JNCC Report 737

Towards Indicators of Soil Health

Annex 1:

Review and evaluation of existing soil health indicators being used in the UK and internationally

Deeks, L. and Rickson, J.

June 2023

For further information please contact:
Joint Nature Conservation Committee
Quay House
2 East Station Road
Fletton Quays
Peterborough
PE2 8YY
https://jncc.gov.uk/
Communications@jncc.gov.uk

This document should be cited as:
Deeks, L.[1] & Rickson, J.[1] 2023. *Review and evaluation of existing soil health indicators being used in the UK and internationally, JNCC Report 737 Annex 1.* JNCC, Peterborough.
https://hub.jncc.gov.uk/assets/71cece04-eef3-4d34-b118-33ddad50912c#jncc-report-737-annex-1-logical-sieve.pdf

Author affiliations:
[1] Cranfield University

Evidence Quality Assurance:
This document is compliant with JNCC's Evidence Quality Assurance Policy
https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Summary

Soil is a complex system, comprising physical, biological and chemical properties, processes and their interactions. This complexity is compounded by the effects of land use and land management on soil condition and soil functions. The main purpose of the review was to identify and evaluate the strength and weaknesses of existing (and emerging) soil health indicators (SHIs) used in the UK and internationally. To evaluate the current 'state of the art', a collation and critical review of the available evidence was carried out, largely based on the 'grey' and peer-reviewed international literature, with emphasis on soil health indicators used in temperate systems. Many of the findings reflect the outputs from the review of other complex systems (e.g. marine ecology; water) and the development of other environmental indicators (Annex 2).

In the SHI literature reviewed, definitions of soil health/ soil quality are intuitively understood. There is also a general consensus as to the desirable attributes of SHIs. The literature shows a bias towards SHIs most relevant to agricultural land management (i.e. soil properties related to the provisioning function, notably food production). However, SHIs are also needed to reflect the diverse and important non-agricultural functions of soils, including climate-change mitigation and adaptation, water quality and human health.

The SHIs identified in the literature review were evaluated using a 'logical sieve' approach (as described in Defra SP1611: Soil Quality Indicators (physical properties)). The aim was to assess the relevance and suitability of the key SHIs to a broad range of soil types, land uses and different ecosystem goods and services (EGSs). Evidence from the literature review was used to test these SHIs against a series of scientific and technical criteria, including whether the SHIs had clear links to soil functions; their applicability to different land uses; and their performance against a number of 'challenge criteria' (including ease of measurement, sensitivity to change, efficiency and cost, policy relevance, and availability of baseline and threshold data).

This process showed that many popular SHIs have to be converted into meaningful metrics, that can be measured and monitored to reflect changes to soil function and delivery of EGSs, allowing comparisons over space and time. For many soil properties, critical ranges, thresholds and / or scoring curves need to be developed. Also, most novel, emerging soil quality indicators are currently unlikely to be chosen as technological, practical and interpretation related issues need to be overcome before they are ready to be deployed in monitoring programmes. The output of the logical sieve ranked the key SHIs identified in the literature, namely:

- Soil organic matter / soil carbon
- Bulk density
- pH
- Infiltration/ hydraulic conductivity
- Water holding capacity
- Nitrogen
- Soil respiration
- Aggregate stability
- Microbial/ fungal diversity
- Porosity
- Soil structure + aggregate distribution
- Earthworms
- Microbial biomass carbon

The review demonstrated that choosing the best combination of pertinent SHIs is one of the greatest challenges in developing a Soil Health Index, as the index is expected to reflect the full spectrum of soils and their abilities to deliver a range of ecosystems goods and services encompassing food production, carbon storage, water regulation and biodiversity. There is support for identifying reference sites with acknowledged good soil quality for a given function / delivery of ecosystem goods / services. Graphical representation of how well a given soil is able to function (and deliver a particular EGS) is preferred over a single Soil Health Index, as this is much more effective in communicating with stakeholders, soil scientists, target users and the general public. This approach is also preferable for soil scientists, recognising the multifunctionality of soils rather than trying to distil the soil health concept into a single index.

It is important that the process of developing a Soil Health Index is flexible enough to address longer-term goals (e.g. changes in national policy) and/or incorporate additional SHIs (e.g. ones emerging as a result of new innovative methodologies) to allow site or application-specific problem solving and decision making.

Contents

1 Introduction

The main purpose of the review was to identify and evaluate the strength and weaknesses of existing (and emerging) soil health properties and indicators (SHIs) used in the UK and internationally. The methodology used is described in Appendix 1.

Soil properties associated with soil health have been the focus of numerous reviews over the years, for example Merrington *et al.* (2006); Bünemann *et al.* 2018; Chapman *et al.* (2018); Rinot *et al.* (2019); Dui and Stewart (2020) and Lehmann *et al.* (2020). The main conclusion from these reviews is that soil is a complex system, comprising physical, biological, and chemical properties and processes (and their interactions). With soil having both inherent (e.g. soil texture, type of clay, depth to bedrock and drainage class) and dynamic qualities such as variable infiltration rates (USDA 2006). Inherent, more permanent, soil qualities determine the suitability of soils for specific uses (e.g. well drained sandy soils for horticulture), and dynamic soil quality is dependent on how soil changes in response to the way in which it is managed (e.g. bulk density of soils compacted by heavy machinery) (De la Rosa & Sobral 2008). This complexity is compounded by the effects of land management practices and land uses. Soil types respond significantly differently to soil management and also in how they react to climate change, so in combination “different soils in different climate zones will offer different challenges and opportunities to be met by appropriate management” (Bonfante *et al.* 2020).

According to Laishram *et al.* (2012) and Sharman (2017), while definitions of soil health/quality are intuitively understood and the terms often used interchangeably (including by the USDA (2019)), an agreed applicable definition and an agreed method for assessment have not yet been reached by the soil science community globally. Sharman (2017) suggests this is because of:

- Variation in the goals of land/ecosystem management and in the audiences to whom definitions of SHIs may be applicable or appropriate.
- Few scientifically validated methods exist for measuring soils’ integrated 'system' properties.
- Multiple possible methods for 'reductionist' soil health measurements are available.
- Variety in tools and data applied.
- Environmental variability over a range of geographical and temporal scales.
- There is scant understanding of the capacity of soil to function under stress and disturbance.

Inevitably, in a broad review of the international soil health literature, different terminologies are used by different authors to describe individual and combined soil properties linked to soil health (e.g. ‘metric’, ‘property’, ‘attribute’, ‘measure’, ‘index/indices’, ‘indicator’, etc.). Where appropriate, these terms have been standardised in this Annex to reflect the definitions given in the Main Report’s Glossary. For example, in this review, ‘soil health indicator’ is used to refer to a combination / suite of individual soil health properties, rather than the individual properties themselves.

2 Key soil health properties

There have been numerous approaches to the selection of relevant attributes used to describe soil health. Table 1 lists the desirable attributes of properties defined by a range of authors (see Appendix 2). Measurements for SHIs can be quantitative (i.e. numerical attributes) or qualitative (usually categorical attributes). This has implications for the robustness of any statistics that can be applied to the SHIs in showing 'meaningful' differences in values and the associated change in soil functions / delivery of ecosystem goods and services.

However, the list of potential properties can be long. For example, Merrington *et al.* (2006) identified 67 potential properties for measurement, but, with justification, limited this to six chemical, five biological and nine physical, with four 'proxy' variables as those most relevant to the soil's interaction with the environment. The review by Bünemann *et al.* (2018), identified 27 properties from a review of 65 soil quality assessment approaches (having removed repeated mentions of the same set of properties from the same authors). Figure 1 shows the frequency with which each of these 27 properties was mentioned. The most frequently mentioned measurable properties were:

- Total organic matter/carbon (mentioned in 91% of assessments.
- pH (mentioned in 82% of assessments).
- Available P (mentioned in 74% of assessments).
- Water storage (mentioned in 60% of assessments).
- Bulk density (mentioned in 54% assessments).

The first wholly biological property, soil respiration (mentioned in 29% of assessments), was ranked 14th.

