air-dry and pre-wetted aggregates are generally rather small (1–7 μm). According to our current understanding, the process of clay mineral expansion would have a significant impact on aggregate fragmentation during pre-wetting, whilst we would expect slaking to dominate when air-dry aggregates are immersed into the RO water.

Influence of electrolyte composition

Based on tests undertaken on the same day, the DR values for two aliquots of soil

material taken from a soil specimen over the marine alluvium, with a weakly saline pre-wetting and circulating solution, were 46.1 and 54.8 (μm), whilst two other aliquots measured in RO water had DR values of 54.1 and 51.7 (μm). Given the substantially larger overall range of DR values between our sixty soil samples, we considered this was sufficient evidence to conclude that modifying the electrolyte composition of both the pre-wetting and circulating solutions would not have a significant impact on the differences observed between AS at regional scales.

Changes in aggregate stability after storage

We present a scatterplot of DR values for each of the 60 soil samples before and after ten months of further air-dry storage (Figure 7). These results suggest there may be an increase in AS with time; a greater number of the points are above the 1:1 line than occur below it. This suggests a small shift to larger values of DR through storage. In Table 7, we present the results of t -tests to determine whether there were statistically significant changes in mean AS after storage. Considering samples from all PM types, there is evidence to suggest an increase in mean AS with time ($P < 0.0001$) based on results from the t -test. However, for two of the PM types (alluvium and mudstone) there was no evidence for an increase in mean DR value (t -test P -values 0.46 and 0.74 respectively). These results suggest that the processes governing changes in AS over short periods of storage may not be consistent across soil types.

Discussion

Our analyses of soils developed from different PM types suggest that it may be too simplistic to state a single, critical threshold of organic carbon concentration in terms of structural stability (2% suggested by Greenland *et al.* (1975) - see Figure 5), albeit qualified by their reference to the absence of calcium carbonate. Our results suggest that the composition of the soil PM, altered through pedogenesis, is also critical in terms of understanding soil structural stability. Assuming our AS test provides an effective means of assessing the structural stability of cultivated topsoil, it may be more

effective to establish SOC concentration thresholds for soils developed from different groups of PM in temperate regions such as the UK. Based on two predictors (PM class and SOC concentration), our model accounted for 53% of the variance in DR. Although our approach needs testing in a wider range of landscapes and climatic settings, we consider it has potential for the prediction and (or) mapping of AS (and associated uncertainties), at landscape scales.

There was no evidence for a statistically significant interaction between SOC concentration and PM type from our mean DR measurements (P -value=0.115). This suggests that the strength of the bonds between soil organic constituents and the various soil minerals derived from the PM types are not significantly different. Given resource limitations, we could not account for the timing and form of any organic materials added to the soils from our sixty survey sites, which we recognise would have implications for our measurements of AS (Watts *et al.*, 2001).

Although studies have investigated some of the relationships between AS and other critical soil properties which influence a range of soil functions and degradation threats, few have utilised relatively novel soil technologies such as X-ray computed tomography. Such approaches can accurately quantify intra-aggregate pore size distributions in small aggregates (1–2 mm diameter) at voxel resolutions as fine as 10 μm . For example, based on a X-ray CT voxel resolution of 14.2 μm , Wang *et al.* (2012) showed that soils under native vegetation had a broader range of intra-aggregate pore sizes when compared to soil under conventional tillage. Exploring the relationships between AS using our new procedure and intra-aggregate pore size distributions could help to elucidate the role of the former in important soil functions such as gaseous exchange between soil and the atmosphere and soil hydraulic properties.

We found that for less stable aggregates, use of the pre-wetting procedure led to greater reductions in DR values (less stable aggregates) by comparison to use of air-dry aggregates based on aliquots from the same soil specimen. This suggests that processes associated with wetting, most likely clay mineral expansion and Na-related dispersal,

causes greater short-term fragmentation (the water circulation step in our stability test) by comparison to slaking (forces associated with trapped air). This supports recent findings showing that aggregate disruption cannot be explained by pore air compression, but is likely due to other processes (Czachor *et al.*, in press). With the exception of the uppermost surface of the soil which has been subject to periods of dry weather, most field soils are at varying moisture contents when subjected to disruptive forces (tillage, soil fauna, and raindrop impact). Based on our findings, we consider that using pre-wetted aggregates provides a more realistic basis for comparing AS across a broader range of soil types. For example, DR values are consistently larger than 70 μm for topsoil samples from predominantly grassland sites, with large SOC concentrations ($> 5\%$; data not shown).

One potential limitation of our current AS test is that we have not quantified the energy used to disrupt the aggregates during the initial water circulation step. It may be possible to estimate dispersive energy more accurately by adopting a system similar to that developed by Zhu *et al.* (2009), an ultrasonic system with a power-adjusting feedback mechanism. We propose to replace the RO water in our LG instrument with a liquid for which the density can be adjusted to match soil aggregates (e.g. sodium polytungstate; SPT) and also to reduce the circulation speed of the system. Initial test results (not shown) suggest that fragmentation of micro aggregates ($< 250 \mu\text{m}$) circulated through the LG in SPT (density 1.6 g cm^{-3}) is minimal. By doing so we can minimise the dispersive energy associated with hydrodynamic disruption, whilst ensuring accurate aggregate size measurements. We could then apply energy through a sonicator to this system and calculate the dispersive energy required to reduce the MWD of a soil aliquot to the same value based on water circulation (our current approach) applied to another aliquot from the same specimen.

We are not aware of any comparable RM that would be appropriate for testing AS using our procedure. Our palaeosol RM appears to have sufficiently consistent disaggregation properties for its use to monitor the precision and any bias associated

with DR analyses. It would be necessary to undertake tests at a range of laboratories, and over a longer period, to assess its long-term suitability as an aggregate RM.

Conclusions

Using a modified procedure for measuring aggregate stability (DR value; units μm) we showed that the stability of pre-wetted aggregates is closely related to the PM in which the soil developed, and infer that this is due to differences in soil mineralogy. We measured AS of 60 topsoil samples (10 samples from each of six PM types) across a region of cultivated soil from eastern England, maximising the variation in organic matter concentration by resampling sites from a previous survey. Soil aggregates developed over calcareous PM types have large DR values (more stable aggregates) than those developed over sandstone and another group of soils that contain large quantities of free Na which leads to clay dispersion. Our results show that any proposed thresholds of organic matter concentration, below which the physical properties of the soil may deteriorate, are specific to the PM from which the soil formed. There was no evidence to suggest an interaction effect between soil organic matter concentrations and soil PM (mineralogy) on aggregate stability (DR value). A linear regression model with PM group and organic carbon concentration as predictors accounted for 53% of the variation in DR values for soil from the 60 sites. From AS analyses on aliquots from the same soil specimen, with and without the pre-wetting stage, we found that pre-wetting led to smaller DR values (less stable aggregates) suggesting that for our method clay mineral expansion may be more important than slaking in terms of fragmentation processes.

