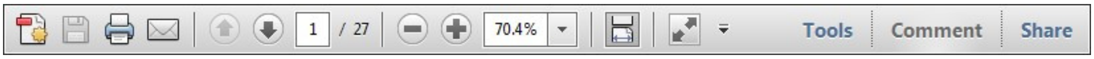
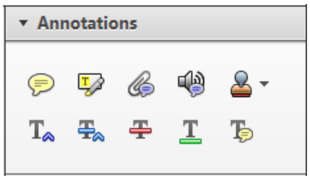
WILEY

Once you have Acrobat Reader open on your computer, click on the Comment tab at the right of the toolbar:



This will open up a panel down the right side of the document. The majority of tools you will use for annotating your proof will be in the Annotations section, pictured opposite. We've picked out some of these tools below:

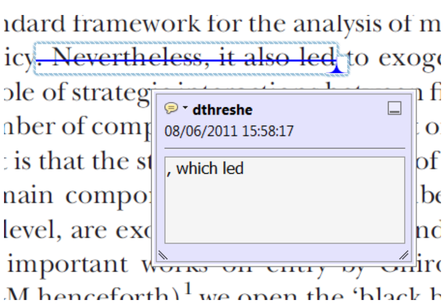


1. Replace (Ins) Tool – for replacing text.

Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it

- Highlight a word or sentence.
- Click on the Replace (Ins) icon in the Annotations section.
- Type the replacement text into the blue box that appears.

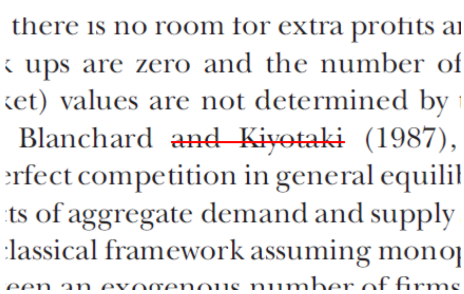


2. Strikethrough (Del) Tool – for deleting text.

Strikes a red line through text that is to be deleted.

How to use it

- Highlight a word or sentence.
- Click on the Strikethrough (Del) icon in the Annotations section.

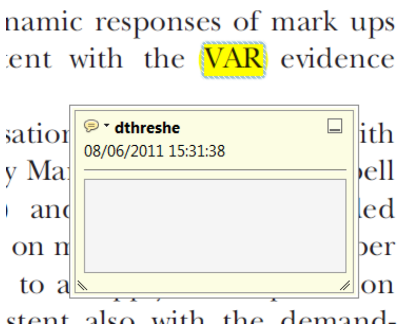


3. Add note to text Tool – for highlighting a section to be changed to bold or italic.

Highlights text in yellow and opens up a text box where comments can be entered.

How to use it

- Highlight the relevant section of text.
- Click on the Add note to text icon in the Annotations section.
- Type instruction on what should be changed regarding the text into the yellow box that appears.

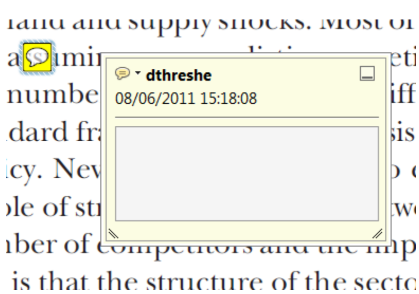


4. Add sticky note Tool – for making notes at specific points in the text.

Marks a point in the proof where a comment needs to be highlighted.

How to use it

- Click on the Add sticky note icon in the Annotations section.
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the yellow box that appears.

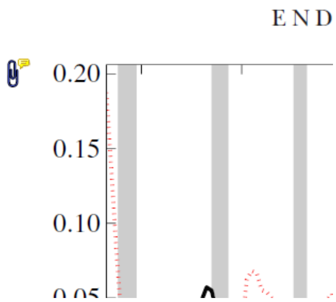


5. Attach File Tool – for inserting large amounts of text or replacement figures.

Inserts an icon linking to the attached file in the appropriate pace in the text.

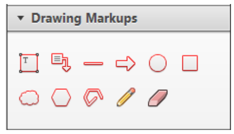
How to use it

- Click on the Attach File icon in the Annotations section.
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.



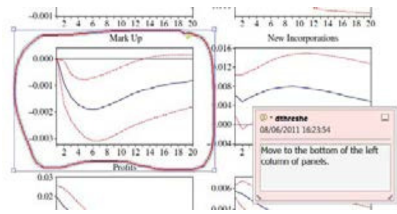
6. Drawing Markups Tools – for drawing shapes, lines and freeform annotations on proofs and commenting on these marks.

Allows shapes, lines and freeform annotations to be drawn on proofs and for comment to be made on these marks.



How to use it

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, move the cursor over the shape until an arrowhead appears.
- Double click on the shape and type any text in the red box that appears.



SUM	12117	WILEY	Dispatch: 3.4.14	CE: Gayathri T.
Journal Code	Manuscript No.		No. of pages: 9	PE: Radhika

Application of a visual soil examination and evaluation technique at site and farm level

M. P. W. Sonneveld[1,†], G. B. M. Heuvelink[1] & S.W. Moolenaar[2]

[1]*Soil Geography and Landscape group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands, and*

[2]*Stichting Kennisoverdracht Bodem, Gouda, PO Box 420, 2800 AK Gouda, Gouda, The Netherlands*

Abstract

Visual soil examination and evaluation (VSEE) techniques are semi-quantitative methods that provide rapid and cost-effective information on soil quality. These are mostly applied at site or field level, but there is an increased need for soil quality indicators at farm level to allow integration with other sustainability indicators. The objectives of this study were to develop and apply a protocol for application of a VSEE technique at site level, to assess the VSEE observations against standardized laboratory analyses and to aggregate VSEE observations to farm level using an appropriate sampling design. The study was conducted at ten dairy farms in a reclaimed polder in the Netherlands with clay and organic soils. A stratified random sampling design was used to account for spatial variability in land use and soil series. VSEE was carried out using the Visual Soil Assessment approach. Results show that 81% of sites were assessed as good and the remainder as moderate to poor. For the clay soils, field observations of soil structure were significantly correlated with pH, bulk density, soil organic matter (SOM) and mean weight diameter of aggregates, whereas for organic soils, soil structure significantly correlated with pH, bulk density, organic C and SOM. The range in overall scores calculated at farm level was smaller than at site level, and most farms were assessed as good.