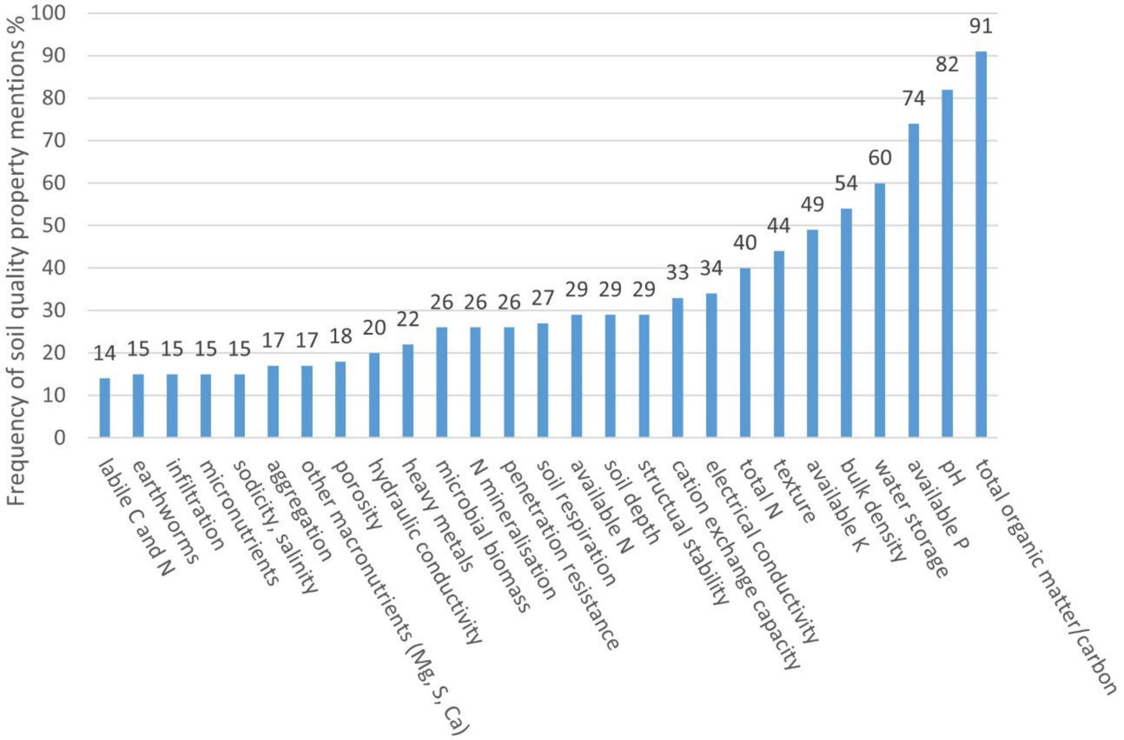

Figure 1. Frequency of soil quality properties mentioned in 62 publications reviewed by Bünemann *et al.* (2018). This represented 65 minimum datasets. Where the same author proposed the same set of properties in more than one publication, these were only considered once (adapted from Bünemann *et al.* 2018).

In the review by Lehmann *et al.* (2020), greater than 75% of 65 soil health testing frameworks included soil organic matter, pH and plant-available phosphorus and potassium; greater than 50% water storage and bulk density; and 33% recommended measurements of soil respiration, microbial biomass, or nitrogen mineralization to characterise soil biological properties and structural stability. Chemical measurements made up at least 40% of the soil quality properties in 90% of the assessment schemes. Biological properties typically still constituted fewer than 20% of the soil health properties mentioned. This is even though others argue that biological properties are sensitive indicators of soil health (De la Rosa & Sobral 2008). It may be that soil biological tests to determine and define soil health are still in their infancy (Wood & Litterick 2017).

Lehmann *et al.* (2020) recommended that soil aggregation, infiltration, earthworm abundance and organic C and N fractions should be more widely adopted in soil health testing. They also recognised a bias in soil properties measured for soil health towards those most relevant to agricultural land management. They recommended new measurements that are mainly geared towards non-agricultural services delivered by soil, such as human health and water quality, need to become part of routine soil-health testing. These included pathogens, parasites, biodiversity, bioavailable and mobile toxins (such as dioxin, polycyclic aromatic hydrocarbons and microplastics) and compounds, and pore-size diversity. Those authors also suggest the need to develop soil health indicators reporting on the functions of soils that relate to climate-change mitigation and adaptation, such as greenhouse gas emissions and carbon sequestration, which they conclude have largely been ignored. They also suggest that diversity measures, whether organismal (biological), molecular (chemical) or structural (physical), are not adequately included in, or integrated into, analytical frameworks of soil health.

Norris *et al.* (2020) chose 28 properties in an assessment across 120 long-term (greater than 10 yrs.) agricultural research sites, spanning sites from north-central Canada to southern Mexico. This wide geographical distribution of sites was in recognition that healthy soils do not represent identical capacities to function across a landscape. The distribution of sites gave a wide range of soils, climates, and land management systems to compare sensitivity of changes against.

Table 1. Desirable attributes for soil quality / soil health properties and indicators.

Attribute type	Desirable attributes of Soil Quality / Soil Health Properties and Indicators	Examples from the literature review
Interpretation of results	Direct links with soil functions and the delivery of ecosystem goods and services	Lehmann *et al.* (2020); Rinot *et al.* (2019); Stott (2019); Bünemann *et al.* (2018); Parisi *et al.* (2005); Nortcliff (2002)
	Measured values are "directionally understood" (e.g. more is better, less is better)	Stott (2019)
	Outcome based thresholds are known (e.g. yield, resilience, risk, environmental)	Stott (2019)
	Clear (absolute) interpretation for a given property	Bünemann *et al.* (2018)
Sensitivity	Able to change detectably and quickly without being reflective of merely short-term oscillations	Lehmann *et al.* (2020); Stott (2019); Dumanski & Pieri (2000)

Attribute type	Desirable attributes of Soil Quality / Soil Health Properties and Indicators	Examples from the literature review
Sensitivity (continued)	Able to detect change in a relevant time scale for decision making	Rinot *et al.* (2019); Stott (2019); Sharman (2017); Parisi *et al.* (2005); Karlen *et al.* (1997)
	Able to reflect variations in soil management, land use, the environment, climate change	Rinot *et al.* (2019); D'Hose *et al.* (2014); Stott (2019); Bünemann *et al.* (2018); Sharman (2017); Parisi *et al.* (2005); Nortcliff (2002); Karlen *et al.* (1997)
Practicality	Practical / ease of sampling, and submission and preparation of samples for analysis	Lehmann *et al.* (2020); Stott (2019); Bünemann *et al.* (2018); Parisi *et al.* (2005); Karlen *et al.* (1997)
	Cheap to measure / analyse (labour and supplies; specialised equipment costs; lab space and time, including overheads)	Lehmann *et al.* (2020); Stott (2019)
	Short turn around for analysis	Lehmann *et al.* (2020); Karlen *et al.* (1997)
	Informative for soil and land management	Lehmann *et al.* (2020)
	Scientifically based data (Physical, biological, and chemical soil properties)	Rinot *et al.* (2019)
	Cost-effective	Rinot *et al.* (2019); Bünemann *et al.* (2018)
	Can be analysed in commercial laboratories	Stott (2019)
	Accurate, precise, and reliable measurements	Rinot *et al.* (2019); Stott (2019); Bünemann *et al.* (2018); Sharman (2017); Nortcliff (2002)
	Methodology clearly understood	Sharman (2017)
	Comparable across space and time	Sharman (2017); Nortcliff (2002)
Other	Target values or potential ranges of values required	Rinot *et al.* (2019); Stott (2019); Laishram *et al.* (2012)
	Avoid duplication i.e. using different attributes that reflect the same outcome	Rinot *et al.* (2019)
	May be combined with other soil properties to represent specific functions / conditions	Stott (2019)
	Comparable to data from other sampling campaigns	Bünemann *et al.* (2018); Nortcliff (2002)

Attribute type	Desirable attributes of Soil Quality / Soil Health Properties and Indicators	Examples from the literature review
Other (continued)	Can be aggregated across different spatial scales (local, regional, national)	Sharman (2017)
	Free, globally available data	Sharman (2017)
	Can substitute better information, where available	Sharman (2017)
	Skills required in use and interpretation	Nortcliff (2002)

A review by Chapman *et al.* (2018), of 240 papers, identified a major gap in knowledge of soil health measurements, that they primarily deal with soil depths of 30 cm or less. This suggests that there is limited knowledge in how soil properties and soil functions at depth respond to changes in land management. In addition, 81% (n = 187) of the studies they looked at were carried out at the field scale, which raises the question of how to measure public goods delivery at the landscape scale. Chapman *et al.* (2018) also observed other limitations, including that 57% of the studies reviewed only considered short-term change in soil health properties, with only 19% of studies considering longer term (greater than 20 years) impact.

The interpretation of the values of proposed soil quality/health variables needs to be well-defined (Bünemann *et al.* 2018). Generalization about critical limits are difficult as the critical limit of an individual soil property (or group of properties) can be ameliorated or exacerbated by limits of other soil properties and the interactions among other soil quality measurements (Arshad & Martin 2002; De la Rosa & Sobral 2008).