Storing aggregates for a period of 9 months led to a small, but statistically significant increase in AS if all samples are considered. However, AS for soil samples from two of the PM groups did not change significantly over the period of storage when considered individually. We showed that a palaeosol RM (from southern England) has consistent disaggregation properties which suggest it may be sufficient to assess precision and bias associated with our AS method for monitoring soil physical quality.

Acknowledgements

We would like to thank Judith Mather, Andy Tye and Jenny Bearcock for assistance with collection of the soil samples and Rachel Efrat for undertaking extra analyses of the aggregate reference material. This paper is published with the permission of the Executive Director of the British Geological Survey (Natural Environment Research Council).

References

- Amezketta, E. 1999. Soil aggregate stability: A review. *Journal of Sustainable Agriculture*, **14**, 83–151.
- Barthes, B. & Roose, E. 2002. Aggregate stability as an indicator of soil susceptibility to runoff and erosion; validation at several levels. *Catena*, **47**, 133–149.
- Blackman, J. D. 1992. Seasonal-variation in the aggregate stability of downland soils. *Soil Use & Management*, **8**, 142–150.
- Blake, G. R. & Gilman, R. D. 1970. Thixotropic changes with aging of synthetic soil aggregates. *Soil Science Society of America Journal*, **34**, 561–564.
- British Geological Survey, 2006. *Digital Geological Map of Great Britain 1:50 000 scale (DiGMapGB-50) data [CD-ROM] Version 3.14*. British Geological Survey, Keyworth, Nottingham.
- Broadbent, F. E. 1953. The soil organic fraction. *Advances in agronomy*, **5**, 153–183.
- Clarke, M. L., Milodowski, A. E., Bouch, J. E., Leng, M. J. & Northmore, K. J. 2007. New OSL dating of UK loess: indications of two phases of Late Glacial dust accretion in SE England and climate implications. *Journal of Quaternary Science*, **22**, 361–371.
- Czachor, H., Charytanowicz, M., Gonet, S., Niewczas, J., Jozefaciuk, G. & Lichner, L. in press. Impact of long term mineral and organic fertilization on water stability,

wettability and porosity of aggregates of two loamy soils. *European Journal of Soil Science*.

Emerson, W.W. 1967. A classification of soil aggregates based on their coherence in water. *Australian Journal of Soil Research*, **5**, 47–57.

European Accreditation, 2003. The Selection and Use of Reference Materials. European Accreditation Report EA-4/14 INF:2003. Accessed 10 September 2014. <http://www.european-accreditation.org/publication/ea-4-14-inf>

Farres, P. J. 1987. The dynamics of rainsplash erosion and the role of soil aggregate stability. *Catena*, **14**, 119–130.

Greenland, D.J. 1995. Land use and soil carbon in different agroecological zones. In: *Soils and global Change*. Lal, R., Kimble, J., Levine, E., Stewart, B.A. (eds). CRC & Lewis Publishers, Boca Raton, Florida. p 1–24

Greenland, D.J., Rimmer, D. & Payne, D., 1975. Determination of the structural stability class of English and Welsh soils, using a water coherence test. *Journal of Soil Science*, **26**, 294–303.

Hall, D.G.M., Reeve, M.J., Thomasson, A.J. & Wright, V.F. 1977. *Water Retention, Porosity and Density of Field Soils*. Soil Survey Technical Monograph No. 9. Soil Survey of England and Wales, Harpenden.

Haynes, R.J. & Knight, T.L. 1989. Comparison of Soil Chemical-Properties, Enzyme-Activities, Levels of Biomass-N and Aggregate Stability in the Soil-Profile under Conventional and No-Tillage in Canterbury, New-Zealand. *Soil & Tillage Research*, **14**, 197–208.

IUSS Working Group WRB, 2006. *World Reference Base for Soil Resources 2006. 2nd edition*. World Soil Resources Reports No. 103. FAO, Rome.

- Kerry, R., Rawlins, B. G., Oliver, M.A. & Lacinska, A. M. 2009. Problems with determining the particle size distribution of chalk soil and some of their implications. *Geoderma*, **152**, 324–337.
- Le Bissonnais, Y. 1996. Aggregates stability and assessment of soil crustability and erodibility: I. Theory and methodology. *European Journal of Soil Science*, **47**, 425–437.
- Lin, L.I-K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, **45**, 255–268.
- Rathore, T.R. Ghildyal, B.P., & Sachan, R.S. 1982. Germination and emergence of soybean under crusted soil conditions II: Seed environment and varietal differences. *Plant and Soil*, **65**, 73–77
- Rawlins, B.G., McGrath, S.P., Scheib, A.J., Breward, N., Cave, M., Lister, T.R., Ingham, M., Gowing, C. and Carter, S. 2012. *The Advanced Soil Geochemical Atlas of England and Wales*. Keyworth, Nottingham: British Geological Survey. Accessed 10 September 2014. <http://www.bgs.ac.uk/ebooks/AdvancedSoilGeochemicalAtlasEbook/index.html\#/1/>
- Rawlins, B. G., Scheib, A. J., Lark, R. M. & Lister, T. R. 2009a. Sampling and analytical plus subsampling variance components for five soil indicators observed at regional scale. *European Journal of Soil Science*, **60**, 740–747
- Rawlins, B.G., Webster, R., Lawley, R., Tye, A.M. & O’Hara, S.O., 2009b. Estimating particle-size fractions of soil dominated by silicate minerals from geochemistry. *European Journal of Soil Science*, **60**, 116–126.
- Rawlins, B. G., Wragg, J. & Lark, R. M. 2013. Application of a novel method for soil aggregate stability measurement by laser granulometry with sonication. *European Journal of Soil Science*, **64**, 92–103.