Keywords: Soil structure, aggregates, soil organic matter, soil quality, indicators

Introduction

The agronomic and environmental performance of most agricultural systems is strongly related to the natural capital of the soil, soil functioning and resulting soil ecosystem services (Dominati *et al.*, 2010). Soil quality can be defined as the fitness of soils to function within natural or managed ecosystem boundaries, to sustain plant and animal productivity, to maintain or enhance water and air quality, and to support human health and habitation (Karlen *et al.*, 1997). Assessment of soil quality is required for evaluation of the overall sustainability of agricultural systems, identification of areas with problems for production, estimation of biomass production and monitoring changes in environmental quality related to agricultural management (Doran & Parkin, 1996). Identification of appropriate soil quality indicators is a first step in the assessment and evaluation of the contribution of soils in the delivery of ecosystem goods and services (Robinson *et al.*, 2012). Thus, the need is to develop indicators for on-site assessment of soil quality by farmers, researchers and extension personnel (Doran & Parkin, 1994). Visual soil examination and evaluation (VSEE) techniques are known to be cost-effective, practical and to provide rapid results (Boizard *et al.*, 2005; Mueller *et al.*, 2009; Ball *et al.*, 2013). A semi-quantitative method to visually examine and evaluate soil quality, in particular soil structure, in the field was proposed by Peerlkamp (1967) and was later modified by Ball *et al.* (2007). Other VSEE techniques include the Visual Soil Assessment (VSA, Shepherd, 2000, 2010b) and the SOILpak assessment (McKenzie, 2012).

Working with soil quality indicators, such as those resulting from VSEE techniques, requires an awareness both of the scale at which observations are made and at which answers are needed (Pieri *et al.*, 1995; Hoosbeek & Bouma,

Correspondence: G. B. M. Heuvelink. E-mail: gerard.heuvelink@wur.nl

†Marthijn Sonneveld sadly passed away in December 2013. He fulfilled his role as first author until the very last stage of preparation, leaving only minor revisions to be made.

Received May 2013; Revised February 2014; accepted after revision February 2014
Editor: M. Goss

1998). There are few published examples of practical soil quality indicators at the farm level (e.g. Bélanger *et al.*, 2012), although the development of sustainable agricultural production systems is especially relevant at this level. On-farm assessment of soil quality using simple and rapid techniques is required to evaluate the overall impact of a new farm management system or to identify potential soil problems at farm level (Sarrantonio *et al.*, 1996). A single integrated soil quality indicator at farm level can also be integrated with other indicators at farm level (Rigby *et al.*, 2001), for example whole-farm nutrient balances (Schröder *et al.*, 2003), energy consumption and emission of CO_2^- equivalents (Thomassen & De Boer, 2005) and other management information (Bélanger *et al.*, 2012). Payments for ecosystem services as part of agricultural policies also address soil quality at farm level.

Developing a simple and practical soil quality indicator at farm level necessitates the application of a suitable spatial sampling scheme, because soil quality assessment must be derived from a limited number of locations on the farm. To avoid bias and subjective decisions as well as quantify the estimation of errors with confidence intervals, sampling needs to be based on statistical sampling theory, ideally taking spatial variation in land use and soil series into account. It is also appropriate to use standardized soil analyses to assess VSEE outcomes as VSEE techniques yield semi-quantitative outcomes which are only proxies of the true soil status.

The objectives of this study were thus to develop and apply a protocol for application of a VSEE technique at site level, to assess these site results against standardized laboratory analyses and to aggregate VSEE observations at site level into an indicator at farm level using an appropriate sampling design.

Materials and methods

Study area

The study area was in the province of North-Holland, the Netherlands, and included the Schermer and the Beemster polders (Figure 1). Both polders were reclaimed from lakes in the early 17th century. Beemster is an exceptional example of reclaimed land in the Netherlands and is a UNESCO World Heritage Site. It has a well-ordered landscape of fields, roads, canals, dykes and settlements, laid out according to classical and Renaissance planning principles. Land use in both polders was dominated by agriculture with grassland being the most important land cover in Beemster, with arable crops, including potatoes and cereals, being the most important land cover in Schermer. Several dairy farmers had implemented rotation schemes on their land involving the cultivation of silage maize and flower bulbs (mainly tulips). For tulips, the bulbs are planted after land is tilled to a depth of 15 cm in autumn. Harvesting of the bulbs takes place in summer. The land is ca. 4 m below sea level, and drainage is artificial. In some locations, subsoil ripening, a pedogenetic process that is related to the dehydration of soft sediments, has not been fully completed. Soils are dominantly Fluvisols with locally very large topsoil clay percentages (>40%), but there are small areas of organic soils (Histosols), especially along the borders of both polders. Many topsoils have been enriched with a peaty mud, which is related to the original peat covering the area. Soils were assessed visually and samples collected on ten dairy farms in the study area. Addresses of these farms were provided by a regional dairy processing company.

VSEE at site level

On the selected sites, a soil sample of 20 × 20 × 20 cm^3 was collected for visual examination in spring or early summer when soil conditions were not too wet and not too dry (Mueller *et al.*, 2009). Visual examination was based on the VSA approach (Shepherd, 2010b) but was slightly modified to include only those indicators that were considered relevant to the study area. Other indicators such as soil erosion were omitted as they were not considered relevant. The selected indicators were soil colour, soil structure, soil mottles, porosity, rooting depth, earthworms, surface cover and tillage pan. For assessing rooting depth and number and colour of soil mottles, the pit created by the removal of the large soil sample was extended to a depth of 50 cm. For each indicator, three classes were defined: 0 (poor condition), 1 (moderate condition) and 2 (good condition). Intermediate figures were also possible, yielding a total of five possible scores. The applied weighing factors, given in Table 1, were derived from Shepherd (2010b). Final soil scores were calculated relative to the maximum soil score (42) and expressed on a 0–100 scale. Final scores below 50 were assessed as poor, scores between 50 and 75 as moderate and scores above 75 as good.