One approach to assessing soil health is to compare soils that have been under a certain use and land management system to natural soils that have not been disturbed (De la Rosa & Sobral 2008). With this method the influence of climate (precipitation and temperature), geomorphology and weathering rate can be accounted for by comparing soils within an agri-ecosystem or soil type. However, this method depends on the availability of a natural reference soil.

Gomez *et al.* (1996) considered an attribute to be at a sustainable level if it exceeds a designated trigger or threshold level where thresholds are tentatively set, based on the average local conditions.

Another approach is to evaluate the advantages and limitations of individual candidate soil quality measurements. Table 2 presents the arguments to both include and exclude specific properties in any given chosen approach.

A list of assessments that integrate a number of soil properties is then provided in Table 3, followed by a list of interactions between soil properties in Table 4.

Table 2. Examples of soil health / quality measurable properties with their specific advantages and limitations (justifying their inclusion / exclusion from soil health measurement and monitoring).

Indicator type	Soil property	Range	Advantages	Limitations
Chemical indicators	**Soil pH**	6.5 to 7.0 most crops but not all (Doll 1964); varies depending on the ecosystem service in question (Merrington *et al.* 2006); 6.5 to 7.5 (Griffiths *et al.* 2018 score card values); varies depending on the plant species (CASH, Moebius-Clune *et al.* 2016)		Biased towards the (agricultural) production function of soil (e.g. calcareous grassland habitat restoration might favour more alkaline soil pH).
	Soil carbon	Carbon in cultivated soil is ~20% less than in uncultivated soils. 2% Soil Organic Carbon (SOC) is the minimum requirement for the maintenance of satisfactory soil aggregate stability, above which no further increases in productivity are achieved (Janzen *et al.* 1992). The quantitative basis for such thresholds is limited (Loveland & Webb 2003). 6 g kg^{-1} SOC is the minimum limit to prevent collapse of soil structure of sandy loams (Janssen & de Willigen 2006). In India, soils with an organic carbon percentage of less than 0.5 are considered low fertility soils, 0.5 to 0.75 are medium fertility soils and greater than 0.75 are high fertility soils (Prasad *et al.* 2003). Crop yield increases by 12% for every 1% of Soil Organic Matter (SOM) in the USA (Magdoff 1998).	Considered as the best surrogate for soil health (Dumanski & Pieri 2000; Magdoff & Van Es 2021). SOC is considered to be a fundamental soil property, determining soil fertility, soil structural stability, retention of organics, sequestration of C and nutrient cycling (Merrington *et al.* 2006). According to Lehmann *et al.* (2020), soil carbon fractions of both unprotected and mineral-protected organic matter allow assessment of soil organic matter vulnerability with respect to soil carbon sequestration, and are indispensable indicators for soils' climate-change function. Sensitive to change.	The magnitude of decline in soil carbon depends on the soil depth used for carbon estimations and time scale of land use change (Ackerman 1993). There is little evidence to show that organic matter contributes to yield on irrigated and fertilized croplands (Sojka & Upchurch 1999). There has been no consensus on what the critical level of soil organic matter should be in an agricultural soil and how this level will vary between soils of different textural classes under different environmental conditions (Nortcliff 2002), although evidence presented by Griffiths *et al.* (2018) and Merrington *et al.* (2006) indicates this is now changing.

Indicator type	Soil property	Range	Advantages	Limitations
Chemical indicators (continued)	**Soil Carbon (continued)**	An index derived from both labile and non-labile carbon fractions is likely to be a more sensitive indicator of land use intensification or land management practices compared with a single measure of soil carbon content (Breland & Eltun 1999). An organic carbon value of 1.5% is considered a lower limit for soils with 40% clay but would be considered high in soils that have less than 10% clay (Verheijen *et al.* 2005). See also the score card values for soil organic matter (by loss on ignition, %) in Griffiths *et al.* (2018); CASH (Moebius-Clune *et al.* 2016) and Magdoff and Van Es (2021) report on handheld colour charts. Darkest soils, almost black, indicate 3.5 to 7% organic matter. A dark brown soil indicates 2 to 3%, and a yellowish brown soil indicates 1.5 to 2.5% organic matter. (Colour may not be as clearly related to organic matter in all regions because the amount of clay and the types of minerals also influence soil colour.) Also report that recently smartphone apps have also been developed to estimate soil colour vs. organic matter content.	See above	Improvements to aggregation increase the risk of loss of agrochemicals to the wider environment (Stevenson 1972; Ross & Lembi 1985; Sojka & Upchurch 1999). As a change in land use is coupled with change in bulk density, the method of calculation of soil carbon is also likely to influence the conclusion on the land use change-carbon stock relationship. An alternative method is to measure bulk density first and then to calculate the sampling depths to obtain the same mass (dry soil) of soil in different land uses (Ellert & Gregorisch 1996). Three distinct types of approaches could be adopted to quantify the change: (i) repeated measurements on a single site (ii) paired sites and (iii) chronosequences where sites are neighbouring. As labile fractions respond to seasonal variations more than total soil organic carbon (Bastida *et al.* 2006), sampling season needs to be carefully considered while using labile organic carbon as an indicator of soil quality.

Indicator type	Soil property	Range	Advantages	Limitations
Chemical indicators (continued)	**Soil carbon (continued)**	Soil texture affects organic matter content. Magdoff and Van Es (2021), suggest sands may be less than 1%; loams (clay loam, UK texture) may have 2% to 3% and clays from 4% to more than 5%. A soil with 20% silt and clay, for example, can store a maximum of 3.6% organic matter, while a soil with 80% silt and clay can hold 6.1% organic matter. This does not include the additional particulate organic matter that may be either subject to rapid decomposition (active) or protected from decomposition by soil organisms inside small (micro) aggregates (part of the passive organic matter). However, the clay content and type of clays present influence the amount of organic matter particles "stored" inside micro-aggregates. See also Merrington *et al.* (2006).	See above	Molecular and soil structural diversity are not yet explored but are important for soil organic carbon persistence and seques-tration (Lehmann *et al.* 2020). Total organic carbon, satisfies three criteria (informative / interpretational, effective / practical, relevant / conceptual), but typically does not change very quickly (is not sensitive). Additional indicators, such as organic carbon fractions, that are more sensitive are therefore also requireed (Lehmann *et al.* 2020) SOC may change slowly, over 3 to 5 years in sub-humid temperate climates and slower under drier conditions (Stott 2019). Organic matter provides a gross indication of soil biological functions at best (Haney *et al.* 2018).
	Labile carbon fractions	See CASH (Moebius-Clune *et al.* 2016).	Labile carbon fractions are highly correlated with soil microbial biomass and the availability of labile nutrients such as nitrogen, phosphorus and sulphur. They vary over shorter time scales and may be sensitive to change in climate and land management (Norris *et al.* 2020).	

Indicator type	Soil property	Range	Advantages	Limitations
Chemical indicators (continued)	**Extractable P**	See the Griffiths *et al.* (2018) score card values, CASH (Moebius-Clune *et al.* 2016) and Merrington *et al.* (2006). P > 5 mg/kg could indicate species-poor, nutrient-enriched pastures (Goodwin *et al.* 1998).	Based on agricultural requirements for P (Merrington *et al.* 2006)	
	Extractable K	See Griffiths *et al.* (2018) score card values and CASH (Moebius-Clune *et al.* 2016).		
	Extractable Mg	See Griffiths *et al.* (2018) score card values and CASH (Moebius-Clune *et al.* 2016).		
	Saline and sodic	See CASH (Moebius-Clune *et al.* 2016).		
	Total soil nitrogen	See Merrington *et al.* (2006).		
	C/N ratio	See Merrington *et al.* (2006).		
Biological indicators	**Soil microbial properties:** **Phospholipid fatty acid profile (PLFA)** **Ester-linked fatty acid methyl ester profile (EL-FAME)**	See Griffiths *et al.* (2018) score card values for soil microbial biomass.	Microbial biomass is correlated with anaerobically mineralized C (Hart *et al.* 1986, Stockdale & Rees 1994). Microbial biomass is considered as an integrative indicator of microbial significance of soils (Powlson 1994). Biological diversity in particular has been recognised as important for soil and human health Bünemann *et al.* 2018). Among different soil quality parameters, biological parameters are more responsive to changes in land management practices (Amat *et al.* 2021).	There is a huge degree of spatio-temporal variation within a given land use / ecosystem observed (Parkin 1993; Khan & Nortcliff 1982). There are issues relating to the ephemeral nature of biological measurements and difficulty in justifying target ranges (e.g. New Zealand; Sparling *et al.* 2004). Variation in soil microbial biomass may not be necessarily correlated to soil quality (Martens 1995; Dilly & Munch 1998).

