2.1 Introduction

In this work package, we evaluate the properties of those indicators which were identified as candidate physical SQIs through the logical sieve process in WP1. The 18 candidate physical SQIs were filtered further at a project team meeting (Appendix 1) whose aims were to:

- Revisit the outcome of WP1, especially the Logical Sieve exercise to ensure the results were sensible and that no indicators were disqualified unduly.
- Move from the ‘narrative’ of WP1 to a more ‘numerical’ approach in WP2 i.e. given the remaining SQIs, are data available to test the robustness of the candidate SQIs through statistical/modelling analysis?
- Rationalise the remaining SQIs in terms of duplications / surrogacy (e.g. total porosity ≈ 1- BD/particle density); overlaps /double counting (e.g. no. of erosion features and rate of erosion); and linkages (e.g. BD and depth – the Logical Sieve considers all candidate SQIs as independent of each other – are composites useful?)
- Consider scale issues e.g. scaling up / aggregating SQI performance at the field (or equivalent on non-agricultural land) to larger landscape units.

The remaining indicators and justification for their inclusion can be found in Table 2.1. The aim for these SQIs was for WP2 to investigate:

- the uncertainty in their measurement
- the spatial and temporal variability in the indicator as given by observed distributions
- the expected rate of change in the indicator

1 Throughout the report the abbreviation ‘SQI’ refers to soil quality indicators (physical properties), unless shown otherwise.
### Table 2.1. Soil Quality Indicators (physical properties) carried forward from WP1

<table>
<thead>
<tr>
<th>Candidate physical SQIs from WP1</th>
<th>Comments following WP2 project meeting on 18/04/12</th>
<th>Further consideration in WP2?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bulk density</td>
<td>BD alone is an adequate indicator of soil physical quality, but potentially it can be improved by deriving other indicators that are more closely related to soil quality and performance e.g. packing density.</td>
<td>✓ (as input to packing density and depth of soil)</td>
</tr>
<tr>
<td>2. Depth of soil</td>
<td>Effective soil depth defines the volume of soil in which roots can grow. Critical depth is where plant cover achieves values above 40%. Crucial depth is where perennial vegetation can no longer be supported. Affected by rate of soil formation, deposition and degradation processes (erosion, compaction, landslides etc.). Direct impact on the regulation and production function, although this may be only once a critical depth is reached, which in turn depends on site and time specific condition. Some data are available to test this.</td>
<td>✓</td>
</tr>
<tr>
<td>3. Infiltration / drainage capacity</td>
<td>Measurements are highly site specific and very temporally variable. So, the robustness and reliability of measurement using currently available techniques are questionable.</td>
<td>✗</td>
</tr>
<tr>
<td>4. Soil water retention characteristics</td>
<td>Powerful measures of the capacity of the soil to regulate water (hydrology) and produce biomass. Consider pedo-transfer functions as a cheaper proxy for the more expensive lab measurements. This will be explored using a hydrological model. The value of S (Dexter’s S) is indicative of the extent to which the soil porosity is concentrated into a narrow range of pore sizes. In most soils, larger values of S are consistent with the presence of a better-defined microstructure. The S value can be considered as an overall index of physical and structural quality in managed soils. Effectively predicted using pedotransfer functions.</td>
<td>✓</td>
</tr>
<tr>
<td>5. Number of erosion features</td>
<td>Captured by another SQI. Rate of erosion (8) is likely to include a survey / count of erosion features.</td>
<td>✗</td>
</tr>
<tr>
<td>6. Packing density</td>
<td>A measure of dry bulk density (BD), modified by texture. PD = BD + 0.009 C; PD is the packing density (denoted as dimensionless in some texts or as Mg m(^{-3}) or equivalent in others), BD the dry bulk density (Mg m(^{-3}) or equivalent) and C the clay content (wt.%). Direct indicator of soil degradation (soil compaction). Requires data on BD (1) and soil texture / particle size distribution (12). Data holdings were identified as LandIS and some of the ADAS experimental data.</td>
<td>✓</td>
</tr>
<tr>
<td>7. Profile description / visual soil evaluation</td>
<td>Visual soil evaluation (VSE) is based on the qualitative or semi-quantitative evaluation of soil properties Used to assess morphological, physical, biological and chemical soil properties, which are visible or possible to distinguish without laboratory analysis. Insufficient data at the national scale to be able to assess how visual evaluation scores relate to soil function or other indicators of soil physical quality through modelling and data analysis in WP2, but the method should be carried forward as a potential low cost field technique to assess general trends in soil condition at the national scale.</td>
<td>✓</td>
</tr>
<tr>
<td>8. Rate of erosion</td>
<td>Erosion causes detachment and transport of soil particles / aggregates from the in-situ soil mass, leading to a reduction in soil depth / volume (assuming soil bulk density is constant). Mass (tonnes) of soil lost per unit space (hectare) per unit time (year). Direct indicator of soil degradation and loss of function. Monitoring proposed in</td>
<td>✓</td>
</tr>
<tr>
<td>Indicator</td>
<td>Description</td>
<td>Suitability</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>9. Sealing</td>
<td>Degree of impervious cover on the soil surface. Can be assessed through remote sensing (satellite and airborne). A sensitive indicator (a step change in soil functioning is likely following sealing) that can be assessed relatively cost effectively.</td>
<td>✓</td>
</tr>
<tr>
<td>10. Shear strength (Going Stick method)</td>
<td>Limited data available nationally. Only measures the top 10 cm of the soil.</td>
<td>×</td>
</tr>
<tr>
<td>11. Soil structure</td>
<td>Can be captured in 7 above</td>
<td>×</td>
</tr>
<tr>
<td>12. Soil texture</td>
<td>See 6 above</td>
<td>✓ (as input to packing density)</td>
</tr>
<tr>
<td>13. Surface water turbidity</td>
<td>Sediment finger printing has been used in the Demonstration Test Catchment projects to quantify sediment source apportionment. There are questions regarding the sensitivity and indicator responsiveness, as sediment in a river course is spatially removed from the actual source of soil degradation. National coverage of data is limited.</td>
<td>×</td>
</tr>
<tr>
<td>14. Total porosity</td>
<td>Directly related to bulk density (1) and particle density. BD is captured in PD (6) and particle density does not generally change in soils</td>
<td>×</td>
</tr>
<tr>
<td>15. Bulk density (Kopecki ring)</td>
<td>See 1 above</td>
<td>✓ (as input to packing density and depth of soil)</td>
</tr>
<tr>
<td>16. Erodibility/aggregate stability</td>
<td>An assessment of the stability of soil surface structure. The susceptibility of soil to (i) slaking, (ii) differential swelling, (iii) mechanical breakdown by raindrop impact and (iv) physico-chemical dispersion processes can be inferred by measurements of aggregate stability. Data lacking in England &amp; Wales, no standard assessment procedures and highly variable soil property in space and time. However, data may be available from outside E&amp;W – needs further investigation.</td>
<td>✓</td>
</tr>
<tr>
<td>17. Biological status of rivers</td>
<td>Too indirect spatially and temporally from soil sources / degradation processes and difficult to relate to the physical quality of soil.</td>
<td>×</td>
</tr>
<tr>
<td>18. Number of eutrophication incidents</td>
<td>Too indirect spatially and temporally from soil sources / degradation processes and difficult to relate to the physical quality of soil.</td>
<td>×</td>
</tr>
</tbody>
</table>
We will use Power Analysis methods to understand the variability of the indicator as given by the observed distributions. The power of an indicator is its ability to detect a particular change at a particular confidence level given the ‘noise’ or variability in the data (i.e. a particular power to detect a change ‘X’ at a confidence level of ‘Y%’ would require ‘N’ samples). The physical SQIs should ideally be valid across all soil and land management regimes or within a defined sub-set of regimes. The ability of the methods to discriminate change will be tested within such data subsets, since these subsets will reduce variability and hence decrease the size of detectable change for a given sample size. This is further described in the statistical analysis below. The expected rate of change in the indicator will be derived from a combination of the WP1 literature review. The spatial variability associated with the indicator will be assessed at the field and regional scales to determine the intensity of sampling and scale at which this sampling takes place in order to ensure that it is effective, yet remains efficient. This will be informed by assessing different sampling configurations (single, bulked, etc.), as described in the statistics description.

These characteristics determine whether the candidate SQIs will support the implementation of a meaningful soil monitoring programme in England and Wales. We start by considering in more detail the criteria by which to assess indicators and then present a set of ‘fact sheets’ in which we detail the behaviour of each individual candidate indicator. In each case, where appropriate data are available, we do so quantitatively; otherwise this is done (semi)qualitatively or (where no data exist) qualitatively.

Finally, we consider the feasibility of combining a number of indicators and their ‘multivariate’ behaviour, and whether this improves on the behaviour of indicators as single entities.

2.2 Basis for the assessment of SQIs (physical properties)

Physical indicators of soil quality (SQI) need to reflect a change in the soil quality status at a given location. To assess whether a particular SQI is effective, this needs to be contextualized in terms of the expected changes in soil quality, i.e. in the capacity for the soil to function. A functional indicator detects a meaningful change in a given soil function. In order to evaluate the effectiveness of the indicator, we need to establish what this meaningful change is, particularly in the relationships between the SQI and the soil function. Physical SQIs need to reflect meaningful changes in soil quality, but also be meaningful with regard to the soil processes they represent (Figure 2.1). It is important to note here that what we are evaluating is change; change in the SQI itself and relating that change to a change in the processes going on in the soil and in turn relating that change to how the soil functions.

Figure 2.1. Relationship of soil physical property, process and function

There are cases where the indicator itself drives changes in soil function. In the case of compaction, where values of BD exceed a threshold value, some studies suggest this has immediate impact on crop yields (production function) as plant roots cannot effectively penetrate the soil (Table 2.2). In this case, we could consider evaluating an SQI around a critical or target value. This is the approach taken by Merrington et al. (2006), where critical values or target ranges were identified for candidate SQIs and their usefulness evaluated around these ranges. The advantage to this approach is in its simplicity: a value of the SQI needs to be ascertained at a given location and the obtained value compared with the range or critical value. The disadvantage to this approach is that it does not capture the dynamic relationships between SQI and soil functions, and that these may be different for different
soil functions, land uses and soil types (Jones, 1983). As Merrington et al. (2006) already considered this approach to physical SQIs and the general approach has been further developed in WP1, WP2 will focus on the dynamic relationships between soil functions and SQIs, and only revisit the Merrington et al (2006) findings where this is appropriate.

Table 2.2. Critical bulk density considering restriction to root elongation or yield decrease (BDc Rest) for different crops and soil texture for soils from Brazil (Reichert et al., 2009)

<table>
<thead>
<tr>
<th>Source</th>
<th>Texture</th>
<th>Soil type</th>
<th>Crop</th>
<th>BDc Rest (Mg m⁻²)</th>
<th>Macrosc (m³ m⁻³)</th>
<th>Restriction</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Maria et al. (1999)</td>
<td>50</td>
<td>Oxisol</td>
<td>No-tillage</td>
<td>Soybean</td>
<td>1.21</td>
<td>Root elongation</td>
</tr>
<tr>
<td>Sterck (2003)</td>
<td>209</td>
<td>Alfisol</td>
<td>No-tillage</td>
<td>Black bean</td>
<td>1.79</td>
<td>Root elongation/ yield</td>
</tr>
<tr>
<td></td>
<td>555</td>
<td>Oxisol</td>
<td>No-tillage</td>
<td>Corn and wheat yield</td>
<td>1.54</td>
<td>Root elongation</td>
</tr>
<tr>
<td>Bontel et al. (2004)</td>
<td>687</td>
<td>Oxisol</td>
<td>Chisel plow</td>
<td>Rice</td>
<td>1.63</td>
<td>Yield</td>
</tr>
<tr>
<td>Boettler and Centurion (2004)</td>
<td>687</td>
<td>Oxisol</td>
<td>Chisel plow</td>
<td>Soybean</td>
<td>1.08</td>
<td>Root elongation</td>
</tr>
<tr>
<td>Collares (2005)</td>
<td>217</td>
<td>Oxisol</td>
<td>No-tillage</td>
<td>Black bean</td>
<td>1.53</td>
<td>Root elongation/ yield</td>
</tr>
<tr>
<td></td>
<td>607</td>
<td>Oxisol</td>
<td>No-tillage</td>
<td>Soybean</td>
<td>1.49</td>
<td>Root elongation/ yield</td>
</tr>
<tr>
<td>Suzuki (2005)</td>
<td>391</td>
<td>Alfisol</td>
<td>No-tillage</td>
<td>Soybean</td>
<td>1.56</td>
<td>Root elongation</td>
</tr>
<tr>
<td></td>
<td>331</td>
<td>Alfisol</td>
<td>No-tillage</td>
<td>Soybean</td>
<td>1.50</td>
<td>Root elongation</td>
</tr>
<tr>
<td></td>
<td>434</td>
<td>Oxisol</td>
<td>No-tillage</td>
<td>Soybean</td>
<td>1.39</td>
<td>Root elongation</td>
</tr>
<tr>
<td></td>
<td>265</td>
<td>Oxisol</td>
<td>No-tillage</td>
<td>Soybean/corn</td>
<td>1.36</td>
<td>Root elongation</td>
</tr>
</tbody>
</table>

* Soybean (Glycine max); black beans (Phaseolus vulgaris); wheat (Triticum aestivum); corn (Zea mays); rice (Oryza sativa).

There is, however, a limited amount of information in the literature on the dynamic relationships between soil functions and particular SQIs. In some cases, such as soil water retention characteristics and packing density, there is a clear picture of the relationships between these SQIs and particular soil functions, such as the production function (e.g. crop yield). However, there is scarce to no evidence relating the aesthetic or cultural function to any SQI. As a result, we have based our assessment of ‘a meaningful change’ on those functions for which we can reasonably find evidence or make inference; the production and regulation function. These functions are also highly relevant in terms of soil policy (see Work Package 3), given current drivers relating to food security; flood alleviation and carbon losses.

Soil properties are spatially and temporally variable, and that spatial variability is driven by a number of factors, amongst which parent material, climate, land use and management predominate. For example, Table 2.3 shows the variability of bulk density with depth, time and soil water content (Logsdon and Karlen, 2004). This variability introduces ‘noise’ into the signal response (i.e. meaningful change) and does so in two ways:

i) There is a spatial unit over which soil quality status is assessed (plot, field, farm, catchment, national scale). The spatial variability within this unit will introduce variability to the SQI irrespective of whether there are any changes in soil function(s). There are numerous sampling methods and sampling designs possible which minimize the effect of this type of variability. A recent report by Lark et al (2012) reviews the impact of these different designs on one candidate physical SQI (bulk density). In WP2, we consider land units of increasing spatial size, with the assumption that SQI will be used to assess changes in soil quality at the different spatial scales.
Table 2.3. Variability of bulk density with depth, time and soil water content (Logsdon and Karlen, 2004)

<table>
<thead>
<tr>
<th>Depth (mm)</th>
<th>Bulk density (Mg m⁻³)</th>
<th>Water content (m³ m⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.18⁺</td>
<td>0.84⁺</td>
</tr>
<tr>
<td>30</td>
<td>1.22</td>
<td>1.15⁺</td>
</tr>
<tr>
<td>50</td>
<td>1.25</td>
<td>1.31</td>
</tr>
<tr>
<td>70</td>
<td>1.36</td>
<td>1.32⁺</td>
</tr>
<tr>
<td>90</td>
<td>1.36</td>
<td>1.34⁺</td>
</tr>
<tr>
<td>110</td>
<td>1.43</td>
<td>1.43</td>
</tr>
<tr>
<td>130</td>
<td>1.37</td>
<td>1.41</td>
</tr>
<tr>
<td>150</td>
<td>1.36⁺</td>
<td>1.44</td>
</tr>
<tr>
<td>170</td>
<td>1.33⁺</td>
<td>1.41⁺</td>
</tr>
<tr>
<td>190</td>
<td>1.25⁺</td>
<td>1.45⁺</td>
</tr>
<tr>
<td>210</td>
<td>1.34⁺</td>
<td>1.44⁺</td>
</tr>
<tr>
<td>230</td>
<td>1.26</td>
<td>1.25</td>
</tr>
<tr>
<td>250</td>
<td>1.27⁺</td>
<td>1.36</td>
</tr>
<tr>
<td>270</td>
<td>1.26⁺</td>
<td>1.36</td>
</tr>
<tr>
<td>290</td>
<td>1.27</td>
<td>1.33⁺</td>
</tr>
</tbody>
</table>

* Significant differences between adjacent dates at *p* = 0.05.

ii) A second consideration is the impact of a particular land management practice on the effectiveness of an SQI to indicate change in soil quality. In this case, where data are available, we will use a combination of land use, soil type and climate classes to determine the impact these factors may have on the background variability of the SQI we are considering.

2.3 SQI Assessment Structure

In the following fact sheets, we present a number of key items of information for each SQI. This includes a brief summary of salient findings from WP1. For all SQIs investigated, the level of data analysis and modelling is limited by the evidence base available to us. This varies considerably amongst the 7 SQIs considered. However, within each fact sheet:

1. We describe whether the SQI is meaningful with regard to soil functions.
2. We review the evidence from both the literature and the project team’s expert knowledge to determine what relationships exist between the SQI and soil processes and from this review what might constitute meaningful change.
3. We assess the spatial variability of the SQI and the implications for sampling. We do so using a combination of spatial statistics and power analyses, where this is possible (more details below). This will give an indication of the number of physical samples needed (sample size) to determine whether the SQI has changed to an extent where there are corresponding changes in soil quality. As there are different circumstances under which a SQI may be used, we evaluate sample sizes required for different spatial levels (e.g. plot, field, farm) and for a national assessment considering land use and climate. We also describe a general labour effort related to obtaining physical samples and their respective analytical costs.
4. We evaluate alternative (proxy) measurements to the 7 candidate SQIs, either as some linear combination of existing soil measurements (denominated pedotransfer functions) or as readings from sensing systems which correlate with the SQI. This WP reviews a number of measurements that can be made on soil, which are indirect measures of soil properties that are otherwise hard to measure. Examples of this are electro-magnetic induction (Milton and Webb, 1987; Milton et al., 1995), and visual and near infrared (VNIR; Viscarra-Rossel et al., 2008). Using chemometric techniques, these measurements on soil are used to infer values of the soil property of interest. These methods are being used more widely and, coupled to remote sensing, form a powerful
and effective approach to monitoring soil conditions (Kuang et al., 2012). However, where we review their appropriateness as proxy or even complementary measurements to the candidate SQIs, we do not consider these measurements as soil quality indicators per se as they themselves cannot be interpreted in a meaningful way. Their value lies in their correlation to soil measurements which are meaningful with regard to the soil processes they represent, which in turn affect soil quality (Figure 2.1).