- Rickson, R. J., Deeks, L.K. , Corstanje, R., Newell-Price, P., Kibblewhite, M.G.,
Chambers, B., Bellamy, P., Holman, I., James, I.T., Jones, R., Kechavarsi, C.,
Mouazen, A.M., Ritz, K. & Waite, T. 2013. Indicators of soil quality physical
properties. Defra Report SP1611, London. [http://randd.defra.gov.uk/](http://randd.defra.gov.uk/Default.aspx?Menu=$Menu\&Module=$More\&Location=None\&Completed=1\&ProjectID=17595)
[Default.aspx?Menu=\\$Menu\&Module=\\$More\&Location=None\&Completed=1\](http://randd.defra.gov.uk/Default.aspx?Menu=$Menu\&Module=$More\&Location=None\&Completed=1\&ProjectID=17595)
[\&ProjectID=17595](http://randd.defra.gov.uk/Default.aspx?Menu=$Menu\&Module=$More\&Location=None\&Completed=1\&ProjectID=17595). Accessed 10 September 2014.
- Shainberg, I., Rhoades, J. D. & Prather, R. J. 1980. Effect of low electrolyte concentration on clay dispersion and hydraulic conductivity of a sodic soil. *Soil Science Society of America Journal*, **45**, 273–277.
- Verheijen, F. G. A., Bellamy, P. H., Kibblewhite, M. G. & Gaunt, J. L. 2005. Organic carbon ranges in arable soils of England and Wales. *Soil Use and Management*, **21**, 2–9.
- Wang, W., Kravchenko, A. N., Smucker, A. J. M., Liang, W. & Rivers, M. L. 2012. Intra-aggregate Pore Characteristics: X-ray Computed Microtomography Analysis. *Soil Science Society of America Journal*, **76**, 1159–1171.
- Watts, C.W., Whalley, W.R., Longstaff, D.J., White, R.P., Brookes, P.C. & Whitmore, A.P. 2001. Aggregation of soil with different cropping histories following the addition of different organic materials. *Soil Use and Management*, **17**, 1-6.
- Zhao, H. L., Cui, J. Y., Zhou, R. L., Zhang, T. H., Zhao, X. Y. & Drake, S. 2007. Soil properties, crop productivity and irrigation effects on five croplands of Inner Mongolia. *Soil & Tillage Research*, **93**, 346–355.
- Zhu, Z.L., Minasny, B. & Field, D.J. 2009. Measurement of aggregate bond energy using ultrasonic dispersion. *European Journal of Soil Science*, **60**, 695–705.

Figure captions

Figure 1 Sixty locations where soil samples were collected for this study on two occasions: i) an original regional survey (summer months of 1994, 1995 or 1996), and ii) a resurvey in spring 2012. The symbols refer to the six parent material types: ALV=marine alluvium, CK=Chalk, LI=Lias clay, LL=Lincolnshire Limestone, MMG= Mercia Mudstone Group, SSG=Sherwood Sandstone Group. The bounding region of the samples extends from 461300–550790 (Easting) and 347700–455600 (Northing; metres on the British National Grid).

Figure 2 Hierarchical classification of the six parent material types used to define groups for statistical analysis using orthogonal contrasts (see Materials and Methods).

Figure 3 Scatterplot of soil organic carbon concentrations (%) in soil from sixty sampling locations for: i) an original survey between 1994 and 1996, and ii) form a resurvey of the same locations in 2012.

Figure 4 Boxplot of disaggregation reduction (DR) values from analyses of fifty-five aliquots of the palaeosol reference material from five separate containers. The overall mean is shown by the grey dashed line.

Figure 5 Scatterplot of soil organic carbon concentrations (sampled in 2012) and disaggregation reduction values for the 60 soil samples across the study region. The dashed line shows the threshold concentration (2% SOC) below which Greenland *et al.* (1975) considered that soils would be very liable to structural deterioration.

Figure 6 Boxplot of disaggregation reduction (DR; μm) for ten soil samples from each of six parent material types across eastern England.

Figure 7 Scatterplot of disaggregation reduction (DR) values before and after a period of nine months air-dry storage for soil samples from six different parent material types.

Table 1 Timing of activities undertaken in the present study that relate to aggregate stability (AS) testing and sample storage.

Activity	Period since field sampling (months)	Duration (weeks)
1.Soil sampling at 60 selected locations	–	4
2.Laboratory tests to optimise AS technique	1	24
3.Initial AS tests ($n=60$ samples)	9	2
4. ^a Extra AS tests on ALV soils ($n=5$ samples)	13	1
5.Storage period	–	40
6.Repeat AS testing ($n=60$ samples)	18	2

^a The sodic ALV soils were tested in a weak saline solution to determine whether the composition of the electrolyte solution had an effect on their stability.

643 **Table 2** Summary statistics for soil organic carbon (SOC; %) concentration and
644 fine earth bulk density (BDf; g cm⁻³) for a total of sixty cultivated soils developed
645 from six differing parent materials: ALV=marine alluvium, CK=Chalk, LI= Lias clay,
646 LL=Lincolnshire limestone, MMG= Mercia mudstone, SSG=Sherwood sandstone.

	PM code	Min	Max	Mean	Median	<i>n</i>
SOC	ALV	1.6	7.0	3.3	2.9	10
	CK	1.8	3.9	2.9	2.9	10
	LI	1.6	4.3	3.2	3.4	10
	LL	1.6	3.3	2.4	2.4	10
	MMG	1.7	4.8	2.9	2.5	10
	SSG	1.2	2.5	1.6	1.5	10
	All	1.2	7.0	2.7	2.5	60
BDf	ALV	0.9	1.4	1.2	1.2	10
	CK	0.3	1.4	0.8	0.9	10
	LI	0.9	1.5	1.2	1.1	10
	LL	0.4	1.4	1.0	1.0	10
	MMG	0.8	1.4	1.2	1.2	10
	SSG	0.9	1.5	1.2	1.2	10
	All	0.3	1.5	1.1	1.2	60

647

Table 3 - Selected statistics from DR analyses of aliquots of sample material (n=11) taken from five selected containers of the palaeosol aggregate reference material.

Container number	Mean DR (μm)	Std. Dev. (μm)	Coefficient of Variation (%)	Number of sample aliquots
1	19.9	3.63	18.3	11
2	21.5	1.80	8.38	11
3	22.8	3.07	13.5	11
4	21.3	2.34	11.0	11
5	20.7	5.35	25.8	11
All	21.2	3.46	16.3	55

651 **Table 4** - Output from ANOVA analysis in which the DR value from the aggregate
652 RM was the independent variable and container number (between 1 and 5) was a
653 grouping variable or classification.

		Df	Sum Sq	Mean Sq	F value	<i>P</i> -Value
654	Container	4	48.9	12.2	1.019	0.407
	Residuals	50	599	12.0		

655 **Table 5** - Summary output from analysis of variance of mean DR value in which the
656 dependent variables were parent material (PM) class, soil organic carbon concentration
657 and an interaction between them (PM×SOC). The second part of the table shows the
658 partition of the PM effect into five 1-degree of freedom contrasts (C1 to C5 - see
659 Materials and Methods).