Colour assessments were made using a Munsell colour chart, assuming a relationship between soil colour and soil organic matter (SOM) or soil organic carbon (Wills *et al.*, 2007). To define the classes for soil colour, topsoil data from the national soil survey database were used for the soil series in the study area. This yielded soil organic matter data for 111 data points for topsoils with Hue 10 yr and 35 data points for topsoils with Hue 2.5 yr. A single-factor ANOVA was applied to the 10-yr dataset using colour Value to group the data. This showed that mean SOM content was significantly different between the groups, ranging for the most common colour Values from 8.8% (Value 3, n = 50), 7.1% (Value = 3.5, n = 14) to 5.69% (Value = 4, n = 11). Mean SOM contents were much less for Values 4.5 and 5 (1.85% and 1.5%, respectively). For the 2.5-yr data, mean SOM content was also significantly different between groups.

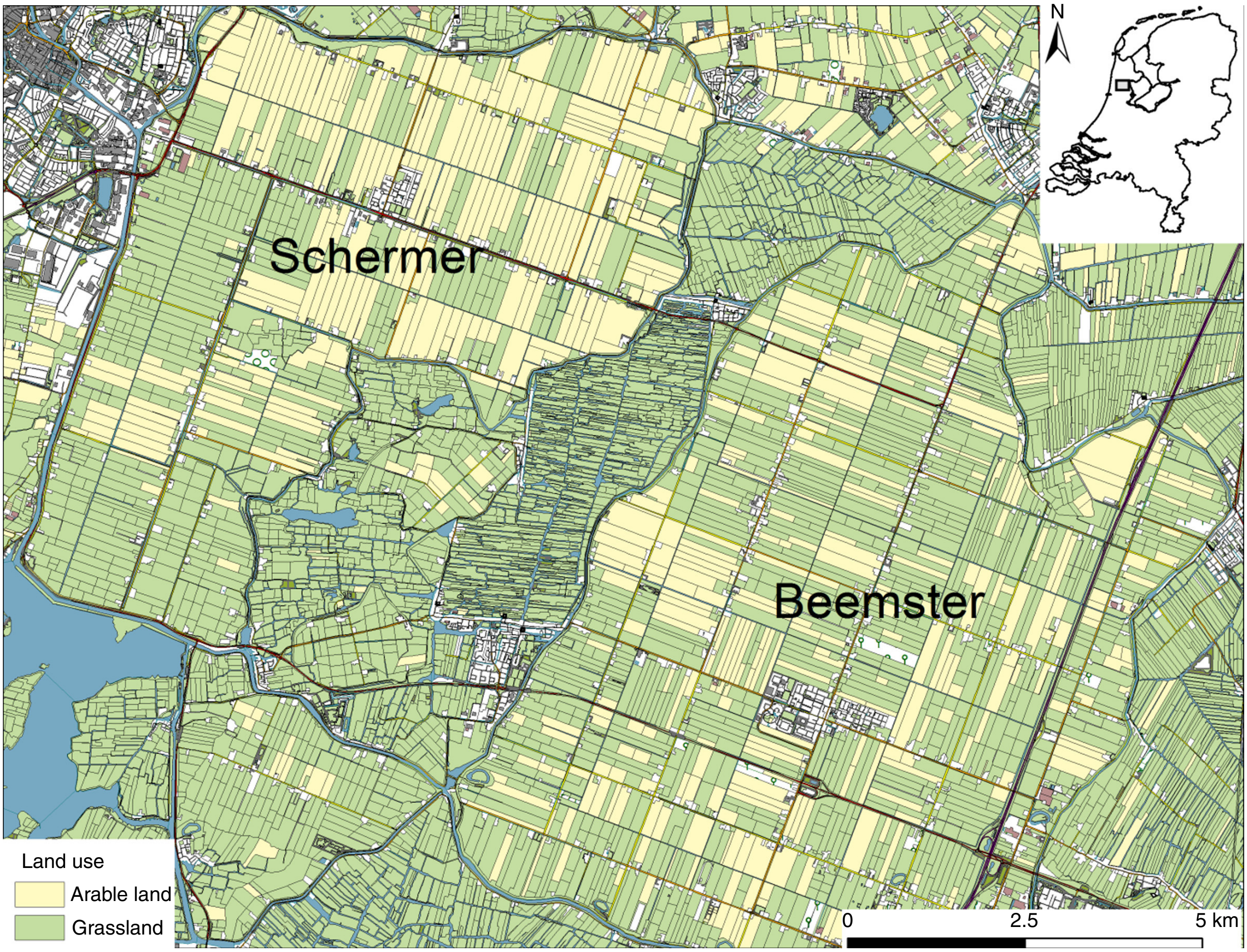


Figure 1 Land use map and location of the Schermer and Beemster polders in the Netherlands.

Table 1 Selected VSA criteria and corresponding weighing factors for this study

Criterion (scores: 0, 1 and 2)	Weighing factor
Soil structure	3
Earthworms	2
Soil porosity	3
Root development	3
Number and colour of soil mottles	3
Surface cover	2
Tillage pan	3
Soil colour	2

Based on this, scores were defined as 0 (Munsell value >4), 1 (Munsell value = 4 or 3.5) and 2 (Munsell value = 3 or less). Ripening of the subsoil was classified as 1 (ripened) or 2 (nonripened), and derived from the soil map. Texture, also derived from the soil map and field estimates, was classified as 1 (clay% ≥25%) or 2 (clay% <25%). After all soil assessments were made, the farmer was interviewed to obtain information on land use history. Land use histories were then classified as 1 (current land use <5 yr) or 2 (current land use >5 yr).

Soil analyses

At each site, a 100 cm^3 core was collected at 10 cm depth, dried (105 C), weighed and then the dry bulk density was calculated. Subsequently, subsamples were taken, placed in crucibles, weighed and heated for 6 h at 550 C to determine SOM content through loss on ignition. These SOM values were not corrected for clay content, as this was not measured in this study. A bulk sample of five aliquots was collected in the field by haphazard removal of soil from an area of ca. 4 m^2 around the selected site. Bulk samples were dried (40 C), sieved (2 mm) and analysed for pH ($CaCl_2$, at 20 °C), total carbon (C/N Analyser) and inorganic carbon using a LECO RC612 analyser. Organic carbon was calculated as total carbon minus inorganic carbon.