2.4 Statistical Background and Sample Size Calculations

The monitoring of soil quality indicators requires sampling strategies which allow assessment of changes in soil quality, taking into account soil heterogeneity, seasonal fluctuations or analytical uncertainties (Arshad and Martin, 2002; Leprêtre and Martin, 1994; Arshad et al., 1996) to be set up.

The objective in this study is to detect a change in an SQI and to do so we implemented a set of power analyses. Statistical power is the probability that a specific difference will be detected at a specified level of confidence. In this case, we computed the sample size required to determine a range of difference at the 95 % confidence level for a two-tailed t-test and so generate a set of figures plotting the sample size needed to detect that difference at 95 % confidence. We compute the sample size for a two tailed t-test, because the directionality of the change is not specified (i.e. we are interested in both increases and decreases in the SQI). We need to specify the random variation of the data $s^2$. The key here is to obtain a measure of $s^2$, which is associated with the change we wish to detect and in principle this is different from the baseline SQI variability (Lark, 2009). However, there is little information in the literature on the relationship between SQIs and soil functions, and even less information around the variability of this change. As such, in this study we will assume that the variance of the baseline data will be the same as the variance of the resampled data. In other words, the baseline variation in soil properties does not change over the monitoring period. Given the available data, this assumption is reasonable and further developed in Lark (2009). This approach has been used elsewhere (e.g. Brazier et al., 2012; Holman et al., 2003).

2.4.1. Obtaining an estimate of the SQI variability.

Through power analysis, Lark (2009) demonstrates that the most effective sampling approach to detect change in a soil property (in that case soil organic C) is a paired sampling approach, where the same location is repeatedly sampled in time. This re-sampling practice reduces the number of samples needed to detect any change, although the accuracy of revisiting the same site can be challenging. Even so, this approach is plausible with monitoring soil organic C content, as a soil property in undisturbed sites. However, most of the physical soil quality indicators considered here are highly dynamic in space and time, not least because they are sensitive to management inputs, as are the soil functions they are indicating. Given this variability and sensitivity, these changes in soil properties, processes and associated functions are better assessed using a randomized sampling design. It is this design that we will consider here, whilst noting that the sample sizes presented in the statistical power graphs would be smaller if a paired sampling approach is considered. A further efficiency is introduced if we consider stratification by land use and soil type, because we have a priori grounds to expect that changes in SQIs may vary for different land use (e.g. arable v. pasture v. woodland) and soil types (e.g. shallow v. deep soils). For example, a change in packing density will have different impacts on different land uses (and even land covers and plant species) and on different soils (e.g. sands v clays). We will investigate these relationships for those SQIs where we have the available data (i.e. packing density and soil depth). Variances in this stratified method were obtained using a linear mixed model in which the fixed effect is the overall mean of the SQI. The random effects in this case are a correlated random variable and the uncorrelated error term. Power Analysis was executed in
JMP software (SAS Institute Inc.) under their Experimental Design procedures. Linear mixed modelling was executed using the nlme library in R.

An alternative approach to obtaining an estimate of the SQI variability is a model-based approach. Here, we consider the spatial correlation of the SQI only, and can obtain from this a model of the spatial correlation, or variation, denoted by the variogram (Figure 2.2). We obtain the variance of a region as the double integral of the variogram over the region of interest. The geostatistical analysis was executed in R software. The variogram procedure was executed in the library Gstat. This approach allows us to consider the sampling requirements for estimating the change in an SQI over increasing areas, from the field scale upwards.

Figure 2.2. Variogram of Packing Density obtained from data from LandIS (NSRI, Cranfield University) and ADAS (Newell Price et al., 2012a).

2.4.2. Applicable analysis to SQIs.

The above discussion is predicated on the availability of appropriate data. This is different for the different SQIs we considered, so the level of analysis varies for each SQI.

**Packing density.** We had access to two large datasets, LandIS (containing data produced by the Soil Survey of England and Wales), which contains the representative profiles for England and Wales and contains approximately 1,250 measurements of bulk density and clay content; these were averaged over soil profiles. We supplemented this dataset with another dataset from ADAS (Newell Price et al., 2012a) which added a further 300 short range measurements of bulk density. From this we obtained a comprehensive data set for packing density and the variogram in Figure 2.2 is based on this dataset. We then obtained estimates of the dispersion variances for areas of increasing size of 5, 10, 25 and 50 km². We also consider here a random sampled stratified approach based on a set of stratum developed for Defra project SP1606 (Graves et al., 2011), in which soil and land use are combined to generate 'supra'-classifications of the soil / land use combinations in England and Wales. This was based on the work by from Hill et al. (2003) who found:
The data grouped by land use and soil order explained about two thirds of the total variability in soil quality data. Similar trends in soil quality were obtained across the 10 regions indicating land use as the major driver of soil quality.

**Soil Water Retention Characteristics.** A total of 2,480 soil profiles for which soil water retention data were available were extracted from the LandIS database. From these, and using hydrological models, the SQIs associated with Soil Water Retention Characteristics were derived. Pedo-transfer functions were calculated based on this dataset.

**Soil Depth.** The only adequate dataset available to us in this case is the LandIS dataset (Soil Survey of England and Wales). These data comprise profiles across England and Wales and depth measurements were restricted to the top 150 cm in most sampled profiles. This means many data points don’t represent accurately the full depth of soil. We therefore only considered those soils identified as ‘shallow soils’ in the soil survey, to avoid the unrealistic (truncated) soil depth data and these are likely to be soils where functions may be compromised if depth changes. This gave 177 soil profiles. Power analysis was based on a variance estimate from the stratified mixed model described above.

**Aggregate Stability.** There is no national dataset on aggregate stability. The most geographically extensive dataset we sourced was generated as part of Defra SP0519: Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration, led by Professor Whitmore and Dr Chris Watts of Rothamsted Research. Data on aggregate stability on a range of soil types and land uses was obtained for a set of test sites across England. The power analysis was based on a variance estimate from the stratified mixed model described above.

**Rate of Erosion.** The statistical analysis and sampling design for monitoring the rate of erosion is comprehensively covered in Brazier et al. (2012) as part of Defra SP1303 Developing a cost-effective framework for monitoring soil erosion in England and Wales. This includes power analyses based on variability (and lack) of data on erosion rates at the national scale. We report a summary of this here.

**Soil Sealing.** This is considered within the concept of remote sensing, so the sample obtained is ‘exhaustive’, i.e. a classified satellite image can cover the entire area of interest, depending on the scale of survey and resolution of image required. As the area of interest is not sampled, power analysis is not appropriate here. We discuss here pixel size (sample support) and different satellite images that are available to determine sealing of soil and degree of imperviousness.

**Visual Soil Evaluation.** This is (semi) qualitative assessment of soil quality and we discuss its suitability to monitor change in soil quality (not just the baseline soil quality), especially as many schemes are based on categorical (i.e. scores) rather than continuous data (e.g. packing density; depth of soil). Comparisons are made with the more quantitative measures discussed earlier.
2.5 Fact Sheet 1. Packing Density.

Background

Packing density is a measure of total soil porosity, rather than pore sizes and their distribution or pore continuity. It is therefore a fairly crude measure of soil porosity with respect to soil functions such as water regulation, biomass production and habitat support (e.g. certain soil fauna rely on soil pores of a certain size). Nevertheless, it provides a good estimate of ‘compaction’, in terms of reduced total porosity within a soil and a degree of degradation, which can be taken as the inverse of soil quality (Huber et al., 2008).

Packing density is, in essence, a measure of dry bulk density (BD) modified by clay content and thus is a better indicator of soil compaction than dry bulk density alone, as it provides an indirect estimate of soil porosity (Hall et al., 1977) and has better correlation with air capacity (Huber et al., 2008). It is calculated as:

$$\text{PD} = \text{BD} + 0.009 \text{C}$$

Where PD is the packing density (Mg m$^{-3}$), BD the dry bulk density (Mg m$^{-3}$) and C the clay content (weight %). PD provides a means of comparing BD between soils of different clay content. Although identified as the only suitable physical SQI by Merrington et al’s (2006), BD values alone are not a ‘good’ indicator of soil physical quality in isolation, but can be used to derive other indicators/variables that are more closely related to soil quality and function e.g. PD.

Detecting change and relationship to soil functions

Measurements of packing density (PD) can detect relatively large changes in soil physical properties, but are less sensitive than other indicators such as macroporosity, due to the fact that packing density provides an estimate of total porosity, rather than any particular pore size range. For example, Kurz et al. (2006) reported that summer grazing by cattle (compared with ungrazed and untrafficked control plots) resulted in a 8-17% increase in BD, but that corresponded to much larger (60-80%) reduction in macroporosity. Also, in a review of soil quality indicators in the Waikato region of New Zealand, Taylor et al. (2010) found that macroporosity was a more sensitive indicator of surface compaction (assumed to be the inverse of soil quality) than dry bulk density. Lower biomass production was observed at < 10% macroporosity (~10 kPa) for grassland and arable soils (Mackay et al., 2006), but no threshold was provided for dry bulk density.

Nevertheless, PD or BD has been used to detect differences in ‘compaction’ between different management practices, such as contrasting tillage systems (Dam et al., 2005; Da Silva et al., 2001). Dam et al. (2005) found that BD was 10% higher in no-till compared with conventional tillage systems, particularly in the 0-10 cm layer.

For a soil of a given clay and organic matter content, PD can be related to biomass production. Crop yield tends to increase with increasing PD up to a certain threshold value (depending on clay and organic matter content and crop type) and will then normally decrease beyond this threshold. For soils with higher organic matter content (e.g. >c.5% for spring crops on arable soils), the PD threshold at which yields decline tends to be higher than for lower organic matter soils (Arvidsson, 1998). This is illustrated in Figure 2.3. Higher organic matter soils are generally more resilient to compaction (Gregory et al., 2009).
Figure 2.3. Relative yield (A=100) in treatment A-D for groups of soils with different a) clay and b) organic matter content. A = no traffic, B = one pass by a tractor wheel, low tyre inflation pressure, C = one pass by a tractor wheel, high tyre inflation pressure, D = three passes by a tractor wheel, high tyre inflation pressure.

In Figure 2.3, traffic treatments (A-D) are used as a surrogate for ‘degree of compactness’ (Hakansson, 1990). The BD in the field is expressed as a percentage of the reference BD determined in the laboratory using a 200 kPa uniaxial compression test.

Sampling effort needed

This WP considers four aspects of each SQI: the uncertainty in measurement; the variability in the indicator; the expected rate of change in the indicator; and the spatial variability associated with the indicator. These factors will determine the sampling effort needed to detect a change in the SQI.

Bulk density measurement is usually carried out using a kopecki ring (or cylinder), which has a diameter and depth of approximately 5 cm. The spatial area is therefore limited and multiple replication (3-5 samples) is usually required at each sampling depth to reduce standard error and detect differences over time or between treatments/management practices for a given soil type (e.g. texture and carbon content).

The Kopecki ring method is reproducible using standard operating procedures. However, the method is also sensitive to variations in implementation and in soil moisture content such that sufficient replication (4-5 samples per experimental plot or soil unit) is normally required to detect a 10-15% change in value.

Packing density values are thus sensitive to soil moisture content, so it is important that samples are taken as close to field capacity as possible. Soil moisture content is calculated by laboratory and on-line methods so can be taken into account to some degree. The latter methods include NIR sensors that can be used to detect changes in moisture content across a field in real time.

Power analysis and sampling size.

In this case we consider two approaches; one based on a set of stratum developed for Defra project (SP1606; Total Costs of Soil Degradation) in which soil, land use and climate are combined to generate ‘supra’-classifications of the soils in England and Wales. We investigated if we could find differences in the uncertainty in measurement; variability in
observed distributions; expected rate of change; and spatial variability stratified by land use and soil type to determine from this the sample size needed to detect a meaningful change in PD.

In other words, if spatial and temporal variability in packing density was dependent on land use, soil type or climate, this implies a different sampling regime is required for different geographical areas if the monitoring is to be statistically robust. This exercise also explored whether meaningful changes in packing density (i.e. ones that reflected a change in soil processes and functioning) might be different for different land use / soil / climate combinations. In other words, would the functioning of arable soils with high clay content be more affected by a change in PD than woodland soils on sand? The results from the power analysis can be found in Figure 2.4 (national data) and 2.5 (data at different spatial scales).

![Figure 2.4. Results from the power analysis based on land use by soil strata. The insert shows the spatial distribution of the data points superimposed on the land use /soil classification (see Defra SP1606 for details regarding the stratification).](image)

We also determined the sample size needed if a change is to be determined over areas of increasing size using a model based approach described earlier (geostatistics; Figure 2.5). In these figures, we can see that if we wish to determine a given change in PD for a given spatial area (field; 5 and 10 km$^2$), farm (25 km$^2$) and at the landscape level (50 km$^2$) we increasingly need a larger sample size to determine this change within this spatial area. If we consider factors which affect the spatial variability of PD such as landuse, then this information will reduce the number of samples needed at the landscape scale. This figure
(Figure 2.5), compared to the 50 km\(^2\) figure above, indicates the smaller number of samples needed to determine any given difference in PD.

**Figure 2.5. Power Analysis using a model based approach in which the variation of different regions (size of spatial unit) is obtained from the spatial correlative structure (variogram in Figure 2.2)**

**Proxy techniques**

Higher resolution (1500 – 2500 samples per hectare) with measurements taken over larger areas can be achieved through proxy techniques such as on-line (mobile) and non-mobile systems. There are essentially three proxy methods:

i. A combination of Visual and Near Infrared (vis-NIR) measurements, combined with Theta probe determinations for soil moisture. This is currently an experimental method and has not been tested on a wide range of soil type/land use combinations.

ii. A combination of soil resistance (penetrometer measurements) and vis-NIR measurements to determine BD (and thus PD when combined with clay content). This is a destructive method that has been successfully trialled on 17 fields representing 10 different soil textures.

iii. An ‘on the go’ system, where sensors are placed on tines behind a tractor and BD is measured as a function of vis-NIR and soil resistance. This method provides a good estimate of BD at the time of cultivation (i.e. post-harvest or prior to seedbed preparation in the spring), but cannot provide within crop measurements without
causing some crop damage. The system has been trialled in most soil texture types in the UK.

All three methods involve multiple sensors and advanced data fusion techniques (Mouazen and Ramon, 2006).
2.6 Fact Sheet 2. Soil Water Retention Characteristics.

Background

A series of important physical SQIs can be derived directly from soil water retention characteristics. These SQIs, sometimes referred to as capacity-based indicators, include plant available water capacity (PAWC), air capacity (AC; or drainable porosity), relative field capacity (RFC), macroporosity (M), porosity of the soil matrix (e.g. Reynolds et al., 2002; 2009) and the soil physical quality index $S$, developed by Dexter (2004a,b,c).

An important feature of these physical SQIs is that they are related to both pore volume and pore size distribution (Reynolds et al., 2009). As such, and as with hydraulic parameters in general, these indicators, including $S$, are likely to be more sensitive to temporal and spatial changes in soil condition and quality than other indicators which are governed by pore volume alone (e.g. bulk density) (Dexter, 2004a; Merrington et al 2006).

Soil physical quality index $S$

$S$ is defined as the modulus of the slope of the soil water release function (plotted against the natural logarithm of the matric potential or pressure head) at its inflection point. The soil water release function used is the Van Genuchten equation (1980) which is fitted to water release curve data:

$$\theta = (\theta_{sat} - \theta_{res})\left[1 + \left(\frac{1}{\alpha h}\right)^n\right]^{-m} + \theta_{res}$$

(1)

Where $\theta$ is the moisture content, $h$ is the pressure head, $\theta_{sat}$ is the saturated moisture content, $\theta_{res}$ (residual moisture content) $\alpha$ and $n$ are fitting parameters and $m = 1 - 1/n$.

Following Dexter (2004a), the water retention function (Eq. 1) is subsequently differentiated with respect to the natural logarithm of the pressure head, $h$, to give an indication of the pore size distribution, maximum pore size and frequency of pores in this maximum pore class:

$$\frac{\partial \theta}{\partial \ln(h)} = -mn(\theta_{sat} - \theta_{res})\alpha\theta^n h^n[1 + (\alpha h^n)^{n-m-1}$$

(2)

Plotting the change in moisture content against the natural logarithm of the pressure head rather than pressure, $h$, is suggested to be a more appropriate measure of air entry into granular material such as soil with a broad pore size distribution (Dexter and Bird, 2001). The inflection point of the water release curve, which is located at the peak of the differentiated function, is the point at which drainage is maximum and has two features, its location and its slope. The location is given by the pressure $h_i$ and according to Dexter (2004a) this occurs at:

$$h_i = \frac{1}{\alpha} \left[\frac{1}{m}\right]^{\frac{1}{n}}$$

(3)

with a corresponding water content, $\theta_i$:

$$\theta_i = (\theta_{sat} - \theta_{res})\left[1 + \frac{1}{m}\right]^{-m} + \theta_{res}$$

(4)

According to Dexter (2004a), the slope of the curve at this point, $S$, is the modulus of:
\[ S = \left| n(\theta_{\text{sat}} - \theta_{\text{res}}) \left[ 1 + \frac{1}{m}\right]^{-1}\right| \]  

Note that if the gravimetric water content is used to plot the curve (as originally done by Dexter (2004a)) instead of the volumetric water content, two different indices (\(S_g\) and \(S_v\), respectively) are obtained. They are however related through the soil dry bulk density, \(\rho_b\), so that:

\[ S_v = \rho_b S_g \]  

The inflection point of the water release curve occurs at the matric potential, \(h_i\), representative of the dominant pore size of the soil where the specific water capacity is highest. Hence, the \(S\) value can be considered as an overall index of physical and structural quality in managed soils (Dexter, 2004a; Reynolds et al., 2009). The \(S\) index derived from soil hydraulic behaviour has been related to other soil physical processes by Dexter and co-workers (Dexter, 2004a, b, c; Gate et al., 2006; Dexter and Czyz, 2007; Dexter and Richard, 2009) and correlated to other soil quality indicators such as bulk density and organic matter content. Dexter (2004a) also showed that the \(S\) index was related to the root growth of soil.

**Relationship to soil functions**

The higher the value of \(S\), the higher the soil physical quality. Dexter (2004a) provided categories of physical quality which classed soils with an \(S_g\) value of \(\geq 0.035\) as having 'good' physical quality and those with \(S_g \geq 0.050\) as having 'very good' physical quality. The maximum value of \(S_g\) found by Dexter and Czyz (2007) was 0.140 and the majority of agricultural soils studied had values of \(S_g\) which fell between 0.015 and 0.060. However, whilst Dexter (2004a) illustrates the widespread applicability of the physical quality index \(S\) for a range of agricultural soils, optimum ranges for other land uses have yet to be developed.