	Degrees of Freedom	Sum of Sq.	Mean Sq.	F value	<i>P</i> -value
PM	5	29743	5949	12.3	< 0.0001
SOC	1	4275	4275	8.84	0.005
PM×SOC	5	4555	911	1.88	0.115
Residuals	48	23219	483		
Partition of PM effects					
Contrast	Degrees of Freedom	Sum of Sq.	Mean Sq.	F value	<i>P</i> -value
C1	1	15769	15769	32.6	<0.0001
C2	1	5071	5071	10.5	0.0022
C3	1	111	111	0.23	0.633
C4	1	7418	7418	15.3	0.0003
C5	1	1372	1372	2.84	0.099
Total	5	29743			

661 **Table 6** Disaggregation reduction (DR; μm) values for aliquots from a selected soil
662 specimen from each of the six parent material types based on use of: i) pre-wetted
663 aggregates, and ii) air-dry aggregates.

	Pre-wetted aggregates	Air-dried aggregates
Marine alluvium	51	101
Chalk	67	59
664 Lias clay	54	61
Linconshire limestone	57	56
Mercia mudstone	49	86
Sherwood sandstone	28	50

Table 7 Results of t -tests applied to mean weight diameter values (MWD; μm) of 1-2 mm sieved soil samples to determine whether there is a significant change in mean size before and after storage period of 9 months. Individual t -tests were applied to samples from within each parent material group ($n=10$) and to all samples ($n=60$). Fundamental particles and aggregates refer to the post-sonication and pre-sonication MWD, respectively.

Parent material	Fundamental Particles				Aggregates	
	t -statistic	P -value	mean difference (μm)	t -statistic	P -value	mean difference (μm)
Marine Alluvium	-0.970	0.357	-1.072	-0.767	0.463	-2.721
Chalk	0.638	0.539	1.556	-4.722	0.001	-12.064
Lias Mudstone	-0.989	0.349	-1.858	-3.371	0.008	-21.261
Lincolnshire Limestone	-1.560	0.153	-1.384	-6.421	<0.0001	-14.706
Mercia Mudstone Group	1.852	0.097	3.402	-0.338	0.743	-2.323
Sherwood Sandstone Group	0.433	0.676	0.811	-2.400	0.040	-3.049
All	0.335	0.739	0.243	-4.828	<0.0001	-9.354

Figure 1:

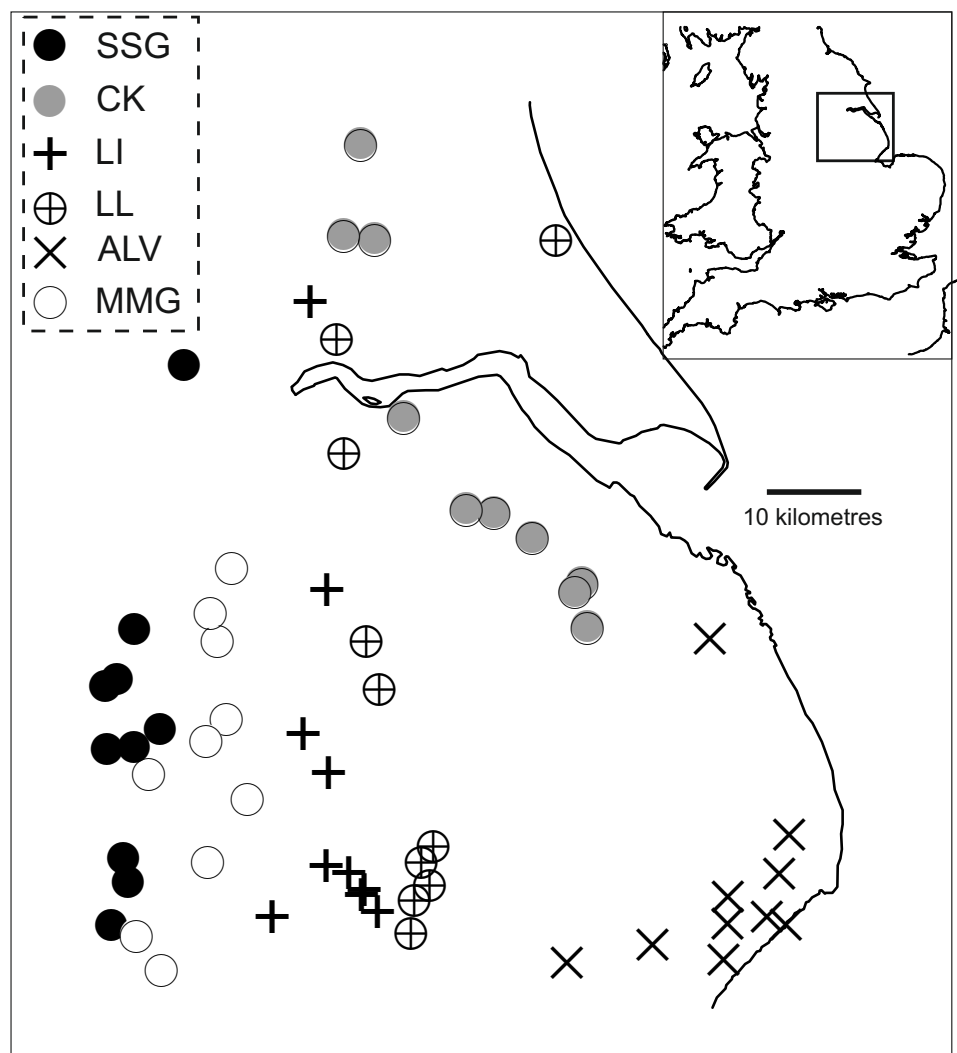


Figure 2:

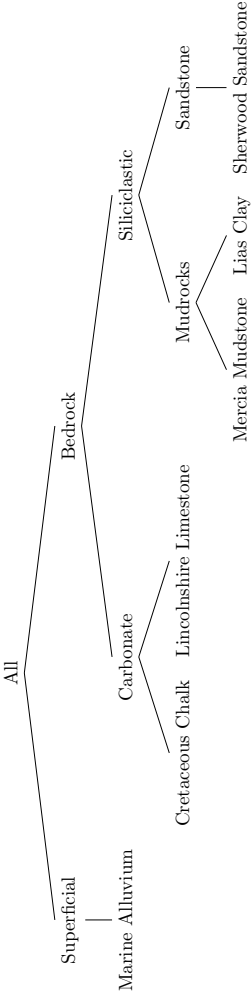


Figure 3:

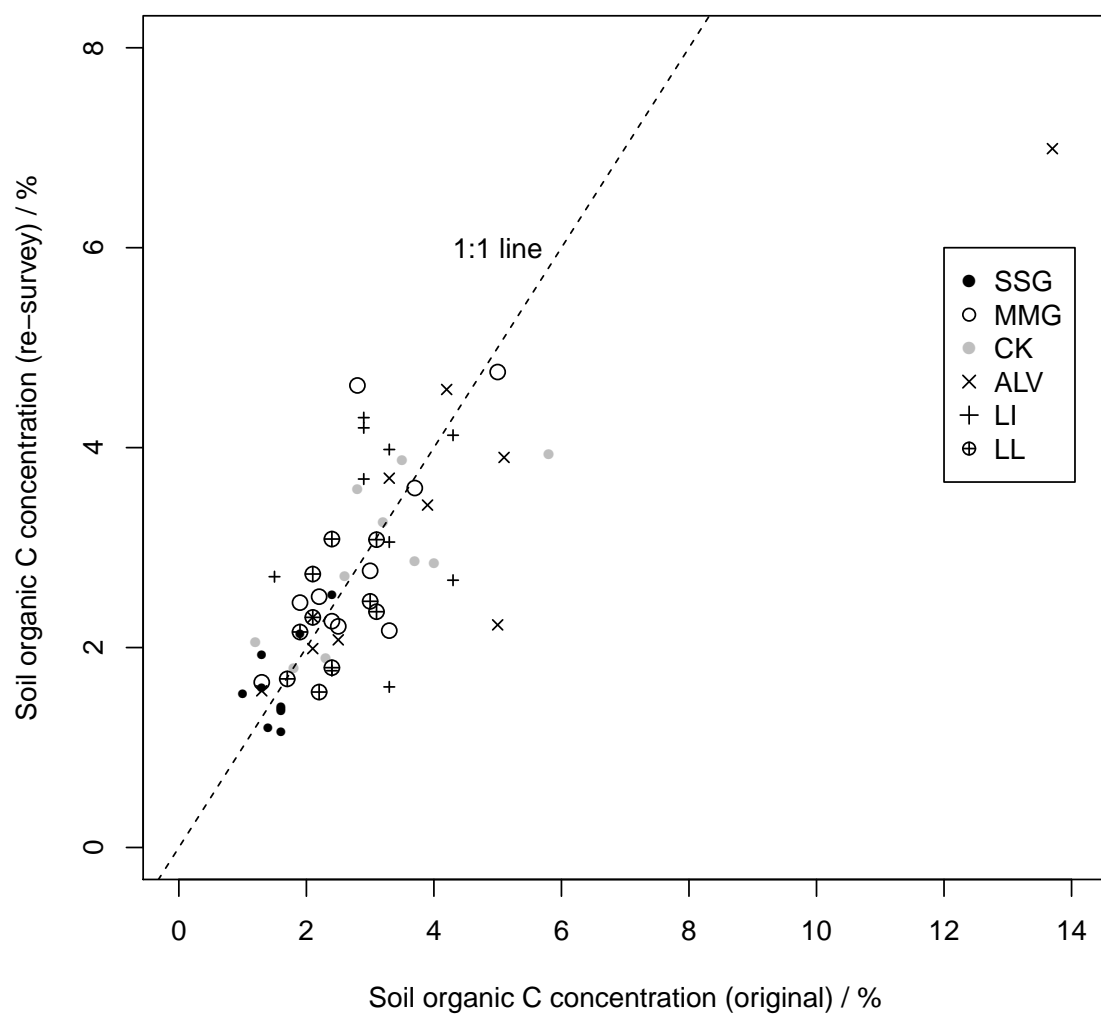


Figure 4:

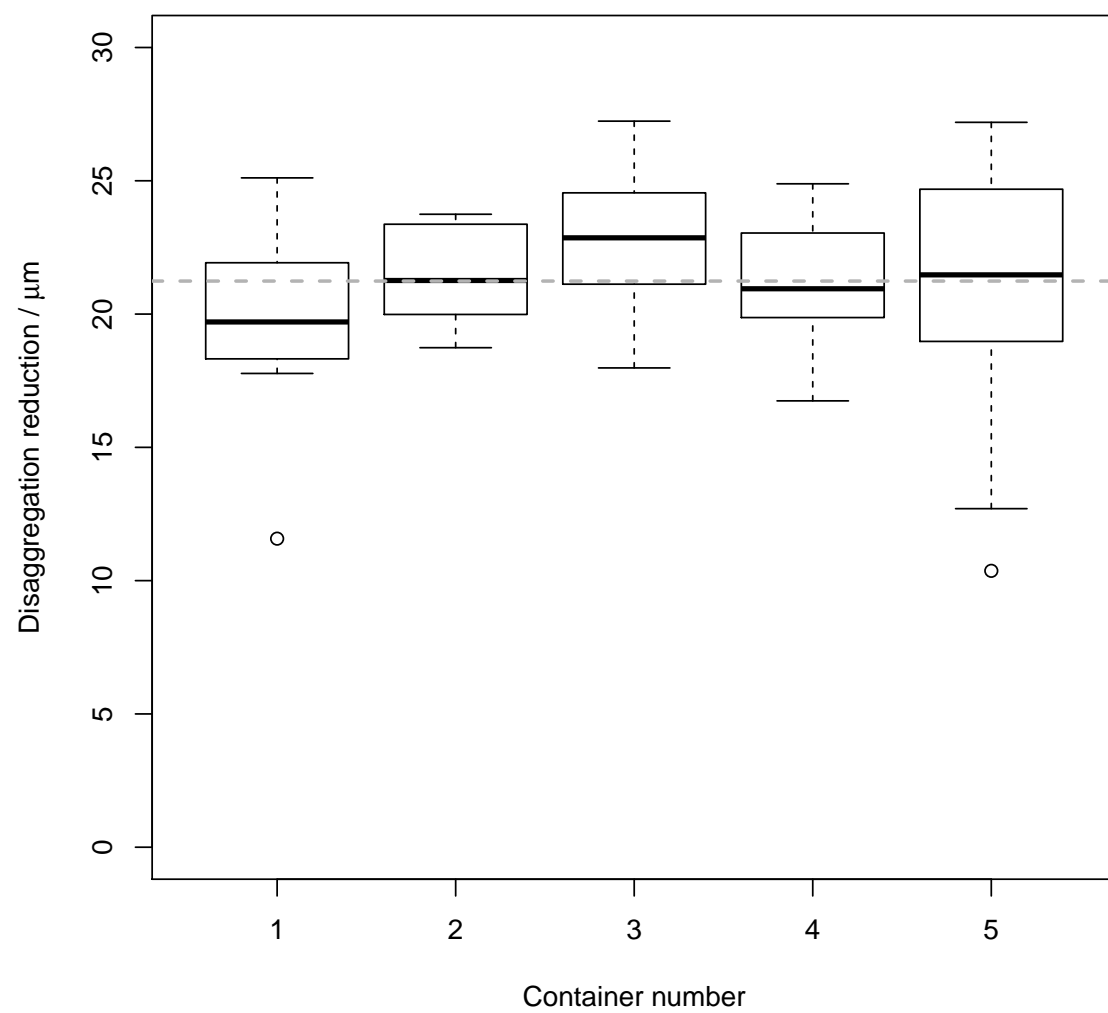


Figure 5:

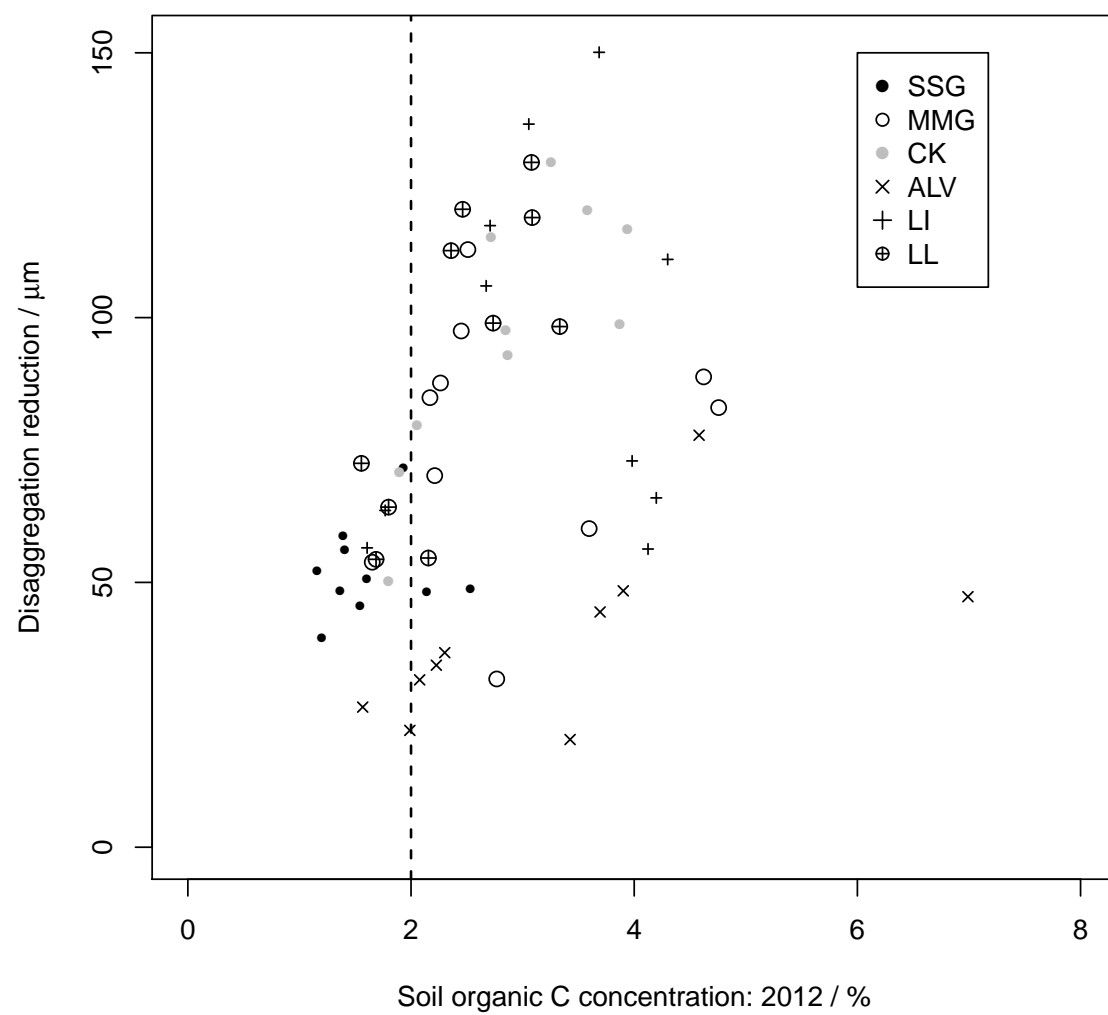


Figure 6:

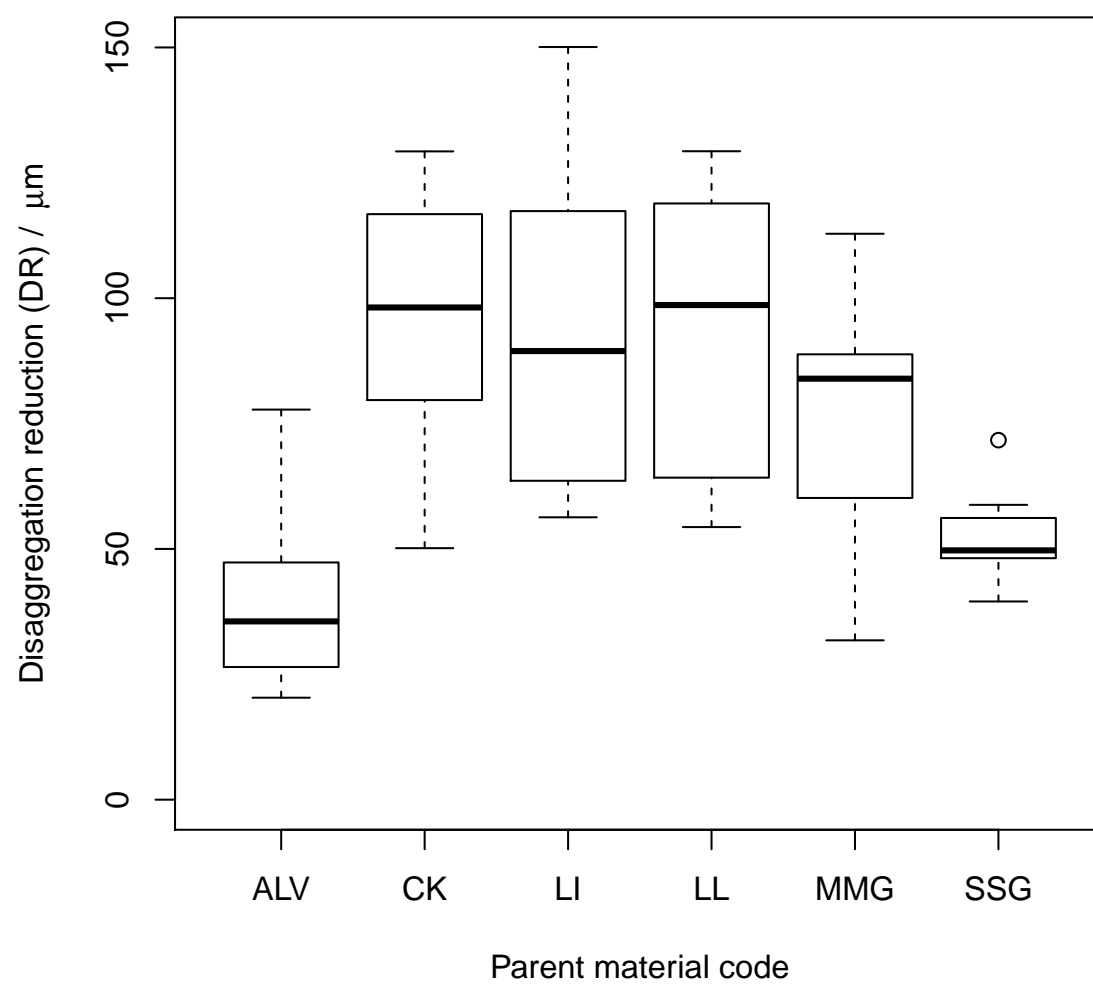
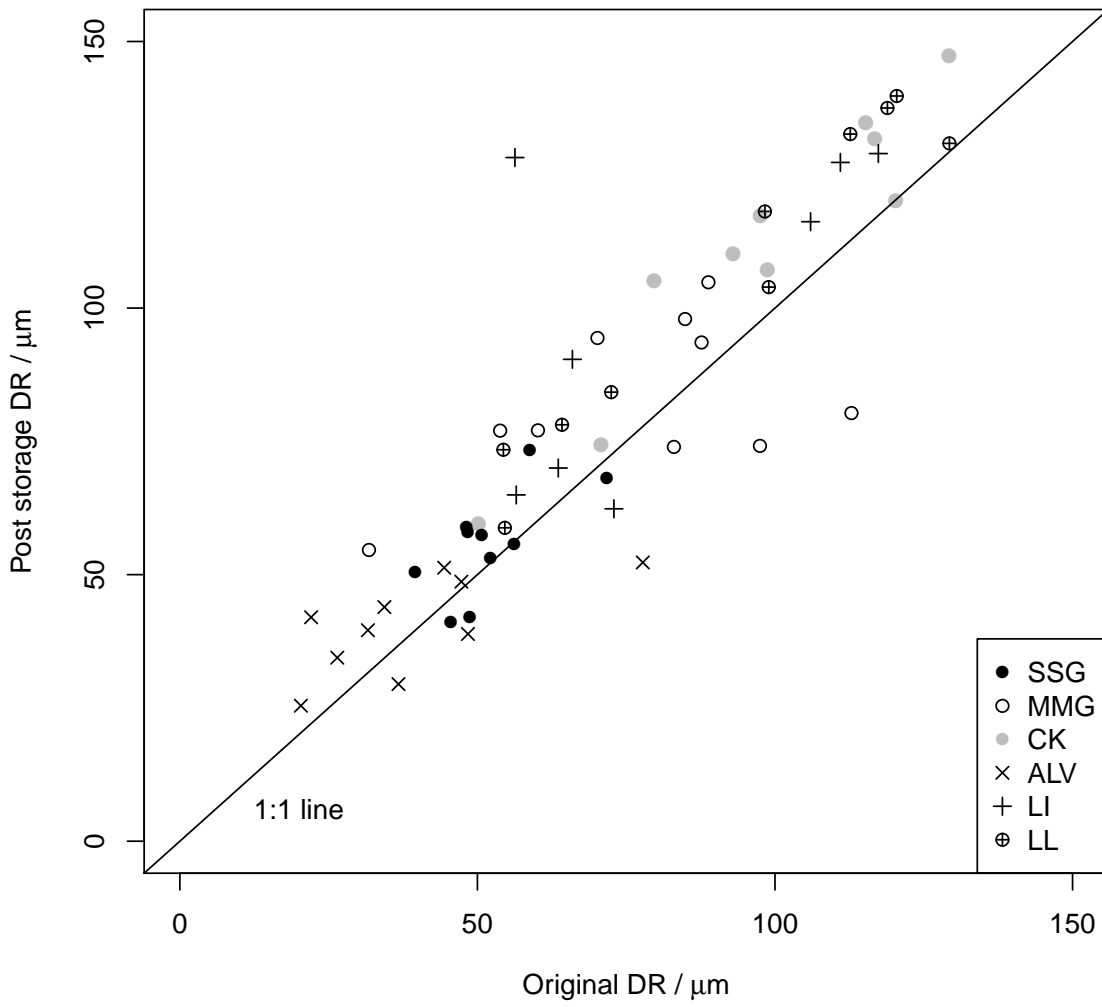


Figure 7:



Appendix 1: Procedure for aggregate stability testing using laser granulometry with aggregate pre-wetting

Our aggregate stability measure is the difference between two measurements of MWD for a soil specimen. The first MWD measurement is made after an aliquot of soil has been subject to circulation in water and the resulting mild disruptive forces (water stable aggregates). The second measurement is made after applying a sonication treatment which subjects the aggregates to strong disruptive forces. The difference between the first and second MWD is called the disaggregation reduction or DR. If two soils differ with respect to the stability of their aggregates in water then we would expect the soil with the more stable aggregates to have the larger DR.

Before undertaking the aggregate stability (AS) test we first determine an approximate mass of soil aggregates (1-2 mm) to be used for each soil specimen. It is more efficient if this is undertaken using aliquots from each soil specimen prior to undertaking any AS tests. Determining a starting mass for the aggregate pre-wetting part of the AS test is necessary because soils have widely differing particle size distributions (psd). If too large a quantity of fine soil material is placed into the Laser Granulometer (LG) instrument, the obscuration of light may be greater than the maximum limit (18%) for accurate particle size analysis. Our repeated tests showed that if the obscuration of light in the first stage of the aggregate stability test was between 1 and 3 %, then the final post-sonication (see below) obscuration value was less than the maximum limit (18%).