Part of the excavated soil at the site was taken for analysis of aggregate size and aggregate stability. All aggregates were dried at 40 °C for ca. 1 week. Because of the hardening of the soil, large aggregates were broken up into smaller ones using a Pulverisette 2® mortar grinder without disturbing the natural aggregates. The dry-sieving method was adopted to calculate the mean weight diameter (MWD):

2

$$\mathrm{MWD} = \sum\nolimits_{l=1}^{L} \bar{x}_l \times s_l \tag{1}$$

where L is the total number of sieves, $\bar{x}_l$ is the mean diameter of the lth sieve and sl is the corresponding weight fraction. Applied sieves had mesh sizes of 4, 3.4, 2, 1, 0.5 and 0.25 mm. Samples were shaken for 5 min at a frequency of 210 cycles per minute. Soil remaining on each sieve was collected and weighed, and the MWD was then calculated. The soil collected in the 1 mm sieve was used for determining aggregate stability using an Eijkelkamp wet sieving apparatus. Of the 1 mm air-dried aggregates, a 4.0 g subsample was weighed out and put into a sieve can with 250 μm sieves. Samples were then moved up and down in distilled water at a frequency of about 36 times/min for 3 min. Collected unstable soil aggregates were dried for 24 h at 105 °C and then weighed. Stable aggregates were again
3 sieved using a dispersion solution (NaPO3)6 and separated from other elements (e.g. roots). The aggregate stability index (ASI) was calculated by dividing the weight of stable soil aggregates over the sum of the stable and unstable aggregate weights. The procedure was repeated for each individual sample, and the mean ASI of two samples was calculated.

Aggregation of VSEE site scores to farm level

To aggregate the VSEE site scores to a VSEE farm score, a design-based statistical approach was taken which involved probability sampling of sites within the farm. In this case, the mean farm VSA score was estimated using stratified simple random sampling (De Gruijter *et al*., 2006). The domain, or target population, was defined as the topsoil of all registered maize and grassland fields on a farm excluding 5-m-wide buffer zones around the field edges. Land that was not owned by the farm, for example rented land, was also excluded, and so was land that was not at least 5 yr in possession of the farmer prior to sampling. This was to ensure that the VSA score reflected the management of the farmer. For practical reasons, the maximum number of site assessments for each farm was limited to eight.

Each farm domain was divided into strata based on soil series and land use. Farm field data were obtained from the national field registration database (Schmidt *et al*., 2003), and soil survey data from the Dutch Soil Information System (De Vries *et al*., 2003). Unique strata consisting of the same soil series and land use were constructed in ArcMAP 10.0 (ESRI 199–2010). In each stratum, a number of sites were randomly selected, with the number of sites per stratum chosen proportional to the stratum size, with a minimum of two sites per stratum.

Because the maximum number of site assessments was limited to eight, the maximum number of strata per farm was limited to four. Farms with more than four strata were treated by randomly selecting four strata and analysing two randomly selected sites for each. In this case, the design is no longer a stratified random sampling design, but a two-stage random sampling design which has implications for the associated statistical inference (De Gruijter *et al*., 2006). As this only applied to a few cases, we did not describe the statistical inference of this design and limit the description below to that of stratified simple random sampling.

For each farm, the estimated mean VSA score was calculated as:

$$\hat{\bar{z}} = \sum\nolimits_{h=1}^{H} \hat{\bar{z}}_h \times w_h \tag{2}$$

where $\hat{\bar{z}}$ is the estimated mean farm VSA score, H is the number of farm strata, $\hat{\bar{z}}_h$ is the estimated VSA mean of stratum h and w_h is the weight of stratum h. This weight was calculated as:

$$w_h = A_h / A \tag{3}$$

where A_h is the area of stratum h and A is the total area of the domain, that is the registered maize and grassland owned by the farmer, excluding field buffer zones.

For the h^{th} stratum, the estimated mean VSA score $\hat{\bar{z}}_h$ was calculated by taking the arithmetic mean of VSA site scores:

$$\hat{\bar{z}}_h = \frac{1}{M_h} \sum\nolimits_{m=1}^{M_h} z_{mh} \tag{4}$$

M_h is the number of sites within stratum h, and z_{mh} is the VSA score of site m in stratum h. The variance of the mean farm VSA score estimation error was estimated as:

$$\hat{V}(\hat{\bar{z}} - \bar{z}) = \sum\nolimits_{h=1}^{H} w_h^2 \times \hat{V}(\hat{\bar{z}}_h - \bar{z}_h) \tag{5}$$

where $\hat{V}(\hat{\bar{z}}_h - \bar{z}_h)$ is the estimated variance of the VSA estimation error for stratum h. This was calculated as:

$$\hat{V}(\hat{\bar{z}}_h - \bar{z}_h) = \frac{1}{M_h \times (M_h - 1)} \sum\nolimits_{m=1}^{M_h} (z_{mh} - \hat{\bar{z}}_h)^2 \tag{6}$$

The standard error of the estimated mean was taken as the square root of the variance of the mean VSA score estimation error. Lower and upper limits of the 95%

symmetric confidence interval were computed from the estimates and standard errors assuming a normal distribution.

The number of sites, M_h, to be selected in stratum h was calculating using:

$$M_h = max(2, \lfloor 8 \times w_h \rfloor) \qquad (7)$$

where $\lfloor .. \rfloor$ means rounding to the nearest integer smaller or equal than its argument. Note that the effect of rounding was not included in the computation of the standard error of the estimated mean farm VSA score, implying that these results are only approximate.

Results

VSEE at site level

Of the 80 sites for which observations were made, seven were organic soils and 73 were mineral soils. Land use at most of the sites with mineral soils was grassland (n = 63), followed by silage maize (n = 8) and flower bulbs (n = 2). Most sites (n = 59) with mineral soil were clay soils (clay % >25%), and the remaining mineral soils (n = 14) were loamy soils (8% < clay% <25%). Of the 73 sites with mineral soils, 46 had ripened subsoils and 27 unripened subsoils. VSA scores for all individual sites ranged between 43 and 96 with a mean VSA score of 81. The distribution of the VSA scores is given in Figure 2. It appeared that 65 sites were assessed as good (81%), 14 as moderate (18%) and one as poor (1%). One sample was omitted from the subsequent farm analysis because of insufficient data.