Sands with unimodal and narrow pore size distributions will be characterised by a large \(S\) index, despite having poor structure and poor water or air capacity, for example. As such, Reynolds et al. (2009) suggest that the quality index \(S\) should be used in combination with other physical quality parameters, especially the other capacity-based indicators: plant available water capacity (PAWC); relative field capacity (RFC); and macroporosity (M). In turn, these can be derived using key features of the water release curve:

- the volumetric moisture content (\(\theta_{\text{fc}}\)) at field capacity, occurring at 0.5 or 1m (5 or 10 kPa) pressure head;
- the saturated moisture content, \(\theta_{\text{sat}}\), at 0m pressure head;
- the moisture content at permanent wilting point, \(\theta_{\text{pwp}}\), occurring at 150m pressure head; and
- the porosity of the soil matrix, \(\theta_{\text{m}}\), occurring at 0.1m pressure head. The ranges of optimal values for these indicators are reviewed below.

**Plant Available Water Capacity (PAWC).** Plant Available Water Capacity (vol / vol; \(\text{cm}^3\cdot\text{cm}^{-3}\)) is an indicator of the soil's capacity to store and provide water that is available to plant roots. It is defined as \(\text{PAWC} = \theta_{\text{fc}} - \theta_{\text{pwp}}\). This simply assumes that, above field capacity, the soil drains readily and that water is lost through percolation and that, below permanent wilting point, plant roots are not able to extract water. The review by Reynolds et al. (2009) suggests that \(\text{PAWC} \geq 0.20\) is required for maximal root growth and function (Cockroft and Olsson, 1997), while \(0.15 \leq \text{PAWC} \leq 0.20\) is 'good', \(0.10 \leq \text{PAWC} \leq 0.15\) is 'limited', and \(\text{PAWC} \leq 0.10\) is 'poor' for root development. These values could be assumed to represent a
meaningful change in this physical SQI as changes of this magnitude are expected to affect root (and therefore crop) growth.

*Macroporosity M.* Macroporosity \( (\text{cm}^3 . \text{cm}^{-3}) \) can be defined as \( M = \theta_{\text{sat}} - \theta_{\text{m}} \) which represents the volume of macropores with an equivalent pore diameter \( \geq 300 \mu\text{m} \). It is an indicator of the capacity of the soil to quickly drain excess water and facilitate root growth (Reynolds et al., 2009). It can also potentially indicate good structure. According to Reynolds et al. (2009), \( M \geq 0.05–0.10 \) is often considered optimal and \( M \leq 0.04 \) has been found in soils degraded by compaction. Reynolds et al. (2009) propose \( M \geq 0.07 \) as an optimal figure and \( M = 0.04 \) as a lower critical limit.

*Relative Field Capacity RFC.* Relative Field Capacity is defined as \( RFC = \frac{\theta_{\text{FC}}}{\theta_{\text{sat}}} \) and is an indicator of the capacity of the soil to store both water and air, relative to the total pore volume. RFC represents the proportion of pores filled with water at field capacity and therefore informs directly on whether the soil tends to be too wet or too dry. Moisture content at field capacity, \( \theta_{\text{fc}} \), on its own does not provide this information, unless it is associated with \( \theta_{\text{sat}} \) or \( \theta_{\text{pwp}} \) as with the indicators AC and PAWC. According to Reynolds et al (2009), for rain-fed agriculture and mineral soils, the optimal balance between available water and air capacity occurs between \( 0.6 \leq RFC \leq 0.7 \). Values outside this range indicate insufficient water (\( \leq 0.6 \)) or air (\( \geq 0.7 \)) and a potential reduction in microbial activity, notably microbial production of nitrate, which can impact crop growth and yield. Again, these values could be assumed to represent a meaningful change in this physical SQI as changes of this magnitude are expected to affect root (and therefore crop) growth.

**Larger spatial scale effects**

To be meaningful SQIs, the soil water retention characteristics listed above should be indicative of soil functions that operate at a large spatial scale (e.g. the water regulation function), beyond the laboratory scale where these relationships are often founded. Degradation of physical soil quality associated with agricultural land management affects hydrological response at field and small catchment scales (e.g. O’Connell et al. 2004a, O’Connell et al., 2007; McIntyre and Marshall 2010). However, after reviewing the scientific literature, O’Connell et al. (2004a) concluded that, although there was substantial evidence of changes in land use and management practices affecting runoff generation at the local and small catchment scale, there was very limited evidence that these local changes were propagated downstream at the larger catchment scale. Clearly there is a significant evidence gap in connecting soil hydrological processes (and how these are reflecting in soil physical properties) at different spatial scales (i.e. laboratory to field to catchment).

O’Connell et al. (2007) and Beven et al. (2008) state that this is because of a combination of uncertainty in estimates of precipitation inputs to a catchment; the nonlinear impacts of changing catchment inputs over time on stream discharges; the uncertainty in measurements of stream discharges (particularly during flood events); the uncertainty in characterising land use / management patterns in space and time; and that significant impacts at the small scale may not necessarily have significant impact at catchment scales, due to landscape connectivity. However, Beven et al. (2008) concluded that the difficulty in identifying consistent change in soil hydrological processes given the limitations of the available data did not mean that change is not happening and should not be taken to imply a policy of doing nothing.

In summary, whilst additional field to small catchment scale experimental research and monitoring and modelling studies have been conducted since the previous reviews of FD2114 (O’Connell et al. 2004) and FD2120 (Beven et al., 2008)), no studies have contradicted the conclusion of FD2120 that the variability between years and inconsistencies in rainfall and flow data appear to dominate any impacts of land use and management.
change (and any effect these might have on soil physical properties such as soil water retention characteristics) on flow characteristics at the catchment scale over time.

Sampling effort needed

This WP considers four aspects of each SQI: the uncertainty in measurement; the variability in the indicator; the expected rate of change in the indicator; and the spatial variability associated with the indicator. These factors will determine the sampling effort needed to detect a change in the SQI.

The standard Soil Survey of England and Wales sampling method for determining soil water retention characteristics is to collect three undisturbed soil samples per soil horizon (Avery and Bascomb, 1982). The samples are collected using a coring device that reduces sample compaction. Samples are ideally collected in winter or spring when the soil is near field capacity so that swelling clay soils are at their maximum expansion. The laboratory measurement for soil water retention can take several months depending on soil texture and the number of observation points on the soil moisture release curve that are to be determined. For example, the initial saturation of the soil samples can take between 1 to 2 days for sandy soils and up to 2 weeks for clayey soils. Samples are then placed on tension tables and allowed to equilibrate to the tension (or suction) applied through the tension table.

Equilibration is determined by weighing the sample every 2 days until their weight change is <100 mg between weighing. For suctions >0.5m head of water (5kPa), the samples are transferred to a pressure membrane apparatus where air pressure is used to force water out of the sample. Water is collected in a reservoir and weighed. When the weight of water <3 mg per day over several days the samples are assumed to be equilibrated. The sample can then be removed and weighted. Finally the sample is oven dried at 105°C to determine bulk density of the sample. The volumetric water content at each suction/pressure point can then be determined.

The soil water release characteristics are influenced by land use and soil management, and these changes can occur over short time scales. It is generally accepted that soil hydrological properties to which soil water release characteristics are related, exhibit both short and long range variability (Nielsen et al., 1973). Samples that are collected close together (within a few meters) are more similar than those collected at greater distances.

Proxy techniques

Obtaining soil water retention curves to derive indicators such as the Dexter’s S is time consuming and requires considerable effort. Alternatively, pedotransfer functions (PTFs) can be used to derive these properties from simple to measure soil characteristics. PTFs are (linear) relationships between the desired, but expensive to obtain soil properties (i.e. S index, AC, PAWC and RFC) and easily measured soil properties of which BD and soil C are the most commonly used (Mayr and Jarvis 1999; Matula et al., 2007). In this case, given the effort needed to obtain Soil Water Retention Characteristics, these SQIs can only be considered if it is feasible to develop a set of PTF’s from which these particular SQIs can be derived. We explore the feasibility of developing such PTF’s in the following section.

Determination of Soil Water Retention Characteristics (index S, AC, PAWC and RFC) from the LandIS data base

A total of 2480 soil profiles for which soil water retention data were available were extracted from the LandIS database. The relevant data available consisted of volumetric moisture content measured at pressure heads (expressed in water height equivalents) of 0.5, 1, 4, 20 and 150 m. In addition, total porosity (derived from measured bulk density and particle
density) was used as an approximation of the moisture content at saturation, \( \theta_{sat} \) (= 0m pressure head). For each soil, the water retention function represented by the Van Genuchten equation (Eq 1) was fitted to the data using \( \theta_{res} \), \( \alpha \) and \( n \) as fitting parameters. The Dexter’s S index was calculated with Eq 5, using the fitted parameters and converted into \( S_g \) using Eq 6. \( \theta_{sat} \), estimated from the porosity, \( \theta_{fc} \) (measured at both 0.5 and 1m pressure head) and \( \theta_{PWP} \) (measured at 150m pressure head) were used to calculate the indicators AC, PAWC and RFC as described above (Table 2.4). It was not possible to calculate macroporosity (M) because measurements of the moisture content, \( \theta_m \) (occurring at 0.1m pressure head) were not given.

**Assessment of pedotransfer functions**

BD, texture (clay, silt and sand content) and organic C content are also available for the same soils in the LandIS database. Two types of PTFs were considered (Figure 2.6): The first represents the standard type PTF and is derived using multiple linear regressions (MLRs). The format of MLR models is:

\[
Y = a + b_1X_1 + b_2X_2 + \ldots + b_XX_X + E
\]  

(7)

where \( Y \) is the dependent variable; \( a \) is a constant; \( b_i \) are coefficients; \( X_i \) are predictor variables; and \( E \) is an error term. To assess the predictive power of the model, a 10-fold cross validation was implemented using the DAAG package (Maindonald & Braun, 2011). The significant variables were chosen by a stepwise selection procedure using the stepAIC function of the MASS package in R (Venables & Ripley, 2002).

The second PTF approach is an extension of MLR model, which can also consider categorical data such as 'Soil Series' and 'Land use' (Table 2.5). Multiple Additive Regression Splines (MARS splines) is a nonparametric regression technique that combines both regression splines and model selection methods (Friedman, 1991). It constructs a set spline basis functions (nonlinear functions) that are entirely determined from the regression data, and determines where they are applicable by automatically selecting appropriate knot values for different variables (in essence a multiple piecewise linear regression, where each breakpoint (or knot value) defines the "region of application" for a particular linear regression equation).

**Table 2.4. Soil Water Retention Characteristics: Fit results from PTFs based on LandIS data (BD, Clay, Silt and Sand and Organic C content)**

<table>
<thead>
<tr>
<th></th>
<th>( S_v )</th>
<th>( S_g )</th>
<th>Drainable Porosity</th>
<th>Plant Available Water</th>
<th>Relative Field Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSQconcR</td>
<td>RSQconcR</td>
<td>RSQconcR</td>
<td>RSQconcR</td>
<td>RSQconcR</td>
<td></td>
</tr>
<tr>
<td>multiple regression</td>
<td>0.56 0.73</td>
<td>0.82 0.85</td>
<td>0.53 0.65</td>
<td>0.58 0.72</td>
<td>0.61 0.78</td>
</tr>
<tr>
<td>MARS splines</td>
<td>0.72 0.75</td>
<td>0.87 0.9</td>
<td>0.71 0.08</td>
<td>0.68 0.82</td>
<td>0.73 0.85</td>
</tr>
</tbody>
</table>

RSQ = \( R^2 \) statistic; conc R = concordance correlation

**Table 2.5. Variable importance in MARS fitting. Numbers refer to rankings of importance**
In summary, we can obtain adequate performance ($R^2$ of 0.5 to 0.6) if we use standard regression approaches. This is significantly improved with MARS regression approaches using data which are available for England and Wales ($R^2 > 0.8$). It is therefore feasible to derive Soil Water Retention Characteristics SQIs using PTFs based on bulk density, texture and organic C measurements. In SP1305 (Subproject B), the sampling requirement, under different sampling regimes, is considered for BD and organic C, and the report suggests that for a plot of 20 by 20 m, 25 aggregated samples would be required. This is indicative of the sampling effort needed to determine the input data for the PTFs.
Figure 2.6. Biplots representing the predicted SQI versus observed SQI based on the MARS pedotransfer functions. Note SVI in these graphs refers to $S_v$ and Dexter S refers to $S_g$. 
2.7 Fact Sheet 3. Aggregate Stability

Background

Aggregate stability is a measure of the resistance of a soil to the destructive effects of rainfall, runoff and wind. The four main mechanisms responsible for surface soil aggregate breakdown by water are: (i) slaking, (ii) differential swelling, (iii) mechanical breakdown by raindrop impact and (iv) physico-chemical dispersion (Le Bissonnais, 1996). The breakdown of aggregates and reorientation of resulting fragments on the soil surface leads to surface crust formation, with higher bulk densities, restricted infiltration and greater runoff erosion hazard. Aggregate stability is an integrative indicator in that it reflects physical, biological and chemical soil properties. Aggregate stability is also multi-faceted as it reflects both soil functioning capacity (e.g. water regulation via infiltration rate) and susceptibility to degradation (e.g. soil erosion). Stable aggregates maintain soil structure in terms of a range of pore sizes, and thus promote soil aeration, water infiltration, drainage, better workability, seed bed quality and root penetration. More stable aggregates, with a higher proportion of large to small aggregates, suggest better soil quality, although this is unlikely to be a linear relationship. The stability of aggregates can be inferred by measurements of aggregate stability (Saygin et al., 2012).

To investigate the uncertainty in measurement; variability in observed distributions; expected rate of change; and spatial variability, we found only very limited datasets to assess aggregate stability across different land uses and soils at the national scale (England and Wales) (Thompson and Peccol, 1995; Merrington et al., 2006). However, Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration) used measurements of aggregate stability to describe and rank the likely behaviour of soil under the influence of rain. A number of simple aggregate stability tests were designed to allow rapid screening of a relatively large number of representative soil type/land use combinations in terms of their susceptibility to aggregate breakdown. General relationships between aggregate stability and measurements of a wide range of key physical and chemical soil properties were deduced, including easily-measurable soil properties such as organic carbon and clay content. The resultant database has been made available to the current project by Dr. Chris Watts and Prof. Andy Whitmore of BBSRC Rothamsted Research. However, as the authors of that work point out, the data generated from the field experiments is limited, in terms of geographical distribution, environmental conditions, simulated plot scale and replication.

Relationship to soil functions

Aggregate breakdown and subsequent aggregate size distribution affect soil processes and functions through their effect on:

i) Biomass production (provisioning function; Figure 2.7). Surface sealing by dislodged soil particles resulting from aggregate breakdown affects crop emergence. Surface seals also reduce infiltration capacity/hydraulic conductivity, so affecting soil moisture content, and availability of plant nutrients and water to roots. Some studies have related crop yields to aggregate stability (e.g. Skukla et al., 2004; Figure 2.7), but further analysis was not possible here, due to lack of data for England and Wales.
ii) Regulation of water and carbon: surface sealing resulting from aggregate breakdown affects infiltration/hydraulic conductivity through soil. Carbon regulation and sequestration are affected due to mineralization and loss of C on aggregate breakdown.

We considered this relationship as data in the Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration) contained measures of aggregate stability and the runoff observed in these experiments. In Figure 2.8a, we represent the observed relationship between aggregate stability (as measured by the 3 different methods used) and surface runoff. These data show only very weak or no relationship between aggregate stability and runoff generated; therefore currently there is no evidence that aggregate stability is an effective SQI of the water regulation function.

iii) Resistance to degradation – aggregate stability is strongly related to erosion susceptibility (it is regarded as the best estimator of erodibility/soil susceptibility to erosion (Bryan, 1968)), soil compaction and loss of C (Stavi et al., 2011). It also determines the degree of and susceptibility to soil surface sealing and capping.

Using data from the Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration) we analysed the relationships between aggregate stability and soil losses under simulated rainfall in the laboratory and in the field (Figure 2.8b). The soil loss results of the field work and laboratory trials were in contrast with those obtained using the three methods of aggregate stability determination, which used highly disturbed beds of aggregates.

**Sampling effort needed**

Usually individual aggregates are tested, although groups of aggregates can be tested under simulated rainfall. One major challenge is extrapolating results from the individual aggregate scale to larger spatial scales at which related soil processes (e.g. infiltration, erosion) and soil functions (e.g. provisioning) take place. For example, when relating aggregate stability to through flow (infiltration) and erosion, we found no clear relationships using the Defra SP0519 data (Figure 2.8b and 2.8c)

The high variability in aggregate stability as reported in the literature suggests a very high sampling intensity is required to detect changes in space and time. Aggregate stability is biologically mediated, so is likely to change during the growing season but there is only limited scientific literature on this.
Figure 2.8a. Biplot and linear fit between % Runoff and aggregate stability. Data from Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration). Experimental details can be found in the project report.

**Power analysis and sampling size.**

For aggregate stability, we base the power analysis on data obtained in the Defra project SP0519 in which particular aggregate fractions were associated with degrees of stability (very unstable, unstable, medium, stable, very stable). Three types of aggregate stability test were considered using various wetting conditions and energies (fast wetting (AS fast); slow wetting (AS slow); and mechanical energy applied after pre-wetting (AS mechanical)). If we consider the resulting degrees of stability as qualitative categories of aggregate stability (Table 2.6), then we can base the power analysis on the differences between these categories (Figure 2.9)
Figure 2.8b Biplot and linear fit between % through flow and aggregate stability. Data from Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration). Experimental details can be found in the project report.

Figure 2.8c Biplot and linear fit between total sediment eroded and aggregate stability. Data from Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration).
Experimental details can be found in the project report.

Table 2.6. Classification of aggregate stability based on mean weight diameter (MWD; after Le Bissonnais, 1996)

<table>
<thead>
<tr>
<th>MWD</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4 mm</td>
<td>Very Unstable</td>
</tr>
<tr>
<td>0.4 - 0.8 mm</td>
<td>Unstable</td>
</tr>
<tr>
<td>0.8 - 1.3 mm</td>
<td>Medium</td>
</tr>
<tr>
<td>1.3 - 2.0 mm</td>
<td>Stable</td>
</tr>
<tr>
<td>&gt;2.0 mm</td>
<td>Very Stable</td>
</tr>
</tbody>
</table>

Figure 2.9 Power analysis of three different measurement techniques to determine aggregate stability. Data from Defra project SP0519 (Critical levels of soil organic carbon in surface soils in relation to soil stability, function and infiltration).

However, aggregate stability measurements cannot be used in isolation as meaningful physical SQIs, as results must be interpreted in conjunction with information on soil type, changes in land use and management and climate (as reported at the Technical Workshop, Science Project SC030265; Merrington et al., 2008). Indeed, the sensitivity of aggregate stability to different factors is shown in data from Moebius et al., (2007; Table 2.7). These
findings are reflected in the database from Defra project SP0519, which shows measurements of aggregate stability are capable of discriminating between different soil types and management practices (Chris Watts, pers.comm.). These relationships are under investigation currently.