Indicator type	Soil property	Range	Advantages	Limitations
Biological indicators (continued)	**Soil microbial properties:** **Phospholipid fatty acid profile (PLFA)** **Ester-linked fatty acid methyl ester profile (EL-FAME) (continued)**	See above	PLFA gives a measurement of total microbial biomass, a broad categorization of bacterial community, and selects active microbial communities. It is also sensitive to identifying differences across a variety of ecosystem types (Norris *et al.* 2020). The PLFA technique has been successful in differentiating land uses across bio-geographical regions (Norris *et al.* 2020). In forest ecosystems, changes in soil health due to compaction and organic matter removal resulted in significantly different soil microbial community structure (Hartmann *et al.* 2012) and the community's potential ability to decompose organic matter (Cardenas *et al.* 2015).	There is a lack of mechanistic understanding of how soil biota relate to soil functions (meeting the \`relevant' criteria). There is a lack of understanding about how soil microbial properties relate to land management decisions (\`informative') and an inability to easily quantify biological indicators (\`effective') (Lehmann *et al.* 2020). Appropriate soil-health indicators and practical quantification methods for soil biota diversity are lacking (Bünemann *et al.* 2018). There is a high cost of analysis. Thresholds or ranges to define relatively poor/good functioning are currently not known, yet could be established if standard procedures are followed (Stott 2019). PLFA has been shown to be both sensitive and insensitive to long-term agricultural management practices, possibly an effect of fertilizer use on microbial communities, which requires further investigation (Norris *et al.* 2020).

Indicator type	Soil property	Range	Advantages	Limitations
Biological indicators (continued)	**Enzymes**	Phosphomonoesterase, chitinase and phenol oxidase, as a group reflect relative importance of bacteria and fungi, as well as the nature of organic matter complex (Giai & Boerner 2007).	Phosphomonoesterase, correlated with microbial biomass (Clarholm 1993; Kandeler & Eder 1993), fungal hyphal length (Haussling & Marschner 1989) and nitrogen mineralization (Decker *et al.* 1999). Enzyme activities have been found to be very responsive to both natural and land management changes (De la Rosa & Sobral 2008). Enzyme assays quantify the potential for reactions in soil that are intimately associated with elemental cycling, e.g. β-glucosidase, N-acetyl-β-D-glucosaminidase, phosphatase and aryl sulfatase are representative of C, N, P and S cycling respectively (Norris *et al.* 2020).	With greater than 500 enzymes, it is difficult to decide which would be the best indicators of soil quality (Scholoter *et al.* 2003). Soil enzyme assays generally provide a measure of the potential activity which is rarely expressed. Microbial population methods are relatively novel and do not have a long-established history with known trends and thresholds (Stott 2019).
	Soil microbial carbon: total organic carbon ratio	Site-specific baseline value for different soil systems (Anderson 2003).	Soil microbial carbon can be linked to physical, biological, and chemical soil properties and processes.	Requires specialist equipment for analysis. Soil microbial carbon may not reflect level of activity / microbial function in the soil.
	Labile fractions, microbial biomass, dehydrogenase activity and ATP levels		May be highly correlated (Nannipieri *et al.* 1990; Garcia *et al.* 1994).	

Indicator type	Soil property	Range	Advantages	Limitations
Biological indicators (continued)	**Faunal groups (e.g. earthworms; nematodes)**	Scoring (after Shepherd 2000): • Earthworms plentiful (score = 2) if greater than 8 earthworms counted; • Moderate earthworm numbers (score = 1) if 4 to 8 earthworms counted; • Few or no earthworms present (score = 0) if less than 4 earthworms counted.	Examples such as nematodes are likely to be an effective indicator of soil quality if they are dominant and occur in all soil types, have high abundance and high biodiversity, and play an important role in soil functioning (e.g. in food webs). Earthworms respond to soil management practices (e.g. decreased abundance under conventional tillage (ploughed) compared with less intense cultivation (Briones & Schmidt 2017).	Based on a 6-year trial of soil quality monitoring in New Zealand, Sparling *et al.* (2004) did not find utility in using earthworms as a measure of soil quality because of the difficulty in the ephemeral nature of such biological measurements. Difficulties in classification of organisms at the species level have a major constraint delimiting use of indicators based on soil organisms, more so the microfauna. A collembola expert can be expected to analyse 5 samples a day and a nematode expert two samples a day (Ekscmitt *et al.* 2003). There is a poor correlation between soil animal community and environmental factors. This could be due to a) the significant influence of autogenous dynamics of the population under consideration, b) the interaction of this population with predators, parasites and competitors, and c) the (presently) indiscernible past conditions (Salt & Hollick 1946). Earthworms will be abundant In the surface 6 to 9 inch layer when it is moist but tend to go deeper into the soil during dry periods (Magdoff & Van Es 2021).

Indicator type	Soil property	Range	Advantages	Limitations
Biological indicators (continued)	**Earthworms**	See the Griffiths *et al.* (2018) score card values.		
	Nematodes and micro-arthropods	See the Griffiths *et al.* (2018) score card values.		
	Indicators of carbon food source: **Per-manganate oxidizable C** **Particulate organic matter** **28-day C minerali-sation** **Water extractable organic C** **Soluble carbohydrates** **Substrate-induced respiration** **Microbial biomass C**		The pool of soil C is a small proportion of SOM, the larger the pool, the larger the microbial population than can be supported (Stott 2019).	Each method measures a slightly different pool of SOC (Stott 2019).

Indicator type	Soil property	Range	Advantages	Limitations
Biological indicators (continued)	**Fungal**		Species level identification is possible, answering “what is there”.	Identification neither implies the microorganisms are alive and active (Blagodatskaya & Kuzyakov 2013) nor does it describe their function (Prosser 2015). Frac *et al.* (2018) suggest that to unravel the function of the community, either (shotgun) metagenomics (Uroz *et al.* 2013; Hannula & van Veen 2016; Castaneda & Barbosa 2017), metatranscriptomics (Damon *et al.* 2012; Turner *et al.* 2013; Hesse *et al.* 2015) or time-intensive culture based methods combined with functional tests must be used (Behnke-Borowczyk *et al.* 2012; Galqzka *et al.* 2017; Wolifiska *et al.* 2017). A fast, but coarse alternative to molecular methods is MicroResp (Creamer *et al.* 2016) or a community level physiological profiles (CLPP) approach (Fraic *et al.* 2017), which gives an indication on substrate use of the total microbial community. Being able to both identify species and characterise their role in the environment is important. The ability to compare functional structures between ecosystems and predict responses to environmental changes and interventions would be a useful advance (Frac *et al.* 2018).

Indicator type	Soil property	Range	Advantages	Limitations
Biological indicators (continued)	**Soil protein**	See CASH (Moebius-Clune *et al.* 2016).	Is a measure of the fraction of the soil organic matter which contains much of the organically bound N. Microbial activity can mineralise this N and make it available for plant uptake.	
	Respiration	See CASH (Moebius-Clune *et al.* 2016).	A measure of the metabolic activity of the soil microbial community.	Does not say anything about diversity of populations
	Active carbon	See CASH (Moebius-Clune *et al.* 2016).	A measure of the small portion of the organic matter that can serve as an easily available food source for soil microbes, thus helping fuel and maintain a healthy soil food web.	
	Potentially mineralizable N	See CASH (Moebius-Clune *et al.* 2016).	A combined measure of soil biological activity and substrate available to mineralise nitrogen to make it available to the plant.	
	Root health bioassay rating	See CASH (Moebius-Clune *et al.* 2016).		
Physical indicators	**Bulk density / Porosity**	Soil macroporosity below 10% (v/v) is reported to decrease pasture production, but whether this threshold is true for other land uses is not known in New Zealand (Sparling *et al.* 2004). See Table 3 for critical values for agricultural crops. See the Griffiths *et al.* (2018) score card values and Merrington *et al.* (2006).		Unintended 'benefits' of increased porosity: Increased risk of pollutant transfer
	Texture			Does not readily change, cannot be easily managed (Lehmann *et al.* 2020).