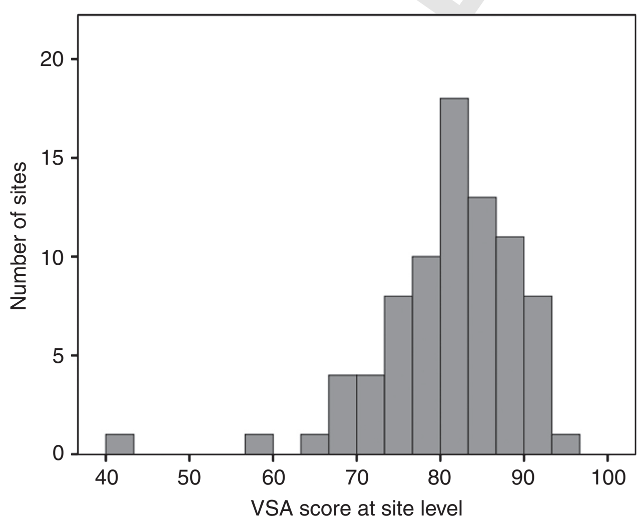


Figure 2 Frequency distribution of VSA scores at site level for all sites.

Relating VSEE results at site level to soil analyses

Descriptive statistics for analyses on soil samples from all sites are given in Table 2. Due to laboratory mistakes, the analysed number of soil samples for bulk density and SOM for the mineral soils was slightly less than for other soil variables. For the mineral soils, SOM ranged between 8% and 19% and most sites received a VSA colour score of 2 (n = 65). Only six sites had a colour value of 1. Mean SOM for colour score 1 (10.1%) was significantly ($P < 0.05$) smaller than mean SOM for colour score 2 (12.0%). Differences between mean values for organic C for both colour scores were not significant.

Individual VSA observations were assessed by calculating Spearman correlation coefficients against measured values of soil properties. Results for the mineral clay soils are given in Table 3. Site VSA scores were significantly negatively correlated with pH ($P < 0.05$). Tillage pans were not observed within 50 cm of the soil surface, so scores for this indicator were always the greatest (2), so VSA scores were not included in Table 3. A significant positive correlation with SOM content was only found for visual scores of soil structure ($P < 0.01$), but negative relations were obtained with pH ($P < 0.01$), bulk density ($P < 0.05$) and MWD ($P < 0.01$). Visual scores for soil porosity or earthworms were not significantly correlated with any of the measured properties. However, visual scores for earthworms were positively correlated with age and age class ($P < 0.01$). As

Table 2 Descriptive statistics for sampling sites with mineral and organic soils

	Number of soil samples	Mean	St. dev.	Min	Max
Mineral soils					
N total (g/kg)	73	4.23	0.90	2.44	7.70
Organic C (g/kg)	73	44.6	8.9	30.3	77.0
pH	73	6.72	0.61	4.84	7.47
Bulk density (g/cm^3)	72	1.07	0.11	0.78	1.30
SOM (%)	71	11.9	2.8	8.2	18.7
Mean ASI	73	0.50	0.03	0.39	0.59
MWD	73	3.00	0.30	1.85	3.46
Organic soils					
N total (g/kg)	7	7.68	2.36	4.94	10.70
Organic C (g/kg)	7	95.4	33.3	57.1	144.0
pH	7	4.96	0.41	4.17	5.37
Bulk density (g/cm^3)	7	0.79	0.17	0.63	1.07
SOM (%)	7	21.4	7.4	14.1	33.7
Mean ASI	6	0.56	0.20	0.43	0.97
MWD	7	2.92	0.62	1.83	3.42

Table 3 Spearman correlation coefficients between site properties, individual VSA scores and other variables at site level for mineral soils

	Age	Age Class	Texture	Subsoil	VSA structure	VSA porosity	VSA earthworms	VSA mottles	VSA roots	VSA surface cover	VSA colour	Site VSA score
Age Class	0.869**											
Texture	0.127	0.050										
Subsoil	−0.213	−0.174	−0.157									
VSA structure	−0.082	−0.137	0.415**	−0.096								
VSA porosity	0.223	0.182	0.171	−0.101	0.208							
VSA earthworms	0.343**	0.315**	−0.174	0.148	0.096	0.015						
VSA mottles	0.359**	0.291*	−0.192	−0.442**	−0.162	−0.038	−0.023					
VSA roots	−0.047	0.007	−0.034	0.061	0.180	0.186	0.186	−0.047				
VSA surface cover	−0.216	−0.287*	0.146	−0.076	0.093	−0.001	−0.110	−0.094	−0.072			
VSA colour	−0.369**	−0.287*	0.146	0.229	0.100	−0.148	−0.157	−0.223	0.156	0.099		
VSA score site level	0.271*	0.215	0.114	−0.124	0.612**	0.496**	0.535**	0.269*	0.338**	0.075	0.057	
pH	0.202	0.185	−0.247*	−0.132	−0.447**	−0.016	−0.051	0.314**	−0.173	−0.109	−0.349**	−0.264*
Organic C	−0.125	−0.014	−0.178	0.039	0.185	0.118	−0.073	−0.067	0.043	−0.225	0.144	0.104
Bulk density	0.340**	0.292*	−0.027	−0.193	−0.275*	0.021	0.200	0.173	−0.280*	0.078	−0.382**	−0.042
SOM	−0.187	−0.103	0.004	0.139	0.355**	0.102	−0.006	−0.161	0.294*	−0.166	0.262*	0.222
ASI	0.060	0.096	−0.210	0.023	−0.127	0.039	0.088	0.035	−0.029	−0.110	0.147	0.036
MWD	0.146	0.094	−0.381**	0.220	−0.348**	0.052	0.158	0.076	0.053	0.006	−0.230	−0.054

*Correlation is significant at the 0.05 level; **Correlation is significant at the 0.01 level.

almost all non-grassland land use classes (that is maize and flower bulbs) had ages <5, this indicates that the number of earthworms increases as the age of the grassland increases. Visual scores for mottles were positively correlated with pH ($P < 0.01$) but also with age ($P < 0.01$) and age class ($P < 0.05$), and negatively with the ripening characteristics of the subsoil ($P < 0.01$). The latter supports our expectation that more hydromorphic features (low VSA scores) are associated with poorly drained site conditions. Visual scores for roots were significantly negatively correlated with bulk density ($P < 0.05$), implying that greater bulk densities are associated with fewer roots. VSA scores for roots were also positively correlated with SOM ($P < 0.05$), but not with organic C content of the soil. Surface cover was negatively correlated with age class ($P < 0.05$). Soil colour score was significantly negatively correlated with pH and bulk density ($P < 0.01$), and positively with SOM ($P < 0.05$). There were no significant correlations between the mean ASI and any of the other VSA observations or measurements.