Table 2.7. Significance of factors affecting aggregate stability (WSA)

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Small WSA (0.25-2mm)</th>
<th>Large WSA (2-8mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>**</td>
<td>***</td>
</tr>
<tr>
<td>Soil type</td>
<td>**</td>
<td>***</td>
</tr>
<tr>
<td>Tillage</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Date x tillage</td>
<td>ns</td>
<td>**</td>
</tr>
<tr>
<td>Date x soil type</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Tillage x soil type</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Date x soil type x soil type</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

| Site 2                  |                         |                   |
| Date                   | ***                     | ***               |
| Tillage                | ***                     | ***               |
| Harvest                | **                      | *                 |
| Date x tillage         | ns                      | *                 |
| Date x harvest         | ns                      | ns                |
| Tillage x harvest      | ns                      | ns                |
| Date x tillage x harvest | ns                | ns                |

| Site 3                  |                         |                   |
| Date                   | ***                     | ns                |
| Soil type              | ns                      | *                 |
| Rotation               | *                       | **                |
| Date x soil type       | ns                      | ns                |
| Date x rotation        | ns                      | **                |
| Soil type x rotation   | ns                      | ns                |
| Date x soil type x rotation | ns          | ns                |

WSA = water stable aggregates. * p = 0.05; ** p = 0.01; *** p = 0.001; ns = not significant

Proxy techniques

Rawlins et al. (undated) describe aggregate stability measurements using a laser granulometer (LG) instrument, which overcomes the problems of sieve-based methods of measuring aggregate stability, namely a) the mass of stable aggregates is measured for only a few discrete sieve size fractions; b) no account is taken of the particle size distribution of the sub sampled material and c) they are labour intensive. LG technology has not been widely applied in aggregate stability testing. The difference in the continuous size distribution of WSA and disaggregated material (DM) is used to quantify the magnitude of the aggregated material. This is done by computing the difference in MWD (μm) between these 2 continuous distributions – referred to as the disaggregation reduction (DR) – the reduction in MWD on disaggregation by sonication. The authors argue this technique is reproducible and rapid, and does not misinterpret the presence of large mineral fragments in the disaggregated material (DM) as aggregates, which can be the case in wet-sieve approaches.
2.8 Fact Sheet 4. Rate of Erosion

**Background**

Soil erosion represents the physical loss of the soil resource, along with the functions (ecosystem goods & services) associated with that resource. This definition suggests erosion is a meaningful indicator of soil quality (Lal, 1998). Erosion is the detachment and transport of soil particles / aggregates from the *in-situ* soil mass, leading to a reduction in soil depth and volume (assuming soil bulk density is constant). Soil erosion is measured as the mass (tonnes) of soil lost per unit space (hectare) per unit time (year) (t ha⁻¹ y⁻¹). The change in the rate of erosion over space and time can also be monitored, although it is normally difficult to split out temporal changes due to management from changes in erosion rate due to changes in rainfall duration and intensity. Rates of soil erosion can be converted into depth of eroded soil using the equation:

\[
\text{Depth of soil loss per annum (mm)} = 0.1 \frac{\text{Mass of soil loss per annum (t ha}^{-1} \text{ y}^{-1})}{\text{Bulk density of the soil (Mg m}^3\text{)}}
\]

e.g. 10 t ha⁻¹ y⁻¹ = 0.7mm depth of soil lost per annum = 70mm/ century, assuming a soil bulk density = 1.4 Mg m³.

To investigate the uncertainty in measurement; variability in observed distributions; expected rate of change; and spatial variability, we reviewed possible data sources. Actual erosion rates are measured on field plots (associated with high levels of inter- and intra-plot variability in space and time; Brazier et al., 2012); through volumetric surveys of rill and gully features (although operator error may occur and accuracy is related to the number of readings; Evans, 2002); and via remote sensing (ground, air and satellite imagery – issues here relate to cost, frequency of image capture, resolution of images to capture erosion features as reported in Brazier et al., 2012).

Many monitoring schemes have concentrated on arable land and are biased as they tended to only consider land susceptible to erosion. Sheet or interrill erosion (caused by non-concentrated overland flow) is usually not monitored, but this can contribute a significant proportion of soil loss on upper and middle slopes (Morgan, 2005). As such, previous studies and the rates they measured are not representative of erosion rates on all soil / land use combinations in England and Wales.

Brazier et al. (2012) considered the application of relatively innovative Cs137 techniques to supply erosion monitoring data at the national scale (see Quine and Walling, 1991; Walling and Quine, 1991). Using these techniques in Defra SP0411 (Walling et al., 2005) and SP0413 (Walling et al., 2008), erosion rate data were collated for 248 fields, providing a basis for assessing the range of erosion rates occurring on agricultural land in England and Wales (Table 2.8). The data represent the first sizeable dataset of soil erosion rates for agricultural land in England and Wales that explicitly includes sheet erosion and includes values of both gross and net erosion.
Table 2.8. Comparison of estimates of gross and net erosion rates obtained for individual fields using the Cs137 technique with those obtained from more detailed sampling involving multiple transects and grid sampling (From Walling et al., 2005)

<table>
<thead>
<tr>
<th>Location</th>
<th>Land use</th>
<th>Gross soil loss (t ha(^{-1}) y(^{-1}))</th>
<th>Net soil loss (t ha(^{-1}) y(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cs137 technique</td>
<td>Intensive sampling</td>
</tr>
<tr>
<td>Crediton</td>
<td>Arable</td>
<td>6.2</td>
<td>7.1</td>
</tr>
<tr>
<td>Yeovil</td>
<td>Arable</td>
<td>9.7</td>
<td>8.7</td>
</tr>
<tr>
<td>Yeovil</td>
<td>Grass</td>
<td>2.1</td>
<td>1.7</td>
</tr>
<tr>
<td>Cadeleigh</td>
<td>Grass</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Crediton</td>
<td>Arable</td>
<td>8.3</td>
<td>9.1</td>
</tr>
<tr>
<td>Cadeleigh</td>
<td>Arable</td>
<td>6.2</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Given the data reported in the literature, we need to relate these rates of erosion with effects on soil processes and thus soil functions.

**Relationship to soil functions**

All soil functions are degraded by erosion due to the complete loss of soil profile. Production functions are affected by loss of soil depth / volume (see 2.9 Depth of Soil Indicator) and the inconvenience of erosion features interrupting farming operations. The regulation of water and carbon are affected by erosion through impacts on soil infiltration, flood risk increase, water quality decline, irregularity of rivers flow and increase of sedimentation in water bodies. Erosion causes perturbation of ecosystems and reduced ecological connectivity which both affect biodiversity and associated functions. As such, soil erosion indicates a degradation of soil quantity and quality (Huber et al., 2008). The impact of soil erosion on soil function and associated economic costs are reviewed in Defra SP1606. Soil erosion can cause a detectable change in soil quality, depending on:

a) the rate of erosion – low rates have no / little effect if they occur at a tolerable rate defined as “any mean annual cumulative (all erosion types combined) soil erosion rate at which a deterioration or loss of one or more soil functions does not occur” (Verheijen et al., 2009); and

b) inherent soil quality – some soils may not lose functioning as a result of erosion, depending on soil qualities such as soil depth, nutrient status and organic matter content (Lal, 1998). This can be captured in the concept of ‘soil loss tolerance’ (Verheijen et al., 2009)

These two points are captured in the concept of the life span of a soil. This is an extension of the idea of ‘tolerable soil loss’ and has been described as:

\[
T_c = \frac{S}{(E - P)}
\]

where, \(T_c\) is the critical life span of the soil, \(S\) is the initial soil thickness (m) of the A-horizon (see 2.9 Soil depth as an SQI below), \(E\) is the net soil erosion rate given by the difference between erosion and deposition (m yr\(^{-1}\)), and \(P\) is the soil production rate (m yr\(^{-1}\); representing the provisioning function) (Bui et al., 2010). The life span of a soil is the time taken to exhaust the soil below a point where it can no longer support production and occurs as a result of tolerable soil loss being exceeded.

Tolerable soil erosion has been defined by Boardman and Poesen (2006) as the rate of soil erosion no larger than the rate of soil production. However, Verheijen et al. (2009) suggest, that a more appropriate definition of tolerable soil erosion should be more holistic in its inclusion of soil functions and therefore should be ‘any actual soil erosion rate at which a
deterioration or loss of one or more soil functions does not occur’, where actual soil erosion means the cumulative amount of soil lost by all recognised erosion types. The term ‘tolerable soil erosion’ is preferable when referring to soil lost by erosion in the context of soil protection because it includes the removal of soil material by both physical processes (erosion), and biochemical processes (solute/gaseous export of mineral matter and decomposition of organic matter) (Verheijen et al., 2009). In Europe the tolerable soil erosion rate ranges from 0.3 to 1.4 t ha\textsuperscript{-1} yr\textsuperscript{-1}.

Verheijen et al. (2009) estimate that based on soil formation and erosion rates and soil management techniques at the time, topsoil in tilled arable land on hill slopes could be ca. 2 to 30 cm thinner in 100 years’ time (assuming a tolerable rate of 1 t ha\textsuperscript{-1} yr\textsuperscript{-1} and a bulk density of 1.3 t m\textsuperscript{-3}).

Erosion is both a dependent (e.g. OMC affects vulnerability of soil to erosion) and independent (e.g. erosion affects OMC levels; Quinton et al., 2006; Defra SP0519) variable in soil quality assessment.

There are many factors affecting rate of erosion and the change in that rate. This complicates understanding of its responsiveness / sensitivity to a change in any one or a combination of these factors. In other words, there is poor discrimination in the signal. Erosion rates are highly variable due to the many factors affecting rate. These factors are:

i. Rainfall erosivity, which is based on intensity, duration and kinetic energy of rainfall. “Climate – more specifically the combination of soil antecedent moisture conditions, rainfall intensity and duration – is a dominant driver of soil erosion by water. Its variation will lead to large differences in mean erosion rates between years for which measurements are made.”

ii. Soil erodibility, which is correlated to aggregate stability, soil texture, OMC, infiltration / permeability, bulk density and structure, is known to affect gross erosion rates (Table 2.10).

Table 2.10. Responsiveness of erosion rate to soil texture (From Walling et al., 2005)

<table>
<thead>
<tr>
<th>Soil texture</th>
<th>Number of sites</th>
<th>Gross erosion rates ( t ha\textsuperscript{-1} year\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Median</td>
</tr>
<tr>
<td>Clay</td>
<td>8</td>
<td>4.2</td>
</tr>
<tr>
<td>Clay loam</td>
<td>26</td>
<td>7.7</td>
</tr>
<tr>
<td>Silt loam</td>
<td>13</td>
<td>12.8</td>
</tr>
<tr>
<td>Sandy loam</td>
<td>15</td>
<td>9.9</td>
</tr>
</tbody>
</table>

iii. Slope length, gradient, curvature and plan all affect erosion rate (Morgan, 2005). Several equations exist to represent slope characteristics and associated changes in erosion rate. However, Evans and Cook (1986) reported the lack of a clear relationship between erosion rate and slope gradient in an analysis of rill erosion rates on agricultural fields (Figure 2.10).

iv. Land cover / land use (horticulture, arable, pasture, woodland, urban etc.) / land management practices are significant factors in determining soil erosion rates.
Research on soil erosion across England and Wales suggests that land use or land cover type is likely to exert a dominant influence on erosion rates; for example, using $^{137}$Cs as a tracer of soil erosion across 248 fields, it was shown that average erosion rates on cultivated land across England are substantially greater than for grassland (Defra, 2009).

**Figure 2.10. Responsiveness of erosion rate to slope gradient (From Walling et al., 2005)**

**Sampling effort needed**

Available erosion data pose a number of challenges to investigating the uncertainty in measurement; variability in observed distributions; expected rate of change; and spatial variability in erosion.

Monitoring erosion and the data generated occur at a range of spatial scales. “It is very difficult, if not impossible to acquire direct soil erosion measures for large areas” (Van Rompaey et al., 2003). Erosion rates are expressed as t ha$^{-1}$ y$^{-1}$, conveying a mean rate per unit area (i.e. 100m x 100m), but there may be considerable variability within the 1 ha unit of assessment. Another flaw in using this spatial scale of assessment is that erosion often occurs as linear features (rills, gullies) and even areal features (inter-rill / sheet erosion) do not extend to the hectare scale. Erosion has been measured and monitored over different spatial units:

- **Field / subfield plots** (e.g. Woburn, Wolverhampton, EU Life project: Soil and Water Protection in Europe; Cooper, 2007)

- **Volumetric surveys** (e.g. Evans and Boardman studies) – which cover hydrological erosion features alone, or depositional area of feature(s). Volumetric surveys of erosion rate are unlikely to monitor soil loss due to interrill erosion.

- **Fields / sub catchments** (e.g. remotely sensed data – aerial, satellite imagery or instrumented catchments)

One of the biggest challenges in erosion research is to reconcile the different spatial scales used (Table 2.11). Large scale studies (i.e. small field plots) are extrapolated to the ha$^{-1}$ scale, whereas small scale studies (i.e. large catchments) cannot distinguish individual erosion features. “Large extent monitoring is at odds with the scale at which erosion processes (and mitigation measures) operate” (Brazier et al., 2012).
Table 2.11. Range of spatial scales of soil erosion research (Rickson, 2006; after Wickenkamp et al., 2000).

<table>
<thead>
<tr>
<th>Erosion research technique</th>
<th>Area</th>
<th>Dimension descriptors (Wickenkamp et al., 2000)</th>
<th>Dominant processes operating</th>
<th>Selected references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Splash cup</td>
<td>mm²</td>
<td>Nanoscale Subtope</td>
<td>Rain splash dominant; overland flow/deposition limited. No gullies, stream bank erosion or mass movements.</td>
<td>Ellison (1944); Kinnell (1974); Morgan et al. (1988); Salles et al., (2000)</td>
</tr>
<tr>
<td>Laboratory tray</td>
<td>cm²</td>
<td>Nanoscale Subtope</td>
<td>Rain splash dominant?; overland flow/deposition limited. No gullies, stream bank erosion or mass movements.</td>
<td>Idowu (1996)</td>
</tr>
<tr>
<td>Runoff rig</td>
<td>m²</td>
<td>Microscale Tope</td>
<td>Rain splash and overland flow; some deposition possible. No gullies, stream bank erosion or mass movements.</td>
<td>Kamalu (1994); Govers (1991)</td>
</tr>
<tr>
<td>Field plot</td>
<td>m²</td>
<td>Microscale Tope</td>
<td>Rain splash and overland flow; some deposition. Some gullying and mass movements possible; no stream bank erosion.</td>
<td>Wischmeier and Smith (1978); Ciesiolka and Rose (1998); Pierson et al. (1994)</td>
</tr>
<tr>
<td>Field</td>
<td>ha</td>
<td>Mesoscale Chore</td>
<td>Rain splash, overland flow and deposition. Gullying and mass movements possible. No stream bank erosion.</td>
<td>Evans and Boardman (1994); Walling and Quine (1991); Chambers et al. (1992)</td>
</tr>
<tr>
<td>Sub-catchment</td>
<td>ha km²</td>
<td>Mesoscale Chore</td>
<td>Rain splash, overland flow and deposition. Gullying possible. Some stream bank erosion.</td>
<td>Hudson (1981); Rapp et al. (1972)</td>
</tr>
<tr>
<td>Catchment/landscape</td>
<td>km²</td>
<td>Macroscale Region</td>
<td>Rain splash, overland flow and deposition. Some gullying and mass movement possible. Stream bank erosion.</td>
<td>Dickinson and Collins (1998)</td>
</tr>
</tbody>
</table>

Simple ‘scaling up’ is not possible – it appears that the mean value of erosion per unit area will change when the spatial scale is increased, all other factors being equal. Unfortunately, this scaling up (or down) cannot be predicted because, according to van Noordwijk et al. (1998), there are no ‘scaling rules’. In other words, linear additivity is not valid for erosion studies (Pierson et al., 1994), as shown by Smith and Quinton (2000), when comparing erosion rates for different slope lengths (Table 2.12). Despite these apparent anomalies and errors, Smith and Quinton (2000) report that such scale extrapolations are commonly applied.
in erosion and sedimentation modelling. De Coursey and Meyer (1977) note that ‘by minimising the variability [by using smaller scales], we greatly reduce our ability to extrapolate to other areas’.

**Table 2.12. Non-linearity in soil erosion rate as a function of slope length (Smith and Quinton, 2000).**

<table>
<thead>
<tr>
<th>Slope length (m)</th>
<th>Erosion rate (kg/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>446</td>
</tr>
<tr>
<td>20</td>
<td>632</td>
</tr>
<tr>
<td>200</td>
<td>1095</td>
</tr>
</tbody>
</table>

In conclusion, the SI units require extrapolation from small spatial (eroded) areas to derive standard units of t ha\(^{-1}\) y\(^{-1}\). Within the support area, actual erosion rate will be highly variable. Rates of erosion are not linearly related to spatial scale (see Smith & Quinton, 2000). The m\(^2\) spatial unit is too small to account for runoff processes and would need many samples to be representative of a larger area. Erosion rates are often normalized to the ha\(^{-1}\) unit.

**Power analysis and sampling size.**

Brazier et al. (2012) noted that whilst there has been a great deal of erosion monitoring in the UK historically, it has not been placed within an unbiased sampling framework. The project proposed the following sampling design for a cost-effective, national erosion monitoring scheme:

- Frequency = annually, or bi-annually;
- Sample size = 300 sites for each landcover type (i.e. Arable; Improved Grassland; Upland / semi-natural) = 900 sites, based on power analyses to estimate the number of monitoring locations required to detect a specified (i.e. statistically significant) change in soil erosion rate on cultivated land (Table 2.13);
- Locations – Random, stratified (by land cover; not soil type) sample. “Soil erosion is known to vary substantially over short spatial (and temporal) scales and so it is unlikely that the covariances between the observations based on the geographical coordinates of the sampling locations would be of practical use in model based (geostatistical) inference. We therefore chose to pursue a design-based approach where inference is based on probability (random) sampling.”
- Monitoring (repeat sampling) at fixed sites preferred (but not essential). This has been done to a limited extent for upland erosion sites by McHugh (2007).

It should be noted that the data used in the analysis were drawn only from more erodible (cultivated) land use types, as they were the only available data to investigate water erosion over time. As similar datasets do not exist for gross erosion from all pathways (water, wind, tillage and soil loss through co-extraction on vehicles and during harvest), this analysis could not be performed for all erosion pathways. Brazier et al. (2012) note that in the future should these datasets become available, such analysis could be undertaken.