Indicator type	Soil property	Range	Advantages	Limitations
Physical indicators (continued)	**Water content**	See CASH (Moebius-Clune *et al.* 2016).		Largely depends on the immediately preceding weather (Griffiths *et al.* 2018).
	Penetro-meter resistance	See the Griffiths *et al.* (2018) score card values and CASH (Moebius-Clune *et al.* 2016).		
	Aggregate stability	See CASH (Moebius-Clune *et al.* 2016) and the water stable aggregates in Merrington *et al.* (2006).	Aggregates have been shown to be sensitive to soil management, specifically macroagregation is highly influenced by management (Stott 2019). Soil macroaggregate stability is related to soil biology, with microorganisms producing glues (soil carbohydrates) that along with fungal hyphae and fine roots bind the aggregates together. Since aggregation integrates soil biological, chemical and physical properties, it is an important indicator of SH (Stott 2019) differences between treatments can be detected within 1 to 3 years although minor changes in land management may take longer (Stott 2019).	Does not adequately describe the soil habitat or soil architecture experienced by microbes or roots and is not useful to model water transport (Young *et al.* 2001). Slight variations of the standard method used by various laboratories makes comparison of datasets challenging (Stott 2019). The trends are known, however the upper limit of how much water-stable aggregation at the surface is enough or achievable for a healthy soil is known for only a few regions (Stott 2019).
	Infiltration		Rapid infiltration indicates a good soil structure in clay or silt soil textures but is undesirable in sandy textures.	Quick infiltration can increase risk of off-site contamination. Field measurements of infiltration are time consuming and challenging to make as they may require many tens of litres of water (Griffiths *et al.* 2018).

Table 3. Assessments that integrate a number of soil properties.

Assessment type	Properties included	Ranges
Ecomorphological score	See QBS Index	As a general rule, eu-edaphic (i.e. deep soil-living) forms get a score of 20; epi-edaphic forms (surface living forms) are scored as 1. Groups like Protura and Diplura have a single value of 20, because all species belonging to these groups show a similar level of adaptation to soil (Parisi *et al.* 2005)
Visual Soil Assessment	The SRUC VESS (arable) and Healthy Grassland Soil methodology (Griffiths *et al.* 2018). Rated by growers as requiring minimum skill, not too time consuming to do and inexpensive (AHDB, Great Soils project) A range of visual soil assessments exist SOIL.pak (McKenzie 2001), Profil cultural (Roger-Estrade *et al.* 2004), VS-Fast (McGarry 2006), Peerlkamp (Ball *et al.* 2007), VSA (Shepherd *et al.* 2008), VESS (Guimaraes *et al.* 2011) and M-SQR (Mueller *et al.* 2014), and between them they include indicators relating to: **Soil physical properties** • Soil texture • Soil structure • Soil consistence • Aggregate size distribution • Aggregate shape • Slaking/dispersion • Soil porosity • Soil colour • Soil mottles • Available water • Water infiltration	See Griffiths *et al.* (2018) score card.

Assessment type	Properties included	Ranges
Visual Soil Assessment (continued)	**Soil chemical properties** • Soil pH • Labile organic C **Soil biological properties** • Earthworms (no., size) • Potential rooting depth • Root development As well as collecting other useful observations including: • Soil layers • Surface crusting • Surface cover • Surface ponding • Slope • Soil erosion With the VSA method (Shepherd *et al.* 2008) specifically being designed to quantify soil health.	See above
Land quality		(i) the yield gap indicator, which is a measure of the difference between yields under optimum and actual management conditions (ii) soil nutrient balance indicator (Bindraban *et al.* 2000)

Table 4. Interactions between soil properties.

Soil properties	Interactions
pH	Affects nutrient availability Changes how nutrients interact with other constituents of soil CEC and base saturation are pH dependent measurements
Sodium adsorption ratio	Clay dispersion (high SAR increased dispersion) Infiltration (high SAR reduced infiltration) Aggregate stability (high SAR diminished stability
Nutrient analysis of P	Soil mineralogy pH Plant uptake Phosphate=solubilizing microorganisms
C	Climate Soil management Soil texture (inherent soil property) Topography (inherent soil property) Water holding capacity (increased C increases WHC) Infiltration (increased C increases infiltration) Total C linked to total N
Soil microbial diversity	Monocultures Agricultural intensity
Soil microbial biomass	Monocultures Agricultural intensity

Stott (2019) reiterates that standardisation of pre-analytical soil processing (e.g. degree of aggregation, sieving, or grinding), is as important as the analytical methods for understanding results.

Sharman (2017) collated evidence on soil health to devise a star rating system for attributes to indicate soil health, which were ‘scored’ from: * (least accuracy and efficacy) to ***** (greatest accuracy and efficacy). From their review, Sharman (2017) identified a number of attributes that were frequently included in a SHI “minimum data set” (MDS). Other attributes that appeared less frequently in a MDS were labelled ‘optional’. They also distinguished ‘proxy’ variables which were indirect measurements of a soil property (Table 5). Whilst the review is extensive, some parameters that could be candidate indicators, such as soil aggregation or aggregate stability, were not considered as SHIs in this study.

Table 5. Soil properties identified based on a collation of evidence from Sharman (2017).

Theme	Type	Soil Health Indicator	Measurement method	Consensus on value from Sharman's (2017) review
Physical	MDS	Soil texture/structure	Visual Soil Assessment (VSA)	*****A wider soil health test that assesses more than just structure
			Visual Evaluation of Soil Structure (VESS)	***Can give good 'traffic light' soil structure information
	MDS	Water-holding capacity	Field sampling and lab assessment	Unknown
	MDS	Soil rooting potential	Cone penetrometer	**Could be used to 'traffic light' soil structure
	MDS	Infiltration and soil bulk density	Field assessment Remote sensing for bulk density	*****Well-recognised, comparable metric
	Optional	Erosion/waterlogging	Range of scales from manual to satellite imagery	Unknown Requires ground truthing
	Proxy	Soil organic matter (SOM), grain nutrient content and yield in combination	In-field or post-harvest assessment or potentially in-field remote sensing	Unknown Requires ground truthing
Chemical	MDS	Soil organic matter (SOM)	Field sampling and lab. Although remote sensing starting to become available	*****Tests at the field level are generally the strongest approach Satellite sensing would lessen field accuracy and reduce star rating
	MDS	pH	Field and lab	Is not a strong measure of soil health/quality in its own right, but provides useful background and context

Theme	Type	Soil Health Indicator	Measurement method	Consensus on value from Sharman's (2017) review
Chemical (continued)	MDS	Electrical conductivity	Field or remote sensing	**Could be used to 'traffic light' soil structure
	MDS	Extractable nitrogen	Field/lab assessment or potentially in-field remote sensing	***Can be informative but is also variable and difficult to interpret depending on timing
			Grain nutrients	****
	MDS	Phosphorus	Soil phosphorus (P) field/lab	*****Some interpretation of total and available nutrients is needed as well
			Grain nutrients	****Could be informative, but is also variable and would need ground truthing
	Optional	Potassium, other nutrients Macro-nutrients: potassium (K), magnesium (Mg), micro-nutrients: iron (Fe), copper (Cu), boron (B), manganese (Mn), etc.	Field/lab	**Difficult to interpret and value varies with crop and scenario
	Optional	Cation-exchange capacity	Field/lab	Unknown
	Optional	Salinity	Range of scales; often field based but remote sensing is becoming more available	**Can be of value in certain scenarios
	Optional	Detailed soil chemistry	In-field testing or lab	*Can be difficult to interpret;
	Proxy	SOM, grain nutrient content and yield in combination	Would need data sets and ground truthing	Unknown

Theme	Type	Soil Health Indicator	Measurement method	Consensus on value from Sharman's (2017) review
Chemical (continued)	Proxy	Other options for some crops (e.g. nutrient budgets or visual deficiency symptoms)	Would need data sets and ground truthing	Unknown
Biological	MDS	Soil Organic Matter (SOM)	Field and lab. Remote sensing	*****Remote-sensing techniques have great potential for mapping and temporal—spatial monitoring of SOM
	MDS	Soil biology: macro- and microbiological properties	Soil microbial assessment/ profiling field/lab	*Typically require laboratory access and can be difficult to interpret
			Number and diversity of macro- and micro-organisms	****Earthworms are a relatively well-recognised and accepted metric
			Number and diversity of mycorrhiza and root colonisation, field/lab	*Typically require laboratory access and can be difficult to interpret
	MDS	Potentially mineral sable nitrogen	Field/lab	Can provide an estimate of biomass, although other soil factors influence
	MDS	Soil respiration	Field/lab	**Some approaches can be a useful laboratory or field indicator, but they are often difficult to interpret
	Proxy	Crop growth or other physical changes	Field to satellite	*Will need ground truthing and would be insensitive

3 Soil health indices / indexes / indicators and approaches

Choosing the best combination of soil properties to devise an overall ‘Soil Health Indicator’ is challenging, as the indicator is expected to reflect soils’ ability to deliver a range of ecosystems services from crop production to habitat protection. Rinot *et al.* (2019) suggest the following issues with the use of an SH Indicator:

- Focusing on a specific soil threat, function or ecosystem service cannot capture the overall status of the soil.
- Scoring approaches used to assign SH indicators are not consistent and could be manipulated, based on expert (or otherwise) opinions or inappropriate statistical methods.
- A transparent, repeatable methodology to assess and subsequently manage soils is needed.
- An indicator tailored to a specific soil type may fail to project SH status, conditions, and dynamic changes at different scales (regionally and globally).
- An SH indicator needs to account for soil heterogeneity in space and time.
- SH needs to be assessed over a large scale to reach constructive decisions for sustainable land management.
- Interpolation is essential using high-resolution mapping via, for example, remote sensing or advanced geostatistical methods.
- There is a need for standard soil sampling schemes and analysis procedures.
- Models have limitations related to selected attributes, algorithms and assumptions used.
- The expert-based approach assumes that a given expert understands the complexity of the mechanisms studied and that their knowledge can be translated accurately into the model.