Spearman correlation coefficients between individual VSA observations, calculated VSA scores at site level and other soil analyses for the organic soils ($n = 7$) are given in Table 4. VSA scores at site level for the organic soils were significantly positively correlated with age ($P < 0.05$), but not with any of the measured soil properties. Individual VSA scores for soil structure were negatively correlated with pH and bulk density, but positively with organic C and SOM ($P < 0.05$). They were also positively correlated with age ($P < 0.01$). Individual VSA scores for earthworms were significantly positively correlated with MWD ($P < 0.01$), showing that earthworm activity affects soil structural morphology and contributes to macro-aggregation (Castellanos-Navarrete *et al*., 2012). Individual scores for mottles were significantly negatively correlated with mean ASI ($P < 0.05$), and individual scores for roots and surface cover were significantly negatively correlated with pH ($P < 0.05$).

VSEE at farm level

Estimated VSA scores for individual farms ranged from 74 to 87 with a mean score of 81. Thus, all farms were assessed as good, with the exception of one farm. Calculated standard errors of the estimated means at farm level ranged from 0.6 to 6.1. The distribution of VSA scores at farm level as given in Figure 3 shows that overall scores at farm level were confined to a narrower range than were scores at site level. Confidence intervals of VSA score estimates at farm level (Table 5) were especially small for farms 2 and 10, and large for farms 1 and 3. Overall, the confidence intervals were sufficiently narrow to detect meaningful differences between VSA scores of individual farms. For instance, farm 10 had a much larger score than farm 4, and the difference cannot be attributed to sampling errors.

Table 4 Spearman correlation coefficients between site properties, individual VSA scores and other variables at site level for organic soils

	Age	Subsoil	VSA structure	VSA porosity	VSA earthworms	VSA mottles	VSA roots	VSA surface cover	Site VSA score
Subsoil	−0.725								
VSA structure	0.989**	−0.717							
VSA porosity	0.862*	−0.661	0.833*						
VSA earthworms	−0.144	0.242	−0.022	0.000					
VSA mottles	0.011	0.322	0.099	−0.335	0.300				
VSA roots	0.882**	−0.548	0.872*	0.805*	−0.088	0.176			
VSA surface cover	0.882**	−0.548	0.872*	0.805*	−0.088	0.176	1.000**		
Site VSA score	0.804*	−0.364	0.815*	0.872*	0.281	0.070	0.798*	0.798*	
pH	−0.896**	0.577	−0.867*	−0.600	0.418	−0.159	−0.791*	−0.791*	−0.541
Organic C	0.837*	−0.577	0.867*	0.473	−0.100	0.319	0.632	0.632	0.505
Bulk density	−0.777*	0.577	−0.788*	−0.636	−0.020	−0.020	−0.474	−0.474	−0.703
SOM	0.837*	−0.577	0.867*	0.473	−0.100	0.319	0.632	0.632	0.505
ASI	−0.278	−0.207	−0.334	0.087	−0.062	−0.828*	−0.207	−0.207	−0.257
MWD	0.100	0.144	0.217	0.109	0.896**	0.319	0.000	0.000	0.396

*Correlation is significant at the 0.05 level; **Correlation is significant at the 0.01 level.

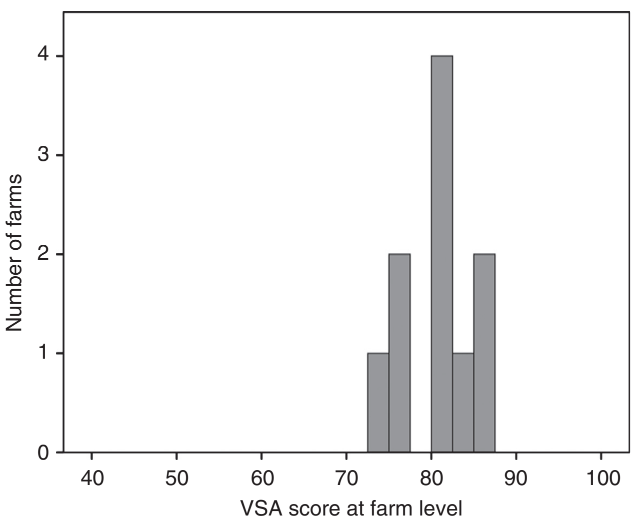


Figure 3 Frequency distribution of estimated VSA scores at farm level.

Table 5 Estimated scores at farm level and confidence intervals (C.I.)

Farm	Estimated mean farm score	Lower limit 95% C.I.	Upper limit 95% C.I.
1	77.5	72.4	80.6
2	80.5	79.8	81.2
3	80.4	75.9	85.0
4	82.0	80.0	83.9
5	74.0	72.7	75.1
6	82.1	79.8	84.3
7	77.0	75.8	78.2
8	85.2	84.8	85.6
9	84.1	82.1	86.2
10	86.9	86.1	87.6

Discussion

VSEE at site level

VSA structure at site level was strongly correlated with measured soil properties, including pH, bulk density, SOM and MWD. This supports the hypothesis that semi-quantitative VSEE techniques are a useful tool in examining soil quality. Yet, from their very nature, VSEE techniques involve some degree of subjectivity and are not a real quantification of soil properties (Guimarães *et al*., 2013; Newell-Price *et al*., 2013). Disruption of a 20 × 20 × 20 cm^3 soil block into aggregates seems particularly dependent on the operator. The drop-shatter procedure as originally defined in the VSA procedure (Shepherd, 2000) did not produce satisfactory results in our study, most likely as a result of the large clay contents (>35%). Instead, individual aggregates were separated using gentle pressure by hand. VSA scores at different sites were assessed independently by different observers (data not shown) to examine observer influence. This yielded maximum differences in VSA scores of 1.

Mottles are most often a result of hydromorphic conditions. Farmers often tend to keep such wetter sites in pasture, which explains the correlation between mottles, and age and age class, as in this study, increased ages are mostly associated with grassland. Mottles were also strongly correlated with the degree of ripening of the subsoil, which shows that this diagnostic property in Dutch soil survey classification is relevant to soil hydrology. For both mineral soils and organic soils, we found that pH was strongly correlated with either individual VSA scores or the overall site VSA score. Given that pH also correlated with other measured soil properties suggests that it is relevant to include an on-site pH determination within a VSEE protocol for similar regions.