**Table 2.13. Results of the power analyses to detect specified change; 0.1 is 10% and 1 is 100% (from Brazier et al., 2012)**

<table>
<thead>
<tr>
<th>Total Sample Size</th>
<th>Power (p=0.05) to detect specified change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple Random Sampling</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>100</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>0.38</td>
</tr>
<tr>
<td>150</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
</tr>
</tbody>
</table>
According to Brazier et al. (2012) a common target power is 80%. The results show that a total sample of 300 would be adequate to achieve this power to detect a 50% increase in erosion rate (two-date comparison). Between 900 and 1000 samples would be needed to detect a 25% increase with the same power. A 10% increase is not detectable as a two-year comparison with 1000 samples, and clearly many more would be needed.

In any case, what constitutes a meaningful change in erosion rates still requires data on the impacts of erosion on soil functions in England and Wales, and will be dependent on highly variable factors such as specific crop (rooting habit, crop demands - water, nutrients); soil properties (water holding capacity; nutrient status) and external inputs (chemical fertilisers; irrigation).

**Proxy techniques**

Monitoring soil erosion by remote sensing (satellite, airborne or ground based) assumes erosion features produce detectable changes in the electromagnetic radiation recorded by the remote sensor (Milton and Webb, 1987; Milton et al., 1995), which must be distinguishable from radiometric distortions arising from the sensor (Collins and Walling, 2004). Measurements have to be at scales suited to erosion features (Kirkby, 2010), but most of the freely-accessible, regional-scale satellite systems (e.g. Landsat, ASTER) deliver data at spatial resolutions coarser than this. The new generation of sensors such as IKONOS may be more suitable, with spatial resolutions appropriate for erosion studies (e.g. <4 m). Remote sensing shows great potential for erosion monitoring, but image processing algorithms have to be better at considering pixel-to-pixel interactions.

Airborne survey data carry higher acquisition and processing costs, and the finer spatial resolution means reduced spatial extent. National coverage will generate vast volumes of data, with associated processing time. Airborne erosion surveys are most cost-effective at the catchment scale or less. Collection of data is rapid (e.g. 6,000-8,000ha can be acquired in a day), but rare erosion events are either over- or under-represented, unless regular monitoring takes place.

Evans (2005) describes two erosion monitoring projects using aerial photography which are the closest to a ‘national’ monitoring scheme. The first project (1982 – 1986) involved 17 transects throughout England and Wales, covering a range of landscapes and most lowland soil types. Each year ≈700 km$^2$ was photographed at a scale of 1:10,000. Rill and gully erosion and sediment deposition were recorded and checked in the field. In 1989 a much smaller scheme (covering ≈11km$^2$) was carried out by ADAS, with some air photos taken of the field sites surveyed (Evans, 2005).

In conclusion, most small scale monitoring (e.g. using satellite imagery) is at odds with the scale of erosion processes, local risk factors and mitigation measures.
2.9 Fact Sheet 5. Depth of Soil

Background

Depth of soil is sometimes considered as a proxy indicator of soil health. Change in soil depth may occur as a result of soil degradation, for example loss of soil through erosion. Whatever the mechanism for change, soil depth as an indicator here relates to a change in the soils’ capacity to provide certain functions relating to storage, regulation and provision. Robinson et al. (2010) note that soil depth would be a helpful addition to a suite of SQIs as soil production vs erosion is not well understood.

Effective soil depth defines the volume of soil in which roots can grow (effective rooting depth) and from where they can retrieve water and nutrients. As depth of soil is reduced, the volume of soil available for plant roots to exploit is diminished. This can limit available water and nutrients to the plant, so stunting above- and below ground plant growth and affecting production yield and quality. Reduced soil depth may also affect the stability of a plant, making it vulnerable to lodging. Soil depth also defines the depth to which micro-organisms are active and may also impact on nutrient and carbon cycling and availability to plants. Soil depth affects the ability to buffer and therefore regulate chemicals due to the reduced residence time for solutes moving through the soil system (although other soil properties will affect this too).

Soil depth is sensitive to land management and therefore reflects management decisions that may affect the soil’s ability to function in a desirable way. A change in soil depth can be an indication of how well a soil is being managed, although there are natural reasons for changes in soil depth such as shrink swell behaviour. A sustainably managed soil will either maintain soil depth or increase in soil depth. Intensive land management can lead to an increased vulnerability of the soil to erode and/or to compact, and as a consequence lead to a reduction in soil depth.

The properties of soil depth over space and time pose a number of challenges to monitoring, which must take into account any uncertainty in measurement; variability in observed distributions; expected rate of change; and spatial variability.

Soil depth varies naturally in the landscape (Table 2.14). It is affected by rate of soil formation, which itself is affected by climate, the characteristics of the underlying bed rock, relief and time. Soil depth, in particular the A horizon, is also dependent on the rate of accumulation of organic material and its decomposition and whether the site is actively eroding or is a depositional zone. Landscape topography therefore influences the depth of soil even at a very local level. For example, hillslopes tend to have shallower soils because the soil is more actively eroding, while valley bottoms have deeper soil because soil tends to be deposited there. Even on a single hillslope there will be local variations due to localised slope irregularities. Depth of soil can also change because of soil compaction caused by a compressive force such as the weight of a vehicle or animal/human traffic. Soil depth is altered by extraction e.g. removal of topsoil during construction or collection of peat for sale. Topsoil depth can also be influenced by the type, depth and frequency of tillage.

Table 2.14 Soil depth classes used in the Australian Soil Classification (Australian Natural Resources Atlas)

<table>
<thead>
<tr>
<th>Class</th>
<th>Soil depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very shallow</td>
<td>&lt;0.25</td>
</tr>
<tr>
<td>Shallow</td>
<td>0.25 - &lt;0.5</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.5 - &lt;1.0</td>
</tr>
<tr>
<td>Deep</td>
<td>1.0 - &lt;1.5</td>
</tr>
<tr>
<td>Very deep</td>
<td>1.5 - 5</td>
</tr>
</tbody>
</table>
The depth of soil is easily measured either by digging a soil profile or by auguring a hole and measuring either total depth of the soil profile or depth of topsoil only. However, boundaries between horizons are not always smooth or distinctive, and a range of depths may need to be noted and then expressed as mean depth and variation. The total depth of soil can be particularly difficult to determine precisely.

Relationship to soil functions

Productivity depends largely on the thickness and quality of the topsoil and on the nature of the subsoil. Productivity declines as soil thickness reduces (Figures 2.11 – 2.13), however, anthropogenic intervention can be used to mask the effect of soil depth loss. For example, more fertiliser can be added to replace lost nutrients and artificial irrigation can be used to overcome reduced soil water storage potential in the soil. Therefore, any comparison between soil depth and productivity (yield) needs to consider land management practices that may have been used to compensate for a reduction in the soils function. This combined information is currently rarely available and complicates the use of soil depth as a meaningful SQI.

Figure 2.11 Relationship between crop production and loss in soil depth (Stallings, 1964)

Figure 2.12 Relationship of wheat biomass production and soil depth measured in hilly areas of Greece (Kosmas et al., 2001)
One approach to link soil depth to soil function is through the concept of soil ‘lifespan’ (see 2.8 Rate of erosion SQI above). Soil depth provides a measure of change in soil volume and therefore a change in soil water storage capacity, organic carbon store and rooting media. In a deep soil, a reduction in soil depth may have little meaning other than a reduction in overall soil volume. This is sometimes described as a tolerable soil loss (see 2.8 above). However, continued reduction in soil depth may eventually reach a critical depth, a point at which the function of the soil is altered. Gomez et al. (1996) defined the trigger or threshold level of soil-depth as being 50 cm or average of similar soil types in the community, but this is theoretical only as in reality it will be highly dependent on crop (type, species and variety; Table 2.15), weather (i.e. water availability) and level of management inputs (e.g. supplementary irrigation and/or nutrients).

From a productivity perspective the plant root system also requires an effective depth of soil over which to forage for water and nutrients i.e. its working depth. Different plants have different working depths (Table 2.15)

Table 2.15. Rooting depths of crops

<table>
<thead>
<tr>
<th>Crop</th>
<th>Rooting depth (m)</th>
<th>Crop</th>
<th>Rooting depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asparagus</td>
<td>3.0</td>
<td>Maize</td>
<td>1.8</td>
</tr>
<tr>
<td>Tomatoes</td>
<td>1.8-3.0</td>
<td>Sugar beet</td>
<td>1.5-1.8</td>
</tr>
<tr>
<td>Cucumber</td>
<td>1.1</td>
<td>Peas</td>
<td>1.1</td>
</tr>
<tr>
<td>Celery</td>
<td>0.3-0.5</td>
<td>Mustard</td>
<td>1.1</td>
</tr>
<tr>
<td>Lettuce</td>
<td>0.4</td>
<td>Onions</td>
<td>0.3</td>
</tr>
<tr>
<td>Potatoes</td>
<td>0.9</td>
<td>Wheat</td>
<td>1.0-1.5</td>
</tr>
</tbody>
</table>
Grass | 0.5-1.5

Monitoring soil depth is further complicated, because given certain physical soil characteristics and underlying parent material, two soil depths can be distinguished: the critical and the crucial soil depth. The critical depth can be defined as the soil depth in which plant cover achieves values above 40% (see Figure 2.14). Extensive studies conducted under semi-arid climatic conditions in hilly areas of Greece with soils formed on various parent materials have shown a critical depth of 25-30 cm. On soil less than that depth the recovery of the natural perennial vegetation is very low and the erosional processes may be very active, resulting in further degradation and desertification of the land. When a hilly landscape of marginal capability is cultivated, agriculture should be abandoned before the soil reaches the critical depth. It should be noted that these data are from Greece where available soil water during key wheat growth stages is often limiting, which is less likely to be the case for wheat production in England and Wales.

![Figure 2.14. Relation of percentage vegetation cover of Sarcopoterium sp and soil depth measured in areas with soils formed in pyroclastics magmatic conglomerates) and schist-marble (Kosmas et al., 2000).](image)

**Figure 2.14. Relation of percentage vegetation cover of Sarcopoterium sp and soil depth measured in areas with soils formed in pyroclastics magmatic conglomerates) and schist-marble (Kosmas et al., 2000).**

While the critical depth is a limit to cultivation, the crucial depth can be defined as a lesser soil depth on which soil-protective perennial vegetation can no longer be supported, and the whole soil is rapidly washed away by wind or water erosion. This is an irreversible process. Crucial depth is affected by the type of parent material in which soil is formed. Soils formed on pyroclastics are the most sensitive in their capability, with a crucial soil depth of 10 cm below which vegetation cannot be supported. Soils formed on schist-marble metamorphic rocks have a higher capability to support perennial vegetation under the same climatic conditions, with crucial soil depth of around only 4-5 cm.

It is important to define these critical limits so that soil is taken out of a degradative land use before it reaches the point of irreversible degradation. These critical limits vary among soils and ecoregions, and are not known for principal soils (Lal, 1993). This demonstrates the challenges of sampling and analysing data on soil depth (or change in depth) at which soil functions are affected, as is required in a meaningful monitoring programme.
(Although the provisioning and regulating functions are the focus in the present study, Davidson and Wilson (2006) identified ‘depth of superficial deposits removed by mineral extraction’ could also be a relevant indicator for the preservation of cultural heritage, with some development).

**Sampling effort needed**

Soil depth is naturally variable even within a field. Therefore several depth measurements across a field/area are needed to obtain an ‘average’ soil depth. It is sometimes difficult to define a lower depth boundary for individual soil profiles because of restrictions due to depth. Observations using survey pits are only practical to 1-2 m, while surveys using soil augers are limited to the depth of the sampler. It is plausible that changes in depth ranging in the tens of centimetres would be possible to detect in a monitoring program. Soil depth can be assessed at a single point or across a landscape by using multiple point samples. Any soil depth sample should be representative of the soil type and the landscape it represents. Measurements of either linear change in soil depth or volume are typically used to monitor for changes in soil depth over space or time. However, as explained above, choice of critical or crucial depths may be used to define tipping points.

Changes in depth may also be based on estimates of soil erosion, although, the reliability of this to describe change in depth, for example over a field area, is not necessarily appropriate when relating it to soil function because erosion may only alter the soil depth over a very limited proportion of the field e.g. linear features such as rill and gully erosion. Pimental et al. (1995) used a simplified method to estimate the impact of soil erosion on soil depth:

\[
\text{Depth of soil loss per annum (mm)} = 0.1 \frac{\text{Mass of soil loss per annum (t ha}^{-1} \text{yr}^{-1})}{\text{Bulk density of the soil (Mg m}^{-3})}
\]

For a soil eroding at a mean rate of 17 t ha\(^{-1}\) yr\(^{-1}\) with a bulk density of 1.25 Mg m\(^{-3}\) the annual loss of soil depth was estimated to be 1.36 mm yr\(^{-1}\). Estimates of soil depth loss for a number of land use/soil type combinations due to soil erosion by water have been given by Graves et al. (2011) for UK soils (see Table 2.16), although other degradation processes affect soil depth too, such as compaction and mineralisation of peat. Pimental et al (1995) equated loss in soil depth to a yield losses of 0.74% per mm of soil (See Table 2.17)

**Power analysis and sampling size.**

In the case of soil depth as a potential SQI, we restrict our analysis to consider a particular stratum or type of soil which is vulnerable to changes in soil depth i.e. those which are denoted as shallow soils in LandIS (Soil Survey of England and Wales). We obtain the measure of spatial variability in depth from the soil profile database and use this to calculate the sample numbers required to determine a change in soil depth. The results from the power analysis are represented in Figure 2.15, which also shows the spatial coverage of the shallow soils in England and Wales.

**Table 2.16 Soil depth loss (due to soil erosion) in each land use/soil type category (Graves et al., 2011)**

<table>
<thead>
<tr>
<th>Land use</th>
<th>Soilscapes</th>
<th>Soil depth loss (mm yr(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>Clay</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Silt</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Sand</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>Peat</td>
<td>0.00</td>
</tr>
<tr>
<td>Horticulture</td>
<td>Clay</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Silt</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td>Sand</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>Peat</td>
<td>4.39</td>
</tr>
<tr>
<td>Arable intensive</td>
<td>Clay</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Silt</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>Sand</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>Peat</td>
<td>5.22</td>
</tr>
</tbody>
</table>
Table 2.17 Yield penalty induced by loss of soil depth (as developed from Pimental et al., 1995; Appendix G of Graves et al., 2011)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Yield penalty (%) per mm loss in soil depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water runoff</td>
<td>0.3</td>
</tr>
<tr>
<td>Nitrogen Phosphorus Potassium</td>
<td>0.3</td>
</tr>
<tr>
<td>Soil depth</td>
<td>0.3</td>
</tr>
<tr>
<td>Organic matter</td>
<td>0.1</td>
</tr>
<tr>
<td>Water holding capacity</td>
<td>0.1</td>
</tr>
<tr>
<td>Soil biota</td>
<td>0.0</td>
</tr>
<tr>
<td>Total on-site</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Figure 2.15. Results from the power analysis for shallow soils in England and Wales. Insert shows the spatial coverage of the data points i.e. the distribution of shallow soils. Note that the y-axis is on a logarithmic scale.
2.10 Fact Sheet 6. Soil Sealing

Background

Soil sealing in this context describes the sealing-over of land (soil) by expanding urban infrastructure (e.g. roads, buildings and pavements). Soil sealing presents one of the greatest threats to soil function. Remote sensing has improved the certainty of measurement and provided base-line data on the proportion of sealed soils within urban areas (e.g. see Wood et al., 2004). However, different urban infrastructure can have different levels of sealing, e.g. roofs, concrete, asphalt are 100% sealed, paving slabs with seep-able joints - 70%; green roofs, gravel, crushed stone - 50%; (Meinel and Hering 2005). Policy decisions (through planning regulation) and individuals’ choice (e.g. block-paving of front gardens for car parking) both affect the levels of soil sealing.

Over time, monitoring the change in the proportion of sealed soils will allow assessment of changes in soil functions. For example, the water regulating function of soil is lost by surface sealing. This increases flood risk due to higher runoff coefficients from paved surfaces. The extent of the built environment can be estimated from high resolution satellite imagery (<1 m ground resolution) and aerial photography. Both are capable of including narrow corridors, such as road and rail systems, as well as providing accurate estimates of residual soil areas within the built environment. The utility of remote sensing data can be improved through integration with other datasets, e.g. OS MasterMap or soil maps, enabling sealed or unsealed areas to be resolved at a finer spatial scale.

The key indicators for soil sealing are:

1. The absolute area of sealed soil (ha)
2. The change/growth rate of area of sealed soil (ha yr⁻¹, ha d⁻¹, % change to baseline)

Regarding expected rate of change, de-sealing (or negative sealing) can also be derived. This is where sealed soil is regained or restored completely or partially by removing artificial soil covers. This will require very high resolution (<1m) monitoring data as the de-sealed areas are often small and fragmented. Examples of de-sealing are the installation of green roofs, porous block paving and grass car parking using geotextiles.

Historically soil sealing was quantified from statistical analysis of national cadastral maps produced by traditional civil engineering survey of buildings and infrastructure using tape measures and theodolites. Later, monitoring of land cover and change was implemented by aerial photo interpretation (API) supported with ground survey. API relies on subjective human decisions, therefore requires skilled operators and is time consuming. Its accuracy is assessed by cross-checking for validation between interpreters and comparison with ground survey. Uncertain cases can be resolved using ‘multiple-eyes’ where a consensus opinion is reached. API is still considered de facto as the most accurate procedure for mapping. The availability of earth observation (EO) remote sensing satellites from the mid 1970’s to their widespread use for spatial analysis of the landscape. The GMES (Global Monitoring for Environment and Security research programme) has facilitated the production of remote sensing based soil sealing layers (e.g. EEA data mentioned above). The primary advantage of satellite remote sensing is large spatial and temporal coverage and lower unit cost of imagery.

Relationship to soil functions

Urban development of inner-city green-space and the conversion of rural and agricultural land at the urban fringe contribute most to deterioration or loss of soil function due to sealing.
Urban sealed surfaces are covered by pavements, buildings and transport infrastructure creating hard, very slowly permeable surfaces. The sealing affects natural processes including the surface water cycle (infiltration, groundwater filtering and evapo-transpiration), geochemical cycles (e.g. reducing soil fertility) and energy transfers. It has climatic impact, through altered surface albedo and air temperature, creating 'urban heat islands'. It increases surface water runoff, increasing erosion and flood risk. It alters and generally reduces the habitats for biodiversity and in some cases the soil is completely removed (Huber et al. 2007). According to Kibblewhite et al. (2007) soil sealing is probably the most serious threat to European soils.

However, as reported in Wood et al (2005), there are few quantified data demonstrating the effect of soil sealing on soil function, which limits detailed analysis of data.