The LG instrument (Beckman Coulter LS 13320, Brea, CA (USA)) was switched on at least two hours prior to analysis to ensure that it was at its operating temperature. Once the instrument had been prepared to measure particle size, a small aliquot (0.7 g) of each air-dried aggregate soil sample is weighed into a weighing boat prior to adding some or all of this aliquot to the reverse osmosis (RO) water circulating through the aqueous vessel of the LG instrument (see Figure 6). The instrument operator observes how the obscuration value changes and progressively adds more of the aliquot from the

weighing boat until between 2 and 3% obscuration is achieved. The LG instrument is then rinsed with RO water (16–17 M Ω ; consistent temperature between 19 and 21°C). The RO water contains very little excess CO₂ which can de-gas from mains water forming bubbles that can cause measurement errors. The operator then determines the mass of aggregates that were added to the aqueous vessel by difference. Based on repeated tests we have found that only a quarter of this mass of dry aggregates is required for use in pre-wetting because wet aggregates fragment more easily in the initial stage of the AS test, giving larger obscuration values.

Another aliquot of the soil specimen of this re-calculated mass is weighed into a weighing boat for the pre-wetting procedure. This involves placing two filter papers inside a clean (92 mm diameter) plastic petri dish. The first filter paper (11 μ m pore size, 90 mm diameter, cellulose No. 42 Whatman, GE Life Sciences UK), is placed on the base of the petri dish. The second filter paper (0.1 μ m, 47 mm diameter, cellulose nitrate, Whatman, GE Life Sciences UK) is placed centrally on top of the first filter. We use this arrangement so that water added to the lower paper is absorbed by the upper paper, and then by the aggregates which are placed onto the upper paper. To transfer the aggregates to the aqueous vessel of the LG, we remove only the smaller (upper) filter paper.

An accurate volume of 1.0 ml of reverse osmosis (RO) water is applied to the first filter paper. The calculated mass of aggregates of air-dried soil aggregates (1-2 mm) were then added to the upper filter paper ensuring that each makes contact with the surface of the paper, and the petri dish lid is placed on top. The aggregates were then allowed to wet by capillary action for 30 minutes; testing the impact of various pre-wetting time periods was beyond the scope of this paper. We found that by using this procedure on 3 samples from each PM it was possible to determine the mean sample mass required for the set of PMs under study, thus reducing the overall time required for the step.

To undertake the stability test, the pump speed on the LG instrument was set

725 to 60 which equates to a flow rate of around 45 ml per second based on measurements
726 we had undertaken at a range of pump speeds. Tests showed that at pump speeds
727 of 60, large aggregates do not settle at the base of the aqueous vessel which could
728 lead to biased measurements of psd. In contrast to our original procedure (Rawlins
729 *et al.*, 2013) the pump speed is not altered at any time during the aggregate stability
730 test. After pre-wetting, the aggregates were transferred to the aqueous vessel of the
731 LG instrument by using a pair of clean stainless steel forceps to remove the upper
732 filter paper from the lower one, and immersing both the upper filter paper and soil
733 aggregates into the RO water in the aqueous vessel. The circulating water removes
734 the aggregates from the filter paper and the filter paper can then be removed from the
735 aqueous vessel.

736 The operator allows exactly 30 seconds to elapse before starting the first particle
737 size analysis. This allows time for aggregate fragmentation by the circulating water. In
738 common with other aggregate stability methods (Zhu *et al.*, 2009), we cannot quantify
739 the mechanical energy leading to the breakdown of the initial aggregates. In our
740 method, this energy comprises: (i) the hydrodynamic forces of the circulating water,
741 (ii) collisions between soil aggregates and the surfaces of the circulating system, and
742 (iii) particle–particle collisions during circulation. The psd of the aggregates is then
743 determined by the LG instrument for 30 seconds based on the volume proportion of
744 particles in 92 size classes from 0.375 μm to 2000 μm . The obscuration value displayed
745 by the LG instrument is recorded.

746 The change in temperature of the water in the aqueous vessel was then recorded
747 from a temperature probe (Traceable digital thermometer, Control Company, USA)
748 inserted into the aqueous vessel. A UP100H sonicator (Hielscher, Germany) was placed
749 inside the aqueous vessel and maximum power (100 W) applied for a period of five
750 minutes. This causes fragmentation of the aggregates leaving only the fundamental
751 particles. In previous work it was shown that the post-sonication psd of samples from
752 two soil types was similar to the psd after removal of organic matter using hydrogen

peroxide (Rawlins *et al.*, 2013). The temperature of the water in the aqueous vessel was also recorded after sonication. The change in water temperature can be used to compute the power used in heating for each test to check that approximately the same quantity of energy is delivered by the sonicator in each analysis (Rawlins *et al.*, 2013).

The psd of the fundamental particles was then determined again for 30 seconds by the LG instrument using the same size classes as in the first analysis, and the final obscuration value was recorded. If the final obscuration value was greater than 18% the test was repeated with a smaller mass of aggregate material in the pre-wetting stage. The aqueous vessel on the instrument was rinsed thoroughly to remove soil particles before starting the next test. Based on repeated testing of the AS procedure using cultivated soils we found there were very few stable aggregates with diameters $> 500 \mu\text{m}$; any large fragments tended to be individual coarse soil particles. We found that these coarse fragments caused large variations in the MWD value because they dominate the volume measurements. We therefore chose to rescale each of the two size distributions to estimate DR for particle diameters $< 500 \mu\text{m}$. The rescaled MWD of each of the two psd measurements was then computed using:

$$\text{MWD} = \sum_{i=1}^n \bar{x}_i w_i, \quad (3)$$

where \bar{x}_i is the mean diameter of each size fraction (μm), and w_i is the volume proportion (expressed as a decimal proportion) of the sample corresponding to that size fraction. The MWD (μm) of the second psd analysis (fundamental particles) was then subtracted from the MWD of the first (aggregates) to compute the disaggregation reduction (DR) value (μm). We found that taking the average DR value of two aggregate stability tests on separate aliquots of each soil sample provides a more robust measurement. For DR measurement of the palaeosol reference material we used the same procedure as above without the pre-wetting stage; 0.15 g of air-dry aggregates (250-500 μm) were added directly to the aqueous vessel. We also used two standard particle size materials (supplied by Beckman Coulter; mean diameter 32 and 500 μm) throughout the series of analysis to check for accuracy and precision of the psd mea-

780 surements by the LG instrument.

Figure 8: Arrangement of apparatus used for the measurement of aggregate stability (not to scale). Arrows depict the circulation of the suspension between the aqueous vessel and the laser granulometer.