VSEE at farm level

Overall scores at farm level are generally higher than at site level because the farms have mostly larger proportions of grassland, which would have influenced positively the final score. One farm was evaluated as moderate. Closer inspection of this farm revealed that of the investigated grasslands, the maximum age was 7 yr. This particular farm had adopted a high degree of crop rotation, including flower bulbs and maize, which had a negative impact on the overall farm score.

Most soil indicators in soil survey databases are related to inherent or relatively static soil properties. This results from the principles behind soil taxonomic systems, because the outcome of a soil classification should not change as a 'result of a single plowing' (De Bakker, 1970). Soil change due to land use and soil management was specifically left out of soil taxonomic systems. Although existing soil survey data can be used for regional systematic land evaluation procedures (Sonneveld *et al.*, 2010), these databases are unsuitable to provide indicators related to soil ecosystem services at the field or farm level. In addition to inherent properties, manageable soil properties (Dominati *et al.*, 2010) are typically organic matter content, bulk density and aggregate size. In this study, the visual soil quality indicators also included such soil properties. Thus, the integration of visual examination techniques at farm level with conventional soil survey information shows how soil change may be linked to existing spatial soil databases, essentially allowing the documentation of phenoforms for any given genoform that is the pedogenetically defined soil series (Sonneveld *et al.*, 2002).

VSEE, soil quality and ecosystem services

Most VSEE techniques have been developed to evaluate soil quality, notably soil structure, in relation to soil management and crop production. For a comprehensive evaluation of the environmental impact of farming systems, indicators of environmental risks must be defined, such as nutrient losses to groundwater and surface waters and greenhouse gas emission (Van Der Werf & Petit, 2002). A guide for visual indicators of environmental performance is presented by Shepherd (2010a). Source indicators are derived from data on the use of fertilizers, animal manure and stocking density (Pieri *et al.*, 1995; Sonneveld *et al.*, 2012). Additional laboratory measurements need to include indicators related to the state of soil resources, such as soil P status, and stocks of soil C and N. VSEE outcomes, laboratory measurements and source indicators together determine the weight of evidence of an integrated soil quality assessment.

Conclusions

Application of the VSA technique at site level in our study showed that 81% of the sites with mineral soils were assessed as good and the remainder as moderate to poor. Field observations of soil structure were significantly correlated with pH, bulk density, soil organic matter content and mean weight diameter of the aggregates. Thus, visual observations of soil quality were validated using standardized laboratory analyses. Moreover, inherent soil properties derived from the soil map, such as texture and degree of subsoil ripening, were also significantly correlated with field observations of soil structure and mottles, respectively. This supports the view that stratification at farm level on the basis of existing soil series data is relevant in establishing overall farm scores. The range of estimated scores at farm level was less than at site level, and most farms (9 out of 10) were assessed as good. Calculated confidence intervals are in some cases large and may be reduced by increasing the number of observations. The approach as described can deliver a quick unbiased estimate of soil quality at farm level which may be used as input in overall sustainability assessments of agricultural systems.

Acknowledgements

This work was financially supported by the SKB programme Sustainable Development of the Subsurface in which both public and private partners are involved in knowledge development. SKB is the Dutch funder of the SNOWMAN Network: this work was carried out within the SNOWMAN project SAS-STRAT (www.snowmannetwork.com). The authors wish to thank Maricke van Leeuwen, Henrieke Paul, Nicole Meijer and Renee van Dongen for their assistance in preparing and performing the fieldwork. Laboratory staff of Wageningen University are thanked for their assistance in soil sample analyses.

References

Ball, B.C., Batey, T. & Munkholm, L.J. 2007. Field assessment of soil structural quality - A development of the Peerlkamp test. *Soil Use and Management*, **23**, 329–337.

Ball, B.C., Munkholm, L.J. & Batey, T. 2013. Applications of visual soil evaluation. *Soil and Tillage Research*, **127**, 1–2.

Bélanger, V., Vanasse, A., Parent, D., Allard, G. & Pellerin, D. 2012. Development of agri-environmental indicators to assess dairy farm sustainability in Quebec, Eastern Canada. *Ecological Indicators*, **23**, 421–430.

Boizard, H., Batey, T., Richard, G., Roger-Estrade, J., Ball, B.C., Cattle, S., Hasinger, G., Munkholm, L.J., Murphy, B.W., Nievergelt, ?.?.?., Peigné, J. & Shepherd, G. 2005. *Field Meeting visual soil structure assessment, held at the INRA research station*,
6 INRA, ?????, France, 25–27

Castellanos-Navarrete, A., Rodríguez-Aragonés, C., De Goede, R.G.M., Kooistra, M.J., Sayre, K.D., Brussaard, L. & Pulleman, M.M. 2012. Earthworm activity and soil structural changes under conservation agriculture in central Mexico. *Soil and Tillage Research*, **123**, 61–70.

De Bakker, H. 1970. Purposes of soil classification. *Geoderma*, **4**, 195–208.

De Gruijter, J.J., Brus, D.J., Bierkens, M.F.P. & Knotters, M. 2006. *Sampling for natural resource monitoring*. Springer, Berlin.

De Vries, F., De Groot, W.J.M., Hoogland, T. & Denne, J. 2003. *De Bodemkaart van Nederland digitaal; Toelichting bij inhoud, actualiteit en methodiek en korte beschrijving van additionel informatie*. Alterra, Wageningen, pp. 48.

Dominati, E., Patterson, M. & Mackay, A. 2010. A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecological Economics*, **69**, 1858–1868.

Doran, J.W. & Parkin, T.B. 1994. Defining and assessing soil quality. In: *Defining soil quality for a sustainable environment* (eds J.W. Doran, D.C. Coleman & D.F. Bezdicek), pp. 3–21. Soil Science Society of America, Madison, WI.

Doran, J.W. & Parkin, T.B. 1996. Quantitative indicators of soil quality: a minimum dataset. In: *Methods for assessing soil quality* (eds J.W. Doran & A.J. Jones), pp. 25–37. Soil Science Society of America, Madison, WI.

Guimarães, R.M.L., Ball, B.C., Tormena, C.A., Giarola, N.F.B. & da Silva, Á.P. 2013. Relating visual evaluation of soil structure to other physical properties in soils of contrasting texture and management. *Soil and Tillage Research*, **127**, 92–99.

Hoosbeek, M.R. & Bouma, J. 1998. Obtaining soil and land quality indicators using research chains and geostatistical methods. *Nutrient Cycling in Agroecosystems*, **50**, 35–50.