**Sampling effort needed**

*Medium-High Resolution EO satellites:*

The most successful and widely used satellite series is NASA’s Landsat which combined high resolution (30 m) with multispectral (7 bands) sensing capability, ideal for land cover classification. Currently there is no fully functioning Landsat imagery collection since Landsat 7 developed a technical fault in May 2003 and Landsat 5 was decommissioned last year. However, there is a large archive of imagery available which can be used to study changes in soil sealing. Among others, the Disaster Monitoring Constellation (DMC) from SSTL is providing similar specification EO imagery, but with only 3 spectral bands (G, R, NIR). SPOT 5 (10 m) imagery is also widely used for land cover mapping. Sensors with R and NIR bands, sensitive to vegetated surfaces (un-sealed) can be used to calculate vegetation indices, the most widely used is NDVI (Normalised Difference Vegetation Index), (Tucker, 1979). At this scale (10 – 30 m) pixel based digital classifications can be used to automatically characterise the landscape in imagery based on probabilistic pixel level digital image processing. The most widely used classification algorithm is Maximum Likelihood (Richards and Jia, 2005). This quantitatively evaluates the spectral response of each pixel in each spectral band and assigns it to a land cover class based on its statistical probability of class membership in multi-dimensional vector space. This technique has proven to be very robust provided imagery is collected at suitable timing where land cover classes exhibit spectral differences. For monitoring soil sealing in the UK context, urban areas are often easily separated due to large spectral contrast between vegetation and urban infrastructure. The disadvantage of medium resolution imagery is that sealed areas less than 1 to 2 times the pixel area cannot be resolved.

*Very High Resolution (VHR) EO data:*

In the last decade there has been increasing availability of very high resolution multispectral satellite imagery (<5 m) and digital aerial photography collections with <1 m spatial resolution. The pixel level digital classification methods are found to be less effective with these very high resolution datasets. This is due to the scale of observation compared to the pixel size, texture affects are visible in the imagery, and the spectral response of a cover class is disaggregated into its constituent spectral signatures. An example would be a building, at medium resolution (e.g. 20 m) every pixel spectral response is a mixture of reflectance values from the different materials and shadows which provide a unique spectral mix for buildings. In very high resolution imagery the different materials and shadows have different spectral responses. The configuration of these gives a textural pattern and also provides a context for a photo-interpreter to make a decision. Using a pixel classifier with very high resolution data, the increased variation in the statistical definition of a ‘building’ class decreases classification accuracy (Caprioli and Tarantino 2003). To attempt to
overcome these limitations alternative image analysis techniques such as object based classifiers have been developed.

**Object based classifiers for high resolution EO data:**

An image object is a group of pixels defined by a homogeneous region based on spectral, textural and contextual information. They are defined by a process known as segmentation (Baatz and Schape 2000). The segments are then classified using various methods including expert knowledge, fuzzy rules (Baatz et al. 2000) or statistical methods. Other polygon data such as OS MasterMap can be used as alternative segments and in some studies as ancillary data (Smith et al. 2007). The biggest challenge for implementation of sealing monitoring using object based classifiers is the cross calibration between different EO imagery from different sources with different timing. The segmentation strategy (which determines the object scale) cannot be transferred between different imagery, therefore requires optimising for each case. Within each image, the segments identified do not always represent real features (Kampouraki et al. 2008), for example, shadows cast by buildings or trees result in separate objects. In contrast, using traditional API the interpreter has the intelligence to identify the different types of shadow and avoid misclassification. The advantage of automatic segmentation is the speed of data processing across an entire image scene. This has led to semi-automated classification, where API is used to manually edit and then classify the objects. This significantly increases the speed of delineating features as the objects only require editing, either merging or splitting. This also provides the opportunity to re-segment complex or uncertain areas at a finer scale.

**Other Remote Sensing Datasets:**

In addition to optical remote sensing, data derived from radar such as LIDAR (Light detection and ranging) or SAR (Synthetic Aperture Radar) at <1 m resolution could be used to determine the elevation of buildings and vegetation canopies or to determine the surface types below tree canopies (and measure through cloud).

**Sampling considerations**

The use of remotely sensed information allows population estimates to be made in the imaged area at the pixel resolution. There is a trade-off between coverage and spatial resolution. For example, VHR IKONOS (1 m) has a coverage of 10 x 10 km. Medium resolution DMC imagery (20 m) has a coverage of up to 250 x 600 km.

Scale is an important consideration here. It is likely that in the rural context, medium resolution (20 m) is sufficient to detect the consumption of land into sealed surfaces. In the urban context very high resolution data (<1m) is desirable to monitor smaller sealed areas, such as gardens converted to driveways, or large patio areas. The de-sealing, such as fitting porous block-paving or fitting a green roof will also require high resolution data to detect changes.

Building development that causes soil sealing takes place over many months or even years, and results in a step change in land cover. Very high resolution imagery collections are scheduled on a rolling basis, with urban areas being updated approximately every 3-5 yrs and rural areas 5-10 yrs. Therefore a monitoring schedule at very high resolution would be at best 3-5 yrs. Medium to High resolution imagery is collected more frequently and has larger coverage, so it would be possible to monitor on a yearly basis.
2.11 Fact Sheet 7. Visual Soil Evaluation

Background

Soil visual evaluation has been used since the advent of soil mapping in which a formal profile description forms the basis of soil classification used by soil surveyors (Hodgson, 1976). However, the semi-quantitative scoring of soil features is a relatively recent innovation. Many of the visible and tactile characteristics used by traditional soil surveyors have been encapsulated into rapid in-field methodologies of describing the impact of land management on soil quality. Various field-based assessments of soil structure quality have been developed in several countries (Ball et al., 2007) although the origins of semi-quantitative visual soil assessment is often attributed to Peerlkamp (Peerlkamp, 1959).

There are a range of methods that vary from the evaluation of soil physical conditions in the topsoil only, through to whole profile descriptions. However, the different methods do have some commonality. Visual soil evaluation (VSE) is based on the qualitative or semi-quantitative evaluation of soil properties and soil indicators e.g. morphological, physical, biological and chemical, which are visible or possible to distinguish without laboratory analysis (Housková, 2005).

VSE is used to assess soil quality by comparing the state of the soil against a reference sample (i.e. either a photograph or a soil sample known to be in optimal condition). Dynamic soil indicators (e.g. aggregate size and stability, water content and porosity) change under different management regimes and land-use pressures. The sensitivity of these soil features to change provides a useful early warning indicator of changes in soil condition and an effective method of monitoring change in soil condition. The methods assess visual indicators of soil quality. For example, the New Zealand Visual Soil Assessment (VSA (NZ)) method (Shepherd, 2008) considers soil texture, soil structure, soil porosity, soil colour, number and colour of soil mottles, earthworm (number and size), rooting depth, surface ponding, surface crusting/cover and soil erosion (wind/water).

VSE is most widely employed to provide a measure of soil structural quality and changes in soil structural quality over time in a field. Soil structure is an important component of soil quality and therefore an important parameter for consideration as a soil quality indicator. The main objective of VSE is to evaluate the effects of cultural operations on soil structure looking at the intrinsic soil quality of soil structural stability. Table 2.18 illustrates the criteria used by each of 10 VSE methods explored by Boizard et al. (2005) to assess soil structure. In addition to these 10 VSE methods, Think Soils (Environment Agency, 2008) is a pictorial reference guide that is designed to help farmers and advisers to identify soil structural problems, but does not provide a system for scoring soil condition.

Relationship to soil functions

Some VSE methods have made direct links to soil functions. The VSA (NZ) index scores have also been shown to be closely related to crop yield, pasture dry matter production, biomass cover and pasture utilisation in New Zealand (Shepherd and Park, 2003). On an experimentally manipulated silt loam and sandy loam arable soils in Germany and Canada, Mueller et al. (2009a) reported that for mean Peerlkamp scores ranging from 3 to 7.5, there was a significant increase in cereal grain yield of 300-350 kg/ha per unit of ‘St’ score.

Holman et al. (2003) developed a scheme, ‘Soil Structural Degradation’, to consider the impacts of soil quality (as assessed through VSE) on the flood regulation function (i.e. runoff generation and transmission). The most useful feature of any VSE method is its ability to reveal the abundance and arrangement of soil aggregates and roots as these properties are affected by soil management and reflect a number of soil functions (Mueller et al., 2012).
Table 2.18 Techniques used to assess soil structure (taken from Boizard et al., 2005)

<table>
<thead>
<tr>
<th>METHOD</th>
<th>Knife</th>
<th>Drop test and sieving</th>
<th>Drop test and visual</th>
<th>Stability test</th>
<th>Soil unit morphology</th>
<th>External porosity</th>
<th>Internal porosity</th>
<th>Color</th>
<th>Water content</th>
<th>Rooting</th>
<th>Biological activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Whole profile assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 SOILpak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Le profil cultural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Peerlkamp score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 CWSSC (B. Murphy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 VSA (NZ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Soil quality scoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 VSA (Denmark)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 FAL method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Guide to better structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data sources: uncertainty in measurement; variability in observed distributions; expected rate of change; and spatial variability.

One criticism of VSE techniques is the challenge of quantifying the findings, thereby providing comparable data in space and time (which is essential for an effective monitoring programme). However, a number of studies have linked VSE findings to quantifiable soil properties. The VSA (NZ) method put forward by Shepherd (2000) has been widely used in New Zealand and has been trialled in at least 10 other countries throughout the world. One such trial of the VSA (NZ) was conducted over 91 sites on 40 soil types from different parent materials, climate, topography and under different land uses including grassland, arable, orchards and forestry. The VSA scores were found to be significantly correlated to corresponding lab-based methods (Shepherd et al., 2002; reported in Boizard et al., 2005) including dry aggregate-size distribution ($r^2=0.91$), saturated hydraulic conductivity ($r^2=0.86$) and air permeability ($r^2=0.80$), and moderately correlated to macroporosity ($r^2=0.69$) and dry bulk density ($r^2=0.64$). VSA (NZ) colour scores are strongly related to total carbon ($r^2=0.80$) and moderately related to anaerobic mineralisable N ($r^2=0.64$) in conventionally cultivated mineral soils.

Muller et al. (2009b) concluded that the scores produced by the Peerlkamp method, Diez score and VSA (NZ) correlated with physical measurements of soil structure including dry bulk density and penetration resistance measurements. Mueller et al. (2009a) also found unfavourable visual structure associated with lower infiltration rates and dry bulk density (Figure 2.16). However, relationships between structural indices and measured parameters were site-specific. Indicators based on several parameters seem to be more robust and reliable to characterise physical soil quality (Muller et al., 2009a).
The Peerlkamp method is likely to be one of the most sensitive to defining change due to land management activity because of the detailed descriptions used to define each score level. Even with the simplified revised version, which defines only five soil quality categories (Table 2.19), the detailed descriptions that accompany the categories enhances the ability of the observer to categorise samples more confidently.

Mueller et al. (2012) found that compaction status of soils with clay contents > 30% could not be reliably determined either by dry bulk density measurements or penetrometer readings. However, they found that visual observations could determine structural differences in these soils.

**Sampling effort needed**

There are a number of techniques used in VSE (Table 2.20) and no standard in terms of the methodology or how the data is used. A review of 10 different VSE methods was presented by Boizard et al. (2005). They reviewed 3 whole profile techniques and 7 topsoil techniques. The methods differed in the depth and area of soil examined and number of replications (see Table 2.20). Those based on soil profile evaluation (0-1.5m depth) have the advantage that the subsoil can be examined and anthropogenic limitation to root growth (e.g. soil compaction), as well as intrinsic quality of deeper layers can be determined. Topsoil examination (0-30cm depth) using a spade may not reveal much about subsoil conditions although some of these methods can be extended down to include lower layers as well (e.g. Peerlkamp and VSA(NZ)).
Field scale variability is taken into account in different ways. For whole profile evaluation, only one pit (or trench) per plot or land unit is normally used, however, the pit can be extended across variable features such as trafficked and untrafficked areas of the field and is assumed to be representative of the whole plot or land unit area.

For topsoil techniques, samples are usually collected either randomly or targeted into areas e.g. trafficked and untrafficked. Number of replications varies from 1 to 10 depending on the methods and the time available. In the evaluation by Boizard et al. (2005) the Peerlkamp method was the only one with a sufficient number of results to enable a statistical evaluation to be made of the differences between the topsoil structure of the plots. However, the number of replicate samples depends on the time allowed to undertake the analysis.

Another uncertainty surrounding VSE outcomes is the continuity in results between different surveyors / monitors (reliability). The VSA (NZ) method has been validated as a robust method whereby laypeople with little or no background in soil science could use the method and come up with similar answers to an ‘expert’ (Figures 2.17 and 2.18; Boizard et al., 2005). Boizard et al. (2005) found that the coefficient of variation in assigning an index score to each soil ranged from 0.05 to 0.24 with a mean of 0.12.
Table 2.20 Sample techniques used, area evaluated and replication of 10 different VSE methods (taken from Boizard et al., 2005).

<table>
<thead>
<tr>
<th>METHOD</th>
<th>Depth</th>
<th>Sampling technique</th>
<th>How layers are delimited</th>
<th>How lateral variability is taken into account</th>
<th>No. of replicate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Whole profile assessment</td>
<td>Up to 1.5m</td>
<td>Soil pit</td>
<td>Delimited by visual porosity, soil strength and soil water content</td>
<td>Pit length flexible</td>
<td>1</td>
</tr>
<tr>
<td>2 SOILpak</td>
<td>Up to 1.5m</td>
<td>Soil pit</td>
<td>Delimited by visual porosity, soil strength and soil water content</td>
<td>Pit length flexible</td>
<td>1</td>
</tr>
<tr>
<td>3 Le profil cultural</td>
<td>Up to 1m</td>
<td>Soil pit</td>
<td>Delimited by surface appearance, visual porosity soil strength and cohesion</td>
<td>Pit length flexible to include each wheeling</td>
<td>1</td>
</tr>
<tr>
<td>4 Peerlkamp score</td>
<td>0.25m</td>
<td>spade</td>
<td>Distinct layers can be given separate scores</td>
<td>Random or selected to be representative</td>
<td>10</td>
</tr>
<tr>
<td>5 CWSSC (B. Murphy)</td>
<td>0.10m</td>
<td></td>
<td>No</td>
<td>Random</td>
<td>?</td>
</tr>
<tr>
<td>6 VSA (NZ)</td>
<td>0.25m</td>
<td>spade</td>
<td>Same as Peerlkamp</td>
<td>Disturbed and non disturbed</td>
<td>Up to 3</td>
</tr>
<tr>
<td>7 Soil quality scoring</td>
<td>0.25m</td>
<td>spade</td>
<td>Same as Peerlkamp</td>
<td>Tracked zone can be scored separately</td>
<td>3</td>
</tr>
<tr>
<td>8 VSA (Denmark)</td>
<td>0.30m</td>
<td>spade</td>
<td>Same as Peerlkamp</td>
<td>random</td>
<td>1</td>
</tr>
<tr>
<td>9 FAL method</td>
<td>0.45m</td>
<td>spade</td>
<td>Delimited by changes in appearance. Distinct layer can be given separate scores</td>
<td>Lateral replicates of samples*</td>
<td>1</td>
</tr>
<tr>
<td>10 Guide to better structure</td>
<td>0.25m</td>
<td>spade</td>
<td>Distinct layers can be given separate scores</td>
<td>random</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 2.17 VSA of the condition of 12 soils under arable cropping – A comparison of ‘expert’ and workshop participant rankings (taken from Boizard et al., 2005).

Figure 2.18. VSA of the condition of 13 soils under pastoral grazing on flat to rolling country – A comparison of ‘expert’ and workshop participant rankings (taken from Boizard et al., 2005).
It is now recognised by the scientific community that the use of good reference material (e.g. photographs) in combination with well-defined descriptions of several soil properties, which can be used to assign a soil structural quality score, provides a more robust and consistent classification system that enhances the reproducibility of the results (Muller et al., 2012). The revised version of the Peerlkamp method is an example of where this has happened to good effect (Table 2.19).

While the reproducibility of visual features can be, to some extent, aided by good visual reference material, the reproducibility of tactile observations (e.g. soil texture) is more dependent on subjective descriptions, for example see Figure 2.19. However, Hodgson et al. (1976) demonstrated that with experience and the use of reference material against which observers can ‘calibrate’ themselves, it is possible to produce a confident estimate of soil texture for a wide range of soils. In their study, they found that 75.5 % of the variation in field estimates of silt content and 85.4 % of the variation of clay estimates could be predicted by hand texturing (Figures 2.20 and 2.21).

![Hand texturing guide](image-url)
Figure 2.20 Relationship between field estimates ($\text{Silt}_e$) and laboratory analysis ($\text{Silt}_a$) of silt (taken from Hodgson et al., 1976)

\[
(\text{Silt}_e) = 1.122 + 0.930(\text{Silt}_a) \\
\quad r = 0.969
\]

Figure 2.21 Relationship between field estimates ($\text{Clay}_e$) and laboratory analyses ($\text{Clay}_a$) of clay (taken from Hodgson et al., 1976)

\[
(\text{Clay}_e) = 2.794 + 0.904(\text{Clay}_a) \\
\quad r = 0.924
\]
Could VSE be made numerical through ratings?

One criticism of using VSE as a physical SQI is the lack of quantification by which meaningful change (i.e. associated change in soil function) in the SQI is detected. There are both qualitative and semi-quantitative VSE methods (see Table 2.21).