Some authors are pessimistic that this is even possible. Sojka and Upchurch (1999) suggest that the search for a single, affordable, workable soil quality indicator is unattainable. Others including Ewing and Singer (2012) have suggested that a multi-property soil quality indicator is required to classify the full spectrum of soils. Karlen and Stott (1994) state that a soil quality indicator can be seen as a decision tool that helps to effectively combine a variety of information for multi-objective decision making. Many soil quality / soil health indicators have been suggested (see Section 2).

Norris *et al.* (2020) evaluated established and newer soil health indicators and soil health evaluation programs, including three soil health indicators:

1) the Soil Health Management Assessment Framework;
2) Haney Soil Test; and
3) Cornell’s Comprehensive Assessment of Soil Health.

These three indicators are of interest because while data for these indicators were found to be easy to collect, analyse and interpret for farmer practitioners, when applied to research sites, the methods did not consistently capture improvements in soil health. This raised questions relating to application, ease-of-use and scope of metrics used (Norris *et al.* 2020). To understand the reasons behind this, Norris *et al.* (2020) undertook a large-scale broad assessment of soil health indicating properties, including those used in the three indicators, across a wide range of soils, climates, and land management systems. Their assessment criteria were for each property to be:

1) Applied regionally and, taking soil inherent properties into account, applied across the continent.
2) Have a clear range of responses based on agricultural goals.
3) Be responsive to varying land management practices.

However, in the final selection not all properties met all of these criteria. The results of this project have not yet been published.

Examples of the approaches / methods used to select meaningful and relevant soil properties for understanding soil health include:

- Weighting / scoring systems (Wienhold *et al.* 2009)
- Logical Sieves (Black *et al.* 2011; Corstanje *et al.* 2017; Rickson *et al.* 2012; Ritz *et al.* 2009)
- Score cards (AHDB/ BBRO: https://ahdb.org.uk/soil-health-scorecard; https://www.agricology.co.uk/resources/testing-soil-health-scorecard-farm-soil-monitoring-2018%E2%80%932019; Romig *et al.*1997; https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/health/assessment/?cid=nrcs142p2_053871)
- Dashboards (Frontier: https://www.frontierag.co.uk/images/CropNutritionAdvice/Soil-Life-Leaflet.PDF; Agrii, Hutchinsons)
- Minimum data sets (Jiang *et al.* 2020; Raiesi 2017)

Rinot *et al.* (2019) and Laishram *et al.* (2012) have defined three main steps of indicator creation and methods for each step (Figure 2). Step 1 involves the selection of the MDS of most pertinent attributes. This could be from Expert Opinion (e.g. Principal Component Analysis). Inclusion of an attribute in an indicator has been based on expert opinion (e.g. through workshops and consultations) and / or by statistical tools relating soil properties to soil functions and delivery of ecosystem goods and services. Statistical tools could include principal component analysis, multiple correlation, factor analysis, cluster analysis and star. Comparing the two approaches, expert opinion has been suggested as being open to the possibilities of disciplinary biases which can be avoided by using quantitative approaches (Bachmann & Kinzel 1992; Doran & Parkin 1996).

Step 2 requires the chosen attributes to be transformed into properties that can be ‘scored’ to allow comparisons of the final score over space and time. Finally (Step 3), the individual attributes (and their scores) are combined to create the overall indicator. The calculation could be a simple summation of all scores, or with weighting added to emphasise a particular aspect (e.g. to reflect current policy drivers such as food production over habitat protection, etc.).

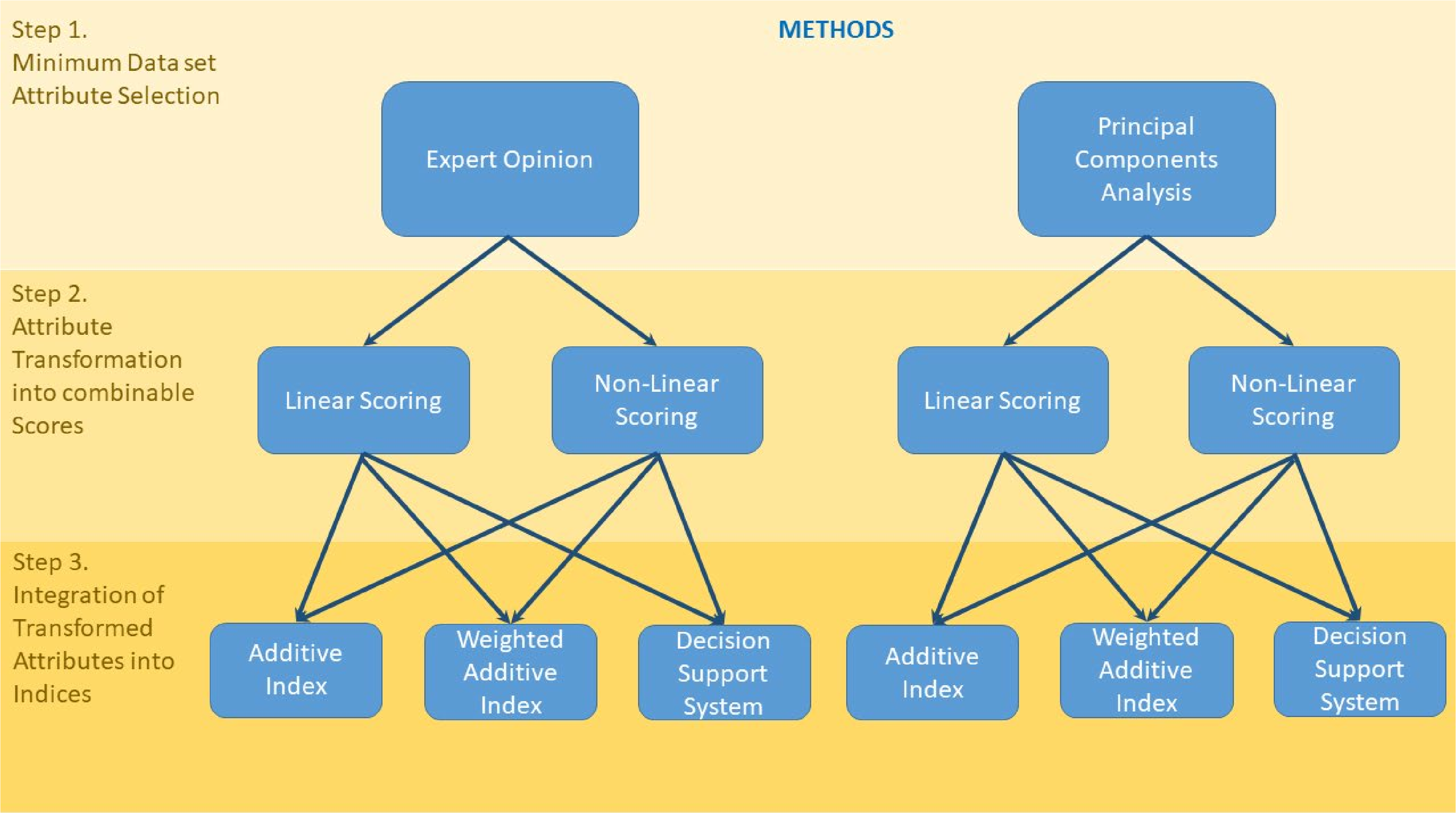

Figure 2. Flow diagram of the three steps of indicator creation and alternative methods for each step (adapted from Laishram *et al.* 2012).

Crookston *et al.* (2021), emphasise the need for appropriate interpretation of soil health assessments through robust databases of soil properties and their variation across large regional areas. Arguing that the interpretation of soil health assessment relies on databases regionally calibrated to edaphic and climatic factors so that values are comparable to similar soils. Reporting uncertainty in temporal and spatial variability is necessary to validate soil condition benchmark values, but few regional databases presently contain these details. Careful consideration of sampling intensity and inherent variability of different soil attributes is required while combining several soil attributes as one synthetic indicator. Warrick and Nielsen (1980) report that two, 110 and 1,300 samples were required to achieve the same level of precision in estimation of bulk density, percent clay and soil hydraulic conductivity respectively. But Crookston *et al.* (2021) showed variation in attributes to be higher for many biological properties than for chemical and physical ones and that the range of variation was dependent on soil taxonomy and texture. Crookston *et al.* (2021) also emphasised the need for long-term (more than 10 years) monitoring to establish definitive soil health temporal variation patterns.