Karlen, D.L., Mausbach, M.J., Doran, J.W., Cline, R.G., Harris, R.F. & Schuman, G.E. 1997. Soil quality: a concept, definition, and framework for evaluation: (A guest editorial). *Soil Science Society of America Journal*, **61**, 4–10.

McKenzie, D.C. 2012. Visual soil examination techniques as part of a soil appraisal framework for farm evaluation in Australia. *Soil and Tillage Research*, **127**, 26–33.

Mueller, L., Kay, B.D., Hu, C., Li, Y., Schindler, U., Behrendt, A., Shepherd, T.G. & Ball, B.C. 2009. Visual assessment of soil structure: evaluation of methodologies on sites in Canada, China and Germany. Part I: comparing visual methods and linking them with soil physical data and grain yield of cereals. *Soil and Tillage Research*, **103**, 178–187.

Newell-Price, J.P., Whittingham, M.J., Chambers, B.J. & Peel, S. 2013. Visual soil evaluation in relation to measured soil physical properties in a survey of grassland soil compaction in England and Wales. *Soil and Tillage Research*, **127**, 65–73.

Peerlkamp, P.K. 1967. Visual estimation of soil structure (St - number). In: *West-European methods for soil structure determination* (eds M. De Boodt, D.E. Leenheer, H. Frese, J. Low & P.K. Peerlkamp), pp. ???–???. ????????????????, ????????????. 7

Pieri, C., Dumanski, J., Hamblin, A. & Young, A. 1995. Land Quality Indicators. In: *World Bank discussion paper* (ed ?.?.?. ??????????), pp. ???–???. World Bank, Washington DC. 8

Rigby, D., Woodhouse, P., Young, T. & Burton, M. 2001. Constructing a farm level indicator of sustainable agricultural practice. *Ecological Economics*, **39**, 463–478.

Robinson, D.A., Hockley, N., Dominati, E., Lebron, I., Scow, K.M., Reynolds, B., Emmett, B.A., Keith, A.M., de Jonge, L.W., Schjønning, P., Moldrup, P., Jones, S.B. & Tuller, M. 2012. Natural capital, ecosystem services, and soil change: why soil science must embrace an ecosystems approach. *Vadose Zone Journal*, **11**, ???–???. 9

Sarrantonio, M., Doran, J.W., Liebig, M.A. & Halvorson, J.J. 1996. On-farm assessment of soil quality and soil health. In: *Methods for assessing soil quality* (eds J.W. Doran & A. Jones), pp.83–105. Soil Science Society of America, Madison, WI

Schmidt, A.M., Schuiling, C. & Bulens, J.D. 2003. *Verificatie kwaliteit BRP gegevensbank*, p. 80. Alterra, Wageningen.

Schröder, J.J., Aarts, H.F.M., Ten Berge, H., Van Keulen, H. & Neeteson, J.J. 2003. An evaluation of whole-farm nitrogen balances and related indices for efficient nitrogen use. *European Journal of Agronomy*, **20**, 33–44.

Shepherd, T.G. 2000. *Visual soil assessment. Volume 1. Field guide for cropping and pastoral grazing on flat to rolling country*, pp. 84. ??????????????, Palmerston North. 10

Shepherd, T.G. 2010a. *Visual indicators of environmental performance under pastoral grazing: Pastures*. Food and Agriculture Organisation of the United Nations, ??????????? 11

Shepherd, T.G. 2010b. *Visual soil assessment: Part 1 Pastures*. Food and Agriculture Organisation of the United Nations, ??????????? 12

Sonneveld, M.P.W., Bouma, J. & Veldkamp, A. 2002. Refining soil survey information for a Dutch soil series using land use history. *Soil Use and Management*, **18**, 157–163.

Sonneveld, M.P.W., Hack-ten Broeke, M.J.D., Van Diepen, C.A. & Boogaard, H.L. 2010. Thirty years of systematic land evaluation in the Netherlands. *Geoderma*, **156**, 84–92.

Sonneveld, M.P.W., De Vos, J.A., Kros, J., Knotters, M., Bleeker, A., Frumau, A. & De Vries, W. 2012. Assessment of N and P status at the landscape scale using environmental models and measurements. *Environmental Pollution*, **162**, 168–175.

Thomassen, M.A. & De Boer, I.J.M. 2005. Evaluation of indicators to assess the environmental impact of dairy production systems. *Agriculture, Ecosystems & Environment*, **111**, 185–199.

Van Der Werf, H.M.G. & Petit, J. 2002. Evaluation of the environmental impact of agriculture at the farm level: a comparison and analysis of 12 indicator-based methods. *Agriculture, Ecosystems and Environment*, **93**, 131–145.

Wills, S.A., Burras, C.L. & Sandor, J.A. 2007. Prediction of soil organic carbon content using field and laboratory measurements of soil color. *Soil Science Society of America Journal*, **71**, 380–388.

Author Query Form

Journal: SUM
Article: 12117

Dear Author,
During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.
Many thanks for your assistance.

Query reference	Query	Remarks
1	**AUTHOR: Please check that authors and their affiliations are correct.**	
2	**AUTHOR: Please check the all equations.**	
3	**AUTHOR: Please check the clarity of the text "Stable aggregates were again sieved using a dispersion solution (NaPO3)6".**	
4	**AUTHOR: The brace ")" has been deleted from the sentence "VSA scores at different...examine observer influence". Please check.**	
5	**AUTHOR: Please check this website address and confirm that it is correct. (Please note that it is the responsibility of the author(s) to ensure that the URL given in this article are correct and useable).**	
6	**AUTHOR: Please provide the given names/initials for the author Nievergelt, publisher location (city) for reference Boizard et al. (2005).**	
7	**AUTHOR: Please provide the publisher name, publisher location, page range for reference Peerlkamp (1967).**	
8	**AUTHOR: Please provide the editor name(s), page range for reference Pieri et al. (1995).**	
9	**AUTHOR: Please provide the page range for reference Robinson et al. (2012).**	
10	**AUTHOR: Please provide the publisher name for reference Shepherd (2000).**	
11	**AUTHOR: Please provide the publisher location for reference Shepherd (2010a).**	
12	**AUTHOR: Please provide the publisher location for reference Shepherd (2010b).**	
13	**AUTHOR: Please check and confirm if the correct version of figure 2 and 3 has been used.**	