**Table 2.21 Comparison of visual soil evaluation methods and criteria of observations (adapted from Boizard et al., 2005)**

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference</th>
<th>Key criteria</th>
<th>Assessment</th>
</tr>
</thead>
</table>
| Whole profile assessment      | Batey (2000)       | • Soil texture  
• Soil colour  
• Development, strength and stability of structure  
• Soil compaction  
• Degree of fissuring  
• Presence of roots | Qualitative. Considers inherent capability of the soil and limitations as a result of the management of the soil |
| SOILpak                       | McKenzie (1998)    | • Size, shape, strength and internal porosity of primary clods and aggregates | Factors are weighted and scored between 0.0 (poor) and 2.0 (good). Beginners use 3-points (0, 1, 2), experienced use 5-points (0.0, 0.5, 0.1...), experts use 21-points (0.0, 0.1, 0.2...,2.0) |
| Le profile cultural           | Richard et al. (1999) | • Transition between the tilled layers  
• Internal structural state of clods or zone  
• Structural state | Qualitative assessment |
| Peerlkamp score               | Peerlkamp (1959)   | • Size, shape and porosity of clods and aggregates  
• Stability and dispersion on the surface  
• Actual or potential root development | Scored from 1 (worst) to 9 (best)  
Modified version (Ball et al., 2007) Scored Sq1 (best structure) and Sq5 (worst structure) with comparison photographs |
| CWSSC                         | Lawri et al (2000) | • Soil texture  
• Stability and resilience of structure | Non-quantitative |
| VSA (NZ)                      | Shepherd (2000)    | • Aggregate size, shape and abundance  
• Porosity  
• Colour  
• Mottles  
• Erosion  
• Earthworm count | Eight indicators are assessed on a scale of 0.0 (poor) to 2.0 (good) (intermediate scores also given e.g. 0.5). Judgments based on comparison with photographs and an undisturbed reference sample. A weighting factor (1, 2 or 3) is applied to exaggerate more important soil condition indicators e.g. soil structure |
| Soil quality scoring          | Ball and Douglas (2003) | • Identification of horizontal layers, depth and thickness  
• Structure  
• Macropores | Separate scores are given for each criteria on a scale of 1 (worst) to 5 (best), using best-fitting descriptions from a
In summary, VSE methods have the potential to be used by researchers, consultants and farmers. Boizard et al. (2005) identified the potential user groups who may, after appropriate training, make use of different VSE methods (see Table 2.22). Most VSE methods can be taught quickly and easily and do not require the skills of a dedicated soil surveyor. Good reference material helps maintain continuity between observers. However, periodic cross checking results against other observers is recommended to maintain consistency and eliminate any bad habits.

Semi-qualitative VSE methods provide a value or description against which temporal changes in score and rankings (the proportion of soils in ‘good’, ‘moderate’ or ‘poor’ condition) can be gauged and that can be recorded as an on-going reference. The use of photographs as visual records of temporal change is recommended i.e. a visual archive that can be related to temporal changes

Semi-quantitative visual evaluation methods could be used to detect large changes in structural quality of the ‘whole topsoil’ over time and for different land uses (e.g. arable and grassland). A significant cost in monitoring schemes is the time taken for surveyors to travel to monitoring sites and take soil samples. For minimal additional cost (45 minutes to 1 hour

<table>
<thead>
<tr>
<th>Method</th>
<th>Scoring Criteria</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSA (Denmark) Munkholm (2000)</td>
<td>- Roots and fauna scoring sheet.</td>
<td>No summary into a single score</td>
</tr>
<tr>
<td>FAL method (see Boizard et al., 2005)</td>
<td>- Size distribution</td>
<td>Score scale 1 (worst) to 14 (best). Pictures and key codes help classify clod and aggregates into 12 different types.</td>
</tr>
<tr>
<td>A guide to better soil structure</td>
<td>- Compaction</td>
<td>Reference to photographic guide assigning degradation class: severe, high, moderate, low</td>
</tr>
<tr>
<td>Soil structural degradation Holman et al. (2003)</td>
<td>- Surface soil condition, Presence of wheeling or tramlines, Extent of poaching, Structural changes within or at the base of topsoil, Presence of erosion and deposition features, Vertical wetness gradient within profile</td>
<td>Qualitative descriptions categorised into 4 classes (severe, high/extensive, moderate/local, low)</td>
</tr>
<tr>
<td>Think Soil EA (2008)</td>
<td>- Surface erosion features, Crop, Soil structure, Root development</td>
<td>Qualitative</td>
</tr>
</tbody>
</table>
per site) visual soil evaluation could be carried out to gain a semi-quantitative assessment of soil structural condition that could be monitored over time for individual sites and for the whole dataset (e.g. does the proportion of sites in good condition change over time?).

Table 2.22 Potential users of each VSE method and the time taken to perform the test in a field (taken from Boizard et al., 2005)

<table>
<thead>
<tr>
<th>METHOD</th>
<th>Research</th>
<th>Consultants</th>
<th>Farmers</th>
<th>Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Whole profile assessment</td>
<td>×</td>
<td>×</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>2 SOILpak</td>
<td>×</td>
<td>×</td>
<td>(×)</td>
<td>30 or 90</td>
</tr>
<tr>
<td>3 Le profil cultural</td>
<td>×</td>
<td>×</td>
<td>(×)</td>
<td>60 to 180</td>
</tr>
<tr>
<td>4 Peerlkamp score</td>
<td>×</td>
<td>×</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>5 CWSSC (B. Murphy)</td>
<td>×</td>
<td>×</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>6 VSA (NZ)</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>30 to 120</td>
</tr>
<tr>
<td>7 Soil quality scoring</td>
<td>×</td>
<td>×</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>8 VSA (Denmark)</td>
<td>×</td>
<td>×</td>
<td></td>
<td>60 to 120</td>
</tr>
<tr>
<td>9 FAL method</td>
<td>×</td>
<td>×</td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>10 Guide to better structure</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>30</td>
</tr>
</tbody>
</table>

As with other physical SQIs, data uncertainty also comes from timing of assessments. Generally, VSE methods are not easily used when the soil water content is high and the soil is plastic. Neither are they easy to use when the soil water content is very low and cohesion may be very strong. Spade techniques are also limited in stony soils particularly when the soil is dry. Concerns over difficulties in breaking up aggregates along natural boundaries by inexperienced operators can be overcome by applying the drop test. However, Guimarães et al. (2011) concluded that both dropping and breaking by hand produced the same score and the latter was easier and quicker.

It is recognised that some visual characteristics take time to develop and therefore a recently compacted site may have fewer visual references than a site where compaction has been present for several seasons. However, repeat observations will help build up a picture over time that will help define if soil degradation is increasing, decreasing or stable.

As with any other assessment of soil physical properties, VSE methods require training before a surveyor can use the method in the field with supervision and eventually unsupervised.
2.12 Integration of physical SQIs

Each candidate physical SQI has been considered independently, but we also considered the relationships between the SQIs. There is evidence that combining the measurements of physical (and biological and chemical) SQIs improves the relationship between the SQIs, soil processes and soil functions. In justifying the use of bulk density / packing density as a meaningful SQI (WP1), we have already seen the empirical evidence suggesting better prediction of air capacity by combining bulk density and clay content (i.e. Packing density) (Figure 2.21).

Figure 2.21. Improving the relationship between air capacity and measures of bulk / packing density (from Huber et al., 2008)

A similar exercise was done here where data were available (i.e. PD, BD, aggregate stability and depth), with the *a priori* reasoning that a number of the physical SQIs would be inter-dependent and therefore exhibit spatial clustering. However, the resulting cluster analysis of the data showed no spatial patterns. This reflects the complexity of relationships between the SQIs and paucity of data. In any case, any relationship between the parameters would have to be expressed in terms of soil processes and functions for it to be meaningful. Even the improved relationship between air capacity and packing density (combining clay content and BD) is only an empirical, rather than explanatory relationship (Huber et al., 2008).

The current project has been focussed on physical properties as soil quality indicators, without explicit reference to biological or chemical SQIs. According to Karlen et al. (2003), soil quality assessment must reflect biological, chemical, and physical properties, processes and their interactions. Taking this more holistic view, Karlen *et al.* (1997) note the interdependence among soil quality indicators and that the concept of soil quality should be “an umbrella concept for examining and integrating relationships and functions among various biological, chemical and physical parameters that are measured and important for sustainable agricultural and environmental systems.” They suggest that “field, farm and watershed-scale evaluations of soil quality require a transition from an experimental mode that contributes to an understanding of soil quality to more interdisciplinary approaches”. Arshad and Martin (2002) identified the interrelationships between a number of key soil quality indicators (physical, biological and chemical; Table 2.23).
Table 2.23. Interrelationships of key soil quality indicators (Arshad and Martin, 2002).

<table>
<thead>
<tr>
<th>Selected indicator</th>
<th>Other soil quality indicators affecting the selected indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregation</td>
<td>Organic matter, microbial activity (especially fungal), texture</td>
</tr>
<tr>
<td>Infiltration</td>
<td>Organic matter, aggregation, electrical conductivity (ESP), exchangeable sodium percentage</td>
</tr>
<tr>
<td>Bulk density</td>
<td>Organic matter, aggregation, topsoil depth, electrical conductivity, biological activity</td>
</tr>
<tr>
<td>Microbial biomass and / or respiration</td>
<td>Organic matter, aggregation, bulk density, pH, texture, electrical conductivity</td>
</tr>
<tr>
<td>Available nutrients</td>
<td>Organic matter, pH, topsoil depth, texture, microbial parameters (mineralisation and immobilisation rates)</td>
</tr>
</tbody>
</table>

These interrelationships have been demonstrated in Fact Sheet 2 (Soil water retention characteristics), where basic soil properties (bulk density, texture and organic C) are used to predict soil water retention characteristics (Dexter’s S, drainable porosity, plant available water and relative field capacity) through pedotransfer functions.

Even so, we need to relate these relationships to soil processes and functions. Shukla et al. (2004) carried out step wise regression analysis with corn grain yield as the dependent variable and measured soil attributes as the independent variables (Table 2.24).

Table 2.24 Linear regression analysis for total dry corn grain yield (Yg, Mg ha\(^{-1}\)) as dependent variable and measured soil attributes as independent variables (P<0.009)

<table>
<thead>
<tr>
<th>Equation</th>
<th>R(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One parameter model</td>
<td></td>
</tr>
<tr>
<td>(Y_g = 3.753 + 0.007 \text{WSA})</td>
<td>0.33</td>
</tr>
<tr>
<td>Two parameter model</td>
<td></td>
</tr>
<tr>
<td>(Y_g = 3.717 + 0.424 \text{AWC} + 0.008 \text{WSA})</td>
<td>0.34</td>
</tr>
<tr>
<td>(Y_g = 3.353 - 0.422 \text{MWD})</td>
<td>0.34</td>
</tr>
<tr>
<td>(Y_g = 14.321 + 0.006 \text{WSA} - 1.435 \text{pH})</td>
<td>0.34</td>
</tr>
<tr>
<td>Three parameter model</td>
<td></td>
</tr>
<tr>
<td>(Y_g = 23.569 + 0.009 \text{WSA} - 0.906 \text{MWD} - 2.808 \text{pH})</td>
<td>0.37</td>
</tr>
<tr>
<td>(Y_g = 24.061 + 0.007 \text{WSA} - 2.809 \text{pH} - 2.449 \text{EC})</td>
<td>0.36</td>
</tr>
</tbody>
</table>

where: AWC – available water (cm); WSA – water stable aggregates (g kg\(^{-1}\)); MWD – mean weight diameter (mm); EC – electrical conductivity (dS m\(^{-1}\))

The criticism with this approach is that a) regressions give no explanation of why combining these soil properties improves prediction of soil functioning (here, yield); b) the relationships and equations would vary for different soil functions (e.g. provisioning v. regulation); c) as reviewed in the WP2 fact sheets above, there are few equivalent data for England and Wales linking soil attributes (i.e. physical SQIs) to soil functions; d) this approach takes no account / makes no allowance of how a change in one or more attributes over time (which is at the core of a soil monitoring programme) may change the interrelationships between soil attributes; e) the current England and Wales database on physical soil attributes is too limited and contains too much variability to derive rigorous regression equations.

It has been demonstrated that for statistical rigour, the sampling design of each SQI is based on the detectable / meaningful change in that property (i.e. power analysis). As a result, different SQIs will require different sampling designs, which undermines the concept of
having multiple, integrative SQIs, unless the largest sampling size is taken for all. Given the cost implications, we conclude it is impossible to design and implement a rigorous sampling scheme for multiple SQIs.

Another limitation of integrative SQIs is error propagation when using more than one SQI. Error in one or more input variables (e.g. measurement / sampling error) means the resulting integrative SQI is not robust (Murray Lark, pers.comm.).

2.13 Conclusions to WP2

Data on the physical SQIs selected by the logical sieve process in WP1 are limited in spatial and temporal extent for England and Wales. This has limited the degree of analysis and modelling that it was possible to test:

- the uncertainty in their measurement
- the spatial and temporal variability in the indicator, as given by observed distributions
- the expected rate of change in the indicator
- the impacts on soil processes and functions associated with a change in any given soil attribute.

Where data are available, power analysis was used to understand the variability of the indicator as given by the observed distributions. This process determines the SQIs ability to detect a particular change at a particular confidence level, given the ‘noise’ or variability in the data (i.e. a particular power to detect a change of X at a confidence level of Y% would require N samples).

The paucity in data has limited the specification of a robust sampling design required to detect changes in the SQI that are associated with a meaningful change in any soil process and/or function. In any case, there is little scientific evidence to define what is meant by a ‘meaningful change’ in any given SQI and this will vary for different soil types, land covers and soil functions. Attempts were made to stratify detectable changes in packing density and soil depth by land cover and soil type (Fact Sheet 1 and 5), but no relationships were found. If this had been the case, we would need to relate these changes to associated changes in soil processes and soil functions, but the evidence base is very weak to do this. Schipper and Sparling (2000) identify the challenge: “a standardised methodology may not be appropriate to apply across contrasting soils and land uses. However, it is not practical to optimise sampling and analytical techniques for each soil and land use for extensive sampling on a national scale”.

Soil hydrological relationships are key to soil functioning, especially in water regulation (availability to plants and avoidance of both flooding and drought; Merrington et al., 2006). Here, the analysis of available data has given promising results regarding the prediction of values of SQIs from easy to measure soil properties (bulk density, texture and organic C), using pedotransfer functions. These measurements also allow estimates of packing density which is also related to soil hydrological properties (e.g. air capacity; Huber et al., 2008). Expanding the evidence base (space and time) should be possible with the development and use of rapid, cost-effective techniques using NIR sensors to gain measurements of bulk density, texture and packing density.

More evidence is needed as to how aggregate stability affects soil processes and thus soil functions. We have found an equivocal relationship with water regulation (runoff generation), which is affected by the method of measurement (low or rapid wetting or mechanical disturbance of the aggregates) for example. There are insufficient data on ‘meaningful’ change in aggregate stability i.e. how a change in that property affects processes and thus
functions. Gathering such data would rely on better, more cost-effective, standardised measurement techniques.

Brazier et al. (2012) used power analyses to estimate the number of monitoring locations required to detect a statistically significant change in soil erosion rate on cultivated land. On the basis of this analysis, the project proposed a cost-effective framework to monitor (and model) soil erosion in England and Wales, with an emphasis on constructing a statistically sound approach to locating sampling sites for soil erosion monitoring. However, what constitutes a meaningful change in erosion rates still requires data on the impacts of erosion on soil functions in England and Wales.

At present there is insufficient quantified evidence for statistical rigour to be applied in determining a sampling strategy for Visual Soil Assessments / Evaluations. Some studies have begun to investigate how VSA might be moved to a quantified scale. Use of remote sensing to detect changes in surface sealing has potential, although the spectral signatures received are not directly measurements of soil quality in terms of physical properties and their influence on soil processes and functions. Wood et al. (2004) notes that in general, remote sensing techniques are most useful when considering indicators that relate to soil quantity and have limited utility for those indicators that relate to soil condition (quality). In any case, as reported in Wood et al (2005), there are few quantified data demonstrating the effect of soil sealing on soil function, which limits detailed analysis of data.

The scientific robustness of these indicators depends on the availability of spatial and temporal data, as these reflect the variability of each property (signal to noise ratio), which in turn determines the sampling strategy required to detect significant change. Whether that change is meaningful depends on the evidence base relating soil properties to soil processes and soil functions. The evidence base is poor: data on meaningful (i.e. what degree of change affects soil processes and functions) and detectable (i.e. what sample size is needed to detect the meaningful signal from the variability or noise in the signal) are lacking at present to propose a rigorous and reliable soil monitoring programme.
References
Bui EN, Hancock GJ, Chappell A and Gregory LJ (2010) Evaluation of tolerable erosion rates and time to critical tipsoil loss in Australia. CSIRO Land and Water, Canberra


Department for Communities and Local Government have detailed the National Planning Policy Framework (March 2012),


Gaţe, O.P., Czyż, E.A., Dexter, A.R. 2006. Soil physical quality, S, as a basis for relationships between some key physical properties of arable soils. *Proceedings of...*
International Soil Tillage Research Organisation, 17th Triennial Conference 28 August–3 September, Kiel, Germany pp 258-264

GeoInformation Group 2012, National Land Use Map
http://www.geoinformationgroup.co.uk/products/national-lu


Houšková B (2005) Visual Soil Assessment. 3rd European Summer School on Soil Survey, co-organized by TAIX, JRC and SZIU.


Watts CW Whalley WR Longstaff, DJ White RP Brookes PC Whitmore AP 2001. Aggregation of soil with different cropping histories following the addition of different organic materials. Soil Use and Management 17, 1–6


APPENDIX 1.
Notes of a meeting held on 18/04/12 to discuss progress on WP2 Data analysis and modelling

Background
To summarize, a set of possible physical SQIs\(^2\) were evaluated through the logical sieve. A number of indicators were identified in WP1 and Defra are generally happy with these. These are made up of the top 25% overall scoring indicators and all SQIs that passed the logical sieving process. So the indicators under consideration are now:

<table>
<thead>
<tr>
<th>Physical SQI’s</th>
<th>Champion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bulk density (FDR and VIS-NIR spectroscopy method)</td>
<td>AM</td>
</tr>
<tr>
<td>2. Depth of soil</td>
<td>JR; LD</td>
</tr>
<tr>
<td>3. Infiltration / drainage capacity</td>
<td>PNP; CK</td>
</tr>
<tr>
<td>4. Moisture storage capacity / Soil water retention characteristics</td>
<td>PNP; CK; TW</td>
</tr>
<tr>
<td>5. Number of erosion features</td>
<td>JR</td>
</tr>
<tr>
<td>6. Packing density</td>
<td>BJ; PNP; LD</td>
</tr>
<tr>
<td>7. Profile description</td>
<td>LD; PNP</td>
</tr>
<tr>
<td>8. Rate of erosion</td>
<td>JR</td>
</tr>
<tr>
<td>9. Sealing</td>
<td>TW</td>
</tr>
<tr>
<td>10. Shear strength (Going Stick method)</td>
<td>IJ</td>
</tr>
<tr>
<td>11. Soil structure</td>
<td>CK LD</td>
</tr>
<tr>
<td>12. Soil texture (wet sieving and sedimentation methods)</td>
<td>CK LD</td>
</tr>
<tr>
<td>13. Surface water turbidity</td>
<td>IH</td>
</tr>
<tr>
<td>14. Total porosity</td>
<td>CK RC</td>
</tr>
<tr>
<td>15. Bulk density (Kopecki ring)</td>
<td>PNP</td>
</tr>
<tr>
<td>16. Erodibility / aggregate stability</td>
<td>JR</td>
</tr>
<tr>
<td>17. Biological status</td>
<td>IH</td>
</tr>
<tr>
<td>18. Number of eutrophication incidents</td>
<td>IH</td>
</tr>
</tbody>
</table>

AM Abdul Mouazen BJ Bob Jones CK Cedric Kechavarzi
IH Ian Holman IJ Iain James, JR Jane Rickson
LD Lynda Deeks PNP Paul Newell Price RC Ron Corstanje
TW Toby Waine

Feedback from Defra / Declan Barraclough of key issues to be addressed in Work Package 2
- Robustness (reliability and accuracy) of the indicator, both statistical and practicability
- Relevance – can it realistically be interpreted in terms of a key soil function?
- Synergy – is there scope for combining with another indicator to greater effect?
- Cost

Moving forward, there are two objectives related to WP2:
- **Objective 2.1**: Evaluation of new scientific evidence, reports, data, SQI selection criteria, sampling techniques and monitoring methodologies aimed at identifying salient physical SQIs (by the end of Project Month 8)
- **Objective 2.2**: Identification of suitable physical SQIs from the findings of Objectives 1.1 and 2.1 above, based on agreed selection criteria, sound science, statistical robustness and feasible data collection methods (by the end of Project Month 9)

\(^2\)Key indicator criteria are (see EPA’s [Report on the Environment](#)):

An indicator is a numeric value derived from actual measurements of a pressure, state or ambient condition over a specified geographic domain, whose trends over time represent or draw attention to underlying trends in the condition (or “health”) of the environment.