3.1 Examples of derived Soil Quality Indicators

Examples of derived soil quality indicators (and constituent soil properties include:

- Microbiological degradation index (Andrews *et al.* 2002; Bastida *et al.* 2006). Includes:
    - Total organic carbon
    - Water soluble carbon
    - Water soluble carbohydrates
    - Microbial biomass carbon
    - Respiration
    - ATP
    - Dehydrogenase
    - Urease
    - Protease
    - Phosphatase
    - Beta-glucosidase activity
- Soil fertility, soil productivity, resource sustainability and environmental quality (Singer & Ewing 1998)
- General indicator of soil quality (GISQ, Velasquez *et al.* 2007). Summed the contributions of each of five sub-indicators (hydraulic properties, chemical fertility, aggregation, organic matter and biodiversity) to derive the general indicator of soil quality (GISQ). 50 soil properties related to macrofauna, chemical fertility, physical state, organic matter fractions and soil morphology. Allows the evaluation of soil quality and facilitates identification of problem areas through the individual values of each sub-indicator:
    - i) PCA analysis of the variables allowing testing of the significance of their variation among land use types;
    - (ii) identification of the variables that best differentiate the sites according to the soil quality;
    - (iii) creation of sub-indicators of soil physical quality, chemical fertility, organic matter, morphology and soil macrofauna, with values ranging from 0.1 to 1.0;
    - (iv) combination of all five sub-indicators into a general one.

- Qualita Biologica del Suolo (QBS) Index (Parisi *et al.* 2005): Evaluation of microarthropods' level of adaptation to the soil environment life rather than the species richness/diversity. Overcomes issues relating to spices classification.
- Vegetation Attributes as a Surrogate to the Soil Quality
- The soil health management assessment framework (Andrews *et al.* 2004): Established a set of decision rules to select the relevant indicators. This set was based on the management goals for each site, associated soil functions and other selection criteria such as regional or crop tolerance, sensitivity, and other inherent properties such as climate conditions, soil taxonomy, etc. "Each indicator has a unique combination of goals, functions, and additional criteria that must be satisfied for it to be suggested as a MDS indicator." These rules can serve as a framework for an expert opinion-based system to select the most relevant attributes for each site. An additive index yields a number between 1 and 10 (Andrews *et al.* 2004). However, if assessed soil functions or ecosystem services rank very differently in importance, then some kind of weighting is mandatory.
- Morphological classification using Visual Soil Assessment (VSA, Moncada *et al.* 2014): Used decision trees, which are based on predetermined thresholds of relevant explanatory variables that should explain SH as the response variable. Often summarized in an overall soil quality rating (McGarry 2006; Mueller *et al.* 2014; Shepherd *et al.* 2008). Scores for the different indicators are summed up, with some weighting applied.
- Comprehensive Assessment of Soil Health (CASH) - Cornell University Soil Health Laboratory (Gugino *et al.* 2009; Idowu *et al.* 2008; Moebius-Clune *et al.* 2016): Designed for farmers, gardeners, agricultural service providers, landscape managers and researchers. It has 39 properties, limited to 12 to 13 physical (e.g. aggregate stability, penetration resistance [surface and subsurface hardness], available water capacity) chemical (pH, P, K, Mg, Fe, Zn, enhanced extractable Al, As, B, Ba, Ca, Cd, Co, Cr, Cu, Mn, Na, Ni, Pb, S, Sr, organic matter content) and biological (soil proteins, soil respiration, soil pathogens, Active C). Compared across three texture classes (coarse, medium, and fine). It mainly compares percentiles to known population distribution in a textural group. By comparing observations against frequency curve values to obtain a number between 0 and 10. All indicator value being averaged to produce a single value of soil health. It has colour coding for different indicators alone or aggregated according to soil function based on threshold frequency values above which a particular indicator exceeds a critical environmental threshold value a traffic light system of three to five colours indicates low, adequate or excessive values for a given indicator. Future work needs to associate thresholds with agronomic and environmental outcomes appropriate to soil, climate, and production system (Stott 2019). Wander *et al.* (2019) suggest it is costly compared to routine tests and still need to be proven.
- Haney Soil Test or Haney Soil Health Tool (HSHT, Haney 2010, 2012; Haney *et al.* 2018): Based on 15 years of soil testing research in the USA. Indices based on a small number of properties (3 to 5). Based on three biological attributes (soil respiration, dissolved organic nitrogen, dissolved organic carbon.) CO_2-C (24 hr); water extractable organic C and N; Oxalic, malic, and citric acid (H3A) extractable P, K, Mg, Ca, Na, Zn, Fe, Mn, Cu, S and Al; total water and H3A extractable NO_3-N, NH_4-N and PO_4-P. The index represents microbial activity and nutrient availability as the main factors describing SH. Wander *et al.* (2019) suggest it is costly compared to routine tests and still need to be proven.
- Fine's Index (Fine *et al.* 2017). Used PCA to examine the relationships between SH indicators measured in distinctively different regions across the US. They found a relatively high number (6) of significant principal components (PCs), where the

dominance of the first PC was relatively low. Consequently, they suggested that each PC tends to represent a relevant process (e.g. physical, biological, or chemical) that is differently expressed in each soil sample. Their findings support the use of indices based on a larger number of indicators rather than smaller.

- Multivariate-complex soil health approach (Rinot *et al.* 2019). Based on the ocean healthy index developed by Halpern *et al.* (2012). Quantifies the relationships between soil attributes and ecosystem services. Allows transition from point measurement to broad scale analysis allowing consideration of different land uses, soil types, climatic regions, and precise definition of effects of SH on human activities. Three-step approach:
    - Taking soil samples from a wide range of soil types, land uses, climatic regions, etc.; measurement of a large number of chemical, biological, and physical attributes to provide a broad database from which minimum data set can be based; use of quantitative statistical models rather than expert-based models, using autocorrelation and principal component analysis (PCA), followed by clustering techniques to identify the most representative attributes.
    - Conversion of raw data into normalized scores; ES rather than soil attributes serving as target values for the assessment of soil functioning; requiring a detailed framework, including methods for setting reference points; this allows evaluation of the quality of the services and benefits provided by the soil; establishment of a comparative SH indicator on a large scale; least squares being used to utilize correlation between selected measured soil attributes.
    - Integration providing a coefficient for each attribute and contribution to each ES and whole model; low contributions being eliminated; holistic model, combining all relevant ES constructed based on relative contribution of each soil ES at a specific site.

  Such models must be based on a wide range of soils, land uses and agroecosystems. Allows identification of the most significant and universal attributes for quantifying the relative contribution of each attribute to each ES and to assess the health of soils.
- Canadian monitoring of soil quality: calculated as the weighted average of the performance indices for erosion, soil organic carbon content, trace elements and soil salinization (Clearwater *et al.* 2016).
- Muencheberg Soil Quality Rating: The weighted sum of the basic indicators is multiplied with values for hazard indicators such as contamination, acidification, and flooding (Mueller *et al.* 2014).
- Soil Management Assessment Framework (SMAF): site-specific interpretations for soil health indicator results for crop and pasture lands. It uses measured soil health indicator data to assess management effects on soil functions by first selecting indicators and appropriate and then interpreting results of indicator measurements. It takes a flexible and context-specific approach to soil quality assessment which has influenced others to apply multivariate statistical methods to select the most relevant indicators (assume but do not assess connections between indicators and soil functions). However, the drawback of flexible approaches is the limited comparability between studies. It uses scoring curves consisting of interpretation algorithms. It includes four microbial or biochemical indicators: SOC, PMN, MBC, and BG, all represented by more-is-better curves (Andrews *et al.* 2004). Measures included:
    - Nematode maturity index
    - Metabolic quotient determined from soil respiration and microbial biomass

  - Bulk density
  - Total organic C
  - Microbial biomass C
  - Potentially mineralizable N
  - Soil pH
  - Soil test P
  - Macroaggregate stability
  - Soil depth
  - Available water holding capacity
  - Electrical conductivity
  - Sodium adsorption ratio
- Great Soils (AHDB): Index based on pH, P, K, Mg, texture, organic matter, CO_2 burst test, earthworm numbers, compaction, crop symptoms.
- Soil health scorecard (AHDB): Threshold values associated with each attribute. A traffic-light system flags up whether anything requires investigation (red), monitoring (yellow) or, if things are good, where no action is needed (green). Key nutrient thresholds were based on values taken from AHDB Nutrient management guide (RB209).
- Frontier Soil Life Report: Each indicator is allocated a score and a weighted average is used to calculate a Soil Health Index for soil.
- New Zealand Visual Soil Assessment (VSA, Shepherd *et al.* 2000): Based on the visual scoring of 7 key physical indicators of soil quality, against a score card. Also considers plant performance.
- Northern Rivers Soil Health Card. The soil health card was developed as part of the Good Soil Project and the Good Worm Project, initiatives of Tuckombil Landcare Inc. It consists of 10 visual soil tests. A soil management tool developed by farmers for farmers.
- Soil Health Card Scheme (National Mission for Sustainable Agriculture, India): Location as well as crop-specific sustainable soil health management.