- the indicator is useful;
- the indicator is objective;
- the indicator is transparent and reproducible;
- the underlying data is characterized by sound collection methodologies, data management systems to protect its integrity, and quality assurance procedures;
- data are available to describe changes or trends; and
- the data are comparable across time and space, and representative of the target population.
Regarding the scientific evidence; there are, for all intents and purposes, two soil functions that are directly affected by changes in soil physical properties: i) the production function and ii) the regulation function. These indicators would be measures that indicate a degree of degradation in these soil functions. What we are trying to get a feel for is the effect-response curve of these indicators to degradation in these soil functions and, from this, the cost effectiveness of the indicator. In summary, what we need to know for each indicator is

1. What is it indicative of; i.e. what function is being degraded?
2. What it is responsive to? How responsive is it? sensitivity, responsiveness
3. What factors may mitigate or accentuate the response? e.g. soil type, land use
4. Is this indicator a first order indicator (i.e. a direct measure of the change in soil quality) or a second, third, etc order indicator (i.e. an indirect measure of the change; e.g. VNIRS, remote sensing)?

Regarding reports, data. In order to meet objective 2.2, we will need to analyse suitable candidates, and for this we will need data, so

5. Please identify existing or suspected data-holdings for each indicator.

Regarding SQI selection criteria, sampling techniques and monitoring methodologies. We will need to reduce this list of indicators as I don’t think it feasible to analyse all 18 SQI’s. The logical sieve has helped us rank and organize much of the information currently, but a national monitoring scheme cannot include 18 physical SQIs due to the resource implications nor will we be able within this project to investigate 18 SQIs

6. How is it measured?
7. What sampling support does it need?
8. What is the sampling intensity required?

Notes of the meeting 18/04/2012
Present: Abdul Mouazen, Paul Newell Price (ADAS), Lynda Deeks, Pat Bellamy, Iain James, Cedric Kechavarzi, Jane Rickson, Ron Corstanje.
Apologies: Ian Holman, Toby Waine.
Aims:
- Revisit the outcome of WP1, especially the Logical Sieve exercise to ensure the results were sensible and that no indicators were disqualified unduly.
- Move from the ‘narrative’ of WP1 to a more ‘numerical’ approach in WP2 i.e. given the remaining SQIs, are data available to test the robustness of the candidate SQIs through statistical/modelling analysis?
- Rationalise the remaining SQIs in terms of duplications (e.g. total porosity = 1 - BD/particle density); overlaps /double counting (e.g. no. of erosion features and rate of erosion); and linkages (e.g. BD and depth – the Logical Sieve consider all candidate SQIs as independent of each other – are composites useful?)
- Consider scale issues e.g. scaling up / aggregating SQI performance at the field (or equivalent on non-ag land) to larger landscape units.

Initially a number of general points were discussed regarding the concept of soil quality indicators. What makes a ‘good’ soil indicator and why?

a) It needs to be sensitive and responsive to the loss or degradation of soil functions. Points were raised regarding the soil functions
   i. It was generally agreed that the most common functions we’ll consider are the production and regulation functions.
   ii. Not every indicator will be appropriate for any given function e.g. loss in depth of soil impacts regulation function immediately, but less immediate on production function if soils are deep and rooting depth is not limiting.
   iii. The basic spatial unit for assessing soil functions is the field scale. This is because it is expected that this is the most appropriate scale to link the relationship between soil properties and the production /regulation function of soils. It is the scale at which management options would be practiced. At the larger scale (e.g. catchment), these links are confounded by other factors. At the smaller scale (e.g. pedon), soil functions are difficult to measure.

b) Sensitive and responsive – this was not discussed in much detail, but we can hypothesise response curves, plotting change in SQI against change in function. Each SQI might produce
different curves for different functions. The statistical analysis of the selected indicators should generate a better understanding of this. The assumption underlying much of the discussion is that the main driver of change in soil quality is land use and land management. Do we need to worry about climatic factors/climate change?
c) There are many qualifiers to the effectiveness of physical indicators of soil quality. Particular indicators might be more meaningful (or more sensitive) under particular circumstances (soil type, landuse). For example, soil depth as a SQI of biomass production capacity is more meaningful on shallow soils on chalk escarpments than where soil depth is not a limiting factor. Peat soils in particular were recognized as a category which might require the development of different indicators than those suitable for mineral soils. We discussed specifying particular indicators to particular circumstance in a matrix format, but this implies physical SQIs may not be universal to all soil/land use/function combinations.
d) We realized that certain indicators are measurements made on the soil that are not necessarily good indicators, rather process measurements which relate to indicators. This allowed us to simplify the list considerably e.g. erodibility is a proxy for rate of erosion; visNIR isn’t the indicator – BD is.

On specific indicators
1. Bulk density (FDR and VIS-NIR spectroscopy method) - Champion AM. AM presented three methods to determine BD in the field:
   a) A combination of Visual and Near Infrared measurements, combined with Theta probe determinations for soil moisture. This is currently an experimental method and has not been tested on a wide range of soil type/land use combinations
   b) A combination of soil resistance (penetrometer measurements) and Vis-NIR measurements to determine BD. This is a destructive method that has been successfully trialled on 17 fields representing 10 different soil textures
   c) The ‘on the go’ system is placed on tines behind a tractor where BD is measured as a function of Vis-NIR and soil resistance. It provides a good estimate of BD at the time of cultivation (i.e. post harvest or prior to seedbed preparation in the spring), but cannot provide within crop measurements without causing some crop damage. The system has been trialled in most soil texture types in the UK.

Note: we had some discussion here about BD as an indicator and AM indicated that for yield, there are relationships between BD and yield representing the production function. This will require further investigation and statistical analysis to test the validity of using BD as a SQI.

The discussion continued throughout the day relating BD to soil functions and culminated in the concept of packing density (see below). In summary, BD values are not a ‘good’ indicator of soil physical quality in isolation, but they can be used to derive other indicators/variables that are more closely related to soil quality and performance e.g. packing density. This requires further investigation as BD was the only meaningful physical SQI identified by Merrington et al. (2006).

2. Depth of soil. JR discussed some published response curves here between depth of soil and crop productivity / biomass / yield, from linear approaches to crucial/critical values. In this case, the use of critical and crucial values is probably a better approach, but these will depend on crop type / species and their sensitivity to soil depth. We also noted that this might be a good indicator for degradation in peat soils. Data holdings identified were LandIS.

In summary, any reduction of soil depth (due to processes such as erosion, wastage/shrinkage, compaction) is likely to have an impact on the regulation function of soil (e.g. water holding capacity, carbon content), but can also impact the production function, although this may be only once a critical threshold (depth) is reached, which in turn depends on site and time specific conditions (e.g. crop and weather). Data are available to test this, so this indicator was retained.

3. Infiltration capacity (CK, PNP). A very pertinent indicator to the regulation function but also the production and support functions. In essence, if the soil cannot infiltrate rainfall or irrigation water then this will contribute to run off and therefore flooding events, erosion, and diffuse pollution. Reduced infiltration can either be linked to a degradation of the soil surface only (capping, crustng, repellency) or the upper layer of the soil profile (compaction, poor structure). For the production function, infiltration capacity can therefore affect water availability to plants. However, the sensitivity of established measurement techniques to temporal changes in infiltration capacity
may be confounded by other factors such as initial moisture content and high spatial variability. In addition these methods are time consuming, costly and require a large number of replicates as the measure itself is highly variable. Methods that can relate the measure to intrinsic soil properties regardless of initial and boundary conditions may be preferred. These would include traditional tension infiltrometry from which near saturated hydraulic functions can be derived but the time required for the measurements is prohibitive. An alternative is the single tension measurement method such as that used by the Decagon infiltrometer. Data – ADAS has data and it may be possible to compare this to a study by Holtan-Kirkpatrick. These measures are highly site specific and very temporally variable. So, despite the importance of this soil property as a SQI, the robustness and reliability of its measurement using currently available techniques are questionable.

4. Moisture storage capacity / soil water retention characteristics indicators - Soil physical and structural indicators derived from the soil water release characteristics (CK, PNP). The soil water release curve is a fundamental soil hydraulic function which informs on soil hydrological properties highly pertinent to the regulation and production functions. From this curve a series of indicators can be derived which inform on pore size distribution and therefore structural quality and the capacity for the soil to store and retain water in appropriate proportions (important for plant production, water regulation, nutrient cycling and pollutant transport). Furthermore, associated with saturated hydraulic conductivity measurements, knowledge of the release characteristics allow modelling of the soil water dynamics and the water balance (infiltration, run off, storage and drainage).

From four single points of the release characteristics (saturation, moisture content at 10 cm, field capacity and permanent wilting point) 6 key indicators can be derived: plant available water capacity, air capacity (or drainable porosity), relative field capacity, macroporosity, porosity of soil matrix and total porosity (derived from saturation rather than bulk density). Alternatively these indicators can be derived individually from either one or two points of the release curve. From the full water release curve (6 to 10 measurement points) a pore size distribution function can be derived mathematically and used to obtain further key indicators of hydraulic quality: the S index, an overall index of structural quality which has been correlated to other soil quality indicators such as bulk density and organic matter, and location and shape parameters.

Optimal ranges for all the above indicators have been developed in the literature for managed soils. As with hydraulic parameters in general, these indicators are more sensitive to temporal changes in soil quality than other indicators such as bulk density because they respond to changes in pore size distribution rather than pore volume alone. However, current techniques are costly and require time consuming lab analyses. There are some novel methods currently being developed that will cut time and costs somewhat. Alternatively there are pedotransfer functions (ptf’s) such as Rosetta or that published by Hollis et al that estimate hydraulic parameters (ie soil release curve parameters) from texture and BD data. These will not always provide accurate properties estimates, especially in structured soils, but may be useful in estimating the influence of changes in BD on hydraulic functioning.

We will consider this suite of indicators as they are such a powerful measure of the capacity of the soil to regulate water (hydrology) and produce biomass. We will consider the ptf’s as a cheaper proxy for the more expensive lab measurements. These indicators go on through for the data analysis. Data sets identified include the SEISMIC dataset held at Cranfield.

5. Number of erosion features (JR). Although this made it through the logical sieve, the number of erosion features (count/ha?) says nothing of the degree / amount / rate of erosion (for example, number of rills v. gullies). Another candidate SQI – rate of erosion features – is likely to include a survey / count of erosion features in estimating volume of soil loss to calculate rate of erosion (t ha⁻¹ y⁻¹).

6. Rate of erosion: t ha⁻¹ y⁻¹ (JR). This has been identified as a key indicator to bring forward, as it is a defined as a direct measure of soil degradation. It affects the production and regulation function, as any physical loss of soil resource is likely to have impact (more so for the regulation function than the production function, depending on the quality of the baseline soil resource). Much of this has been worked through in a related Defra project (Design of a monitoring scheme to assess the
spatial extent and severity of soil erosion - SP1303; Brazier et al. 2012), with actual values of
erosion loss and a two tiered approach to soil erosion monitoring developed. The latter included
the identification of erosion prone areas, reconnaissance using aerial photographs and the
volumetric measurement of erosion features using three-dimensional imagery techniques. Data
are available linking erosion rates with soil functioning (production; regulation – see Defra Costs
of Degradation Project SP1606; Pimental et al.)

7. Packing Density (LD, PNP). This is the one measure all agreed was the best indicator of soil
compaction (see ENVASSO project; Huber et al., 2008). It is, in essence, a measure of bulk
density modified by texture and thus a better indicator of soil compaction. It is calculated as:

\[ L_d = D_b + 0.009 C \]

where \( L_d \) is the packing density (g cm\(^{-3}\)), \( D_b \) the dry bulk density (g cm\(^{-3}\)) and \( C \) the clay content
(wt.%).

Data holdings were identified as LandIs and some of the ADAS experimental data. Also data from
Huber et al 2008 This indicator will be carried through to the analysis phase.

8. Profile description/visual soil assessment. An on-site systematic, semi-quantified visual
assessment of soil properties using established protocols and visual guides. It is semi-subjective,
but the systematic approach and within field replication removes some of the subjectivity and
variability. Visual evaluation of topsoil structure is relatively quick as it only involves examining the
top 0-20 cm of soil with a spade. Full soil profile descriptions are more time consuming as they
require digging a pit to be effective. There is insufficient quantified data at the national scale to be
able to assess how visual evaluation scores relate to soil function or other indicators of soil
physical quality through modelling and data analysis in WP2, but the method should be carried
forward as a potential low cost field technique to assess general trends in soil condition at the
national scale.

9. Aggregate Stability. Theoretically this is a multi-purpose soil property. It is related to erosion (it is
known as the best estimator of erodibility or soil susceptibility to erosion) and determines the
degree of and susceptibility to surface sealing and capping. As such it affects the production
function (e.g. crop emergence; infiltration capacity; soil moisture content) and the regulation
function (Soil hydrological processes such as infiltration). . Jane to review recent work from the
US where wet aggregate stability was identified as a promising physical SQI. However, like
SMSC, data are lacking in England & Wales, there are no standard assessment procedures and
this is a highly variable soil property in space and time. Jane & Cedric to revisit this property to
see if data are available for further analysis / modelling, but it is unlikely to progress further.

10. Shear Strength. The Going Stick (developed for race courses and football pitches) directly
measures a combination of resistance to penetration and resistance to shearing. No data
available wider than that particular landuse. Other limitation is that it can only measure the top 10
cm of the soil.

11. Soil structure. Part of the visual soil assessment (above) although can also be captured using thin
section CAT scans. We noted that the technique was too costly, labour intensive and as such did
not warrant further work into this particular indicator by itself.

12. Soil texture. This is not expected to change as a result of soil degradation and as such is not a
meaningful SQI (although the role of erosion in changing PSD was discussed, but the degree to
which soil PSD is affected by erosion is event-driven and very difficult to measure accurately) and
acts more as a modifier for other indicators. This property has to be measured to determine
Packing Density.

13. Total porosity. This was also disregarded as it is directly (linear) a function of bulk density and
particle density. BD is captured in PD and particle density does generally not change in soils or
changes linearly with BD, so will not reflect changes in SQ.
14. Sealing. This is where the soil gets covered by building, tarmac or other construction and thus is sealed with significant effects on all soil functions. The expectation is that the degree of surface sealing by infrastructure can relatively easily be determined through remote sensing. It is therefore a sensitive indicator (a step change in soil functioning is likely following sealing) that can be assessed relatively cost effectively. It will also be considered. Toby and Ron to discuss further.

15. Sediment Fingerprinting. A method developed at ADAS by Adrian Collins. There are concerns here about being able to effectively trace from the sediment fingerprint the field or spatial unit from which the soil originated. PNP reported that sediment fingerprinting had been used in the Demonstration Test Catchment projects to quantify sediment source apportionment between different land uses rather than between specific fields. There were also questions regarding the sensitivity and indicator responsiveness, as sediment in a river course is spatially removed from the actual source of soil degradation. Ian to pursue further with Adrian.

16. Eutrophication events/Biological status of rivers. These were deemed too indirect spatially and temporally from soil sources / degradation processes and difficult to relate to the physical quality of soil. Relates more to aims of the Water Framework Directive than the Soil Framework Directive.

17. Electromagnetic induction measurements. This indicator was raised in the proposal but did not make it through the logical sieve. We revisited this and generally found that there were too many confounding factors in EMI measurements to make this an effective SQI. Whilst this technique is able to show variations in soil properties, it cannot pinpoint which properties these are (field calibration is necessary to do this).

In summary (Table A4.1)
Indicators carried through are:
- Depth of Soil
- Soil water retention characteristics
- Packing Density
- Visual Assessment
- Rate of Erosion
- Sealing

Indicators still being considered as ‘maybes’ are:
- Aggregate stability
- Sediment fingerprinting

Datasets identified on which to base the statistical/modelling analysis moving forward are
- LandIs (Cranfield)
- Envasso(EU FP7)
- Soil QC (ADAS)
- Grassland compaction (ADAS)
- Seismic (Cranfield)

Action: ‘Champions’ of each of the physical SQIs listed above will collate the data currently available that are directly relevant to the 8 questions raised at the meeting. Any significant previous work / analyses will also be sourced if available.

The project team will investigate the robustness and responsiveness of soil physical indicators and test these indicators against statistical challenge criteria (sensitivity, discrimination and signal-to-noise ratio) where there are available data to do so. The effectiveness of the proposed indicator, as indicated by the challenge criteria, to monitor and detect change is dependent on four main factors which need to be assessed and quantified:

a) the uncertainty in the measurement of the physical SQI - an understanding of the magnitude of the uncertainty in the measurement of each proposed Indicator (relating to field and laboratory analytical error) - change to be detected in the SQI cannot be smaller than the precision of the method; the variability in the indicator We will use statistical methods (such as power analysis) to
estimate the natural variability of the indicator as given by the observed distributions. The physical SQIs should ideally be valid across all soil and land management regimes or stratified within a defined sub-set of regimes – the ability of the methods to discriminate change will be tested within such data subsets, since these subsets will reduce variability and hence decrease the size of detectable change for a given sample size.

b) the expected rate of change in the indicator – this will be derived from a combination of the WP1 literature review and data analysis in order to estimate the degree of change that can be expected in the indicator in response to a degradation in soil function.

c) the spatial heterogeneity associated with each indicator – this will be assessed at the field scale to determine the intensity of sampling and support at which this sampling takes place in order to ensure that it is effective, yet remains efficient. This will be informed by assessing different sampling configurations (single, bulked, etc).

Where potential indicators have the power to adequately detect change, indicative sampling design costs will be established, where appropriate. The output of WP2 will be a written report at the end of Project Month 9 (mid July 2012), identifying and justifying all salient physical SQIs.