Flexible frameworks such as that proposed by Andrews *et al.* (2004) tend to be complex and easily manipulated. Expert opinion-based frameworks contain simple and robust decision rules but weaken the relationship between soil properties and soil functions (Rinot *et al.* 2019).

However, expert-based systems utilising the accumulated knowledge of scholars and practitioners (rather than a specific database) have been used to establish soil quality indicators (Andrews *et al.* 2002), where different management goals or soil functions need representation. Ritz *et al.* (2009) utilised an expert-based framework to select relevant biological measures of soil health appropriate for a national scale monitoring programme. Ritz *et al.* (2009) argued that attributes selected by experts and stakeholders must be capable of addressing national soil/environmental protection policy requirements. However, Bünemann *et al.* (2018) argue that expert-based frameworks may lack methodological transparency and simplicity, precluding wide application of MDS selection across different environments and cropping systems and across wide range of soil attributes.

The North Central Education and Research Activity Committee (NCERA-59) caution on any “rush to enshrine” a standard suite of dated measures as it may be incompatible with longer-term goals (Wander *et al.* 2019). Particularly when there is inadequate data and no agreed upon interpretive frameworks for measurement of soil health. They suggest that methods should be tailored to local conditions in order to effectively apply and interpret attributes for different soil resource regions and land uses.

Where SH indicators are used there remains an issue of interpretation as to what either a single summative result or set of values collectively mean to soil health, and how this information can be utilised to inform appropriate action (Bonfante *et al.* 2020). Bonfante *et al.* (2020) suggest that an integrated approach can be achieved using the separate attributes in a simulation model that address a key soil function, such as its contribution to the ecosystem service of biomass production. In their example they used attributes related to soil structure, water and air regimes, and organic matter content to simulate the soil-water-atmosphere-plant system.

Various modelling approaches have been used, including expert decision tree systems, Neural networks, and process-based models, to assess change in outcome based on change of input parameters (De la Rosa & Sobral 2008). The use of process-based models to evaluate soil health can require high-quality and high-frequency data (De la Rosa & Sobral 2008). However, gaps in data can be filled using either pedotransfer functions or neural networks.

3.2 Minimum data sets used in deriving Soil Health Indicators

The use of minimum data sets (MDS) is often questioned because it inevitably excludes some variables. This exclusivity contradicts the need for wider soil data to represent all interest groups and the diverse soil functions and breadth of ecosystem goods and services derived from soils. However, MDS make practical sense for a national soil monitoring network (Doran & Parking 1996; Arshad & Martin 2002; Merrington *et al.* 2006). The use of a MDS can also be supported through the use of an interpretative framework in which other attributes are not excluded but used for specific purposes (Merrington *et al.* 2006).

According to Merrington *et al.* (2006), the purpose of a MDS is to be “used broadly to assess relatively high-level trends in soil quality and risks or threats to sustainable land management practices.”

The key qualities of a MDS have been suggested to be:

- Readily interpretable (Schipper & Sparling 2000; Seybold *et al.* 2003; Merrington *et al.* 2006; Bünemann *et al.* 2018)
- Relatively precise (Schipper & Sparling 2000; Seybold *et al.* 2003; Merrington *et al.* 2006)
- Not readily estimated from other variables (Schipper & Sparling 2000; Seybold *et al.* 2003; Merrington *et al.* 2006)
- Save time and money (Bünemann *et al.* 2018)

Stable soil attributes – although relevant to soil functioning – may not be included in a minimum dataset (Bünemann *et al.* 2018).

It is important to consider that attributes may carry different weights of information across regions or specific locations in terms of importance and influence (Roper *et al.* 2017; Singh *et al.* 2019).

The NCERA-59 warn against selection of measurements based on widely used methods (Wander *et al.* 2019), arguing that subjective interpretation of test results might be a bigger problem than methodological variability. Wander *et al.* (2019) argue that consistently applied methods are desirable for broad scale inventories but tailored measures are likely to perform better for diagnostic applications. Sampling regimes need to be targeted in space and time to minimize confounding factors to enable interpretation. However, Stott (2019), from USDA, counters Wander *et al.* (2019) by suggesting that once SH attributes have been selected there needs to be a standardisation of sampling and handling procedures in field and laboratory methods and protocols. This, they argue, is to ensure that changes in SH can be measured over space and time and allows for interpretation across soil types and climate zones. It will facilitate the development of a national set of regionally appropriate, interpretation functions (i.e. scoring algorithms), against which raw data can be transformed.

Other approaches include using proportional response and consistent parameters to quantify the overall responsiveness and consistency of various physical, chemical, and biological parameters (Gyawali 2019). Gyawali (2019) identified two to four mm aggregate stability and soil magnesium as being the most responsive on an annual timescale as well as differentiating between tillage and cover crop treatments.

4. Soil properties reflecting diverse ecosystem services and functions

Soils inherently differ in their potential capacity to fulfil ecosystem services because of climate, duration of soil development, organismic influence on the soil (including humans, vegetation, and meso/microorganisms), soil source (parent) material, and topographical relief (Jenny 1994; Crookston *et al.* 2021). A soil's capability is defined by its "natural" state (McBratney *et al.* 2014), which implies before it has undergone any improvement (or degradation) through management.

Much of the research on SHIs and SQIs has related soil properties for provisioning of goods and services, notably food production. However, Nortcliff (2002), suggested that soil must provide the following basic functions:

i. a physical, chemical, and biophysical setting for living organisms;
ii. the regulation and partition of water flow, storage, and recycling of nutrients and other elements;
iii. support for biological activity and diversity for plant growth and animal productivity;
iv. the capacity to filter, buffer, degrade, immobilize, and detoxify organic and inorganic substances; and
v. provide mechanical support for living organisms and their structures.

According to Bünemann *et al.* (2018) the mechanistic relationships between soil properties and soil health indicators and soil functions and ecosystem services are not currently clear. Soil health/quality is not routinely linked to other ecosystem services or to soil degradation (Chapman *et al.* 2018; Bünemann *et al.* 2018). Table 6 shows some soil properties that have been related to other ecosystem goods and services such as water quality, human health, and climate control (after Lehmann *et al.* 2020).

Table 6. Soil health / quality properties included in ecosystem-service assessments (adapted from Lehmann *et al.* 2020). *Soil properties included in more than 20% of soil-health assessments are labelled as 'greater than 20%'. Those included in at least one, but fewer than 20% of assessment methods are labelled as 'less than 20%'. Those that are typically not included, but recommended to be included, are labelled as 'proposed'. Those properties less directly relevant for a certain ecosystem service are marked as' - ', while those that are more relevant are marked with '+'. Type of property is defined as B – biological; C – chemical; GHG – greenhouse gas; P – physical.

Soil property	**Inclusion in soil health assessments reviewed (%)***	**Ecosystem service**				**Type of property**
		Plant production	**Water quality**	**Human health**	**Climate control**	
Nitrogen-/sulphur-/ phosphorus-mineralizing enzyme activity	less than 20%	+	+	-	+	B
Nitrogen mineralization	greater than 20%	+	+	-	+	B
Microbial biomass	greater than 20%	+	+	-	+	B
Pathogens	Proposed	+	+	+	-	B
Earthworms	less than 20%	+	-	+	-	B
Biodiversity	**Proposed**	**+**	**+**	**+**	**+**	**B**

Soil property	Inclusion in soil health assessments reviewed (%)*	Ecosystem service				Type of property
		Plant production	Water quality	Human health	Climate control	
Microbial activity	**greater than 20%**	**+**	**+**	**+**	**+**	**B**
Parasites	**Proposed**	**-**	**-**	**+**	**-**	**B**
Fauna	**Proposed**	**+**	**+**	**+**	**+**	**B**
GHG emissions	Proposed	-	-	-	+	B
Organic toxins	Proposed	+	+	+	-	C
Organic chemical fractions	less than 20%	+	+	-	+	C
Organic nitrogen fractions	less than 20%	+	+	-	+	C
Cation-exchange capacity	greater than 20%	+	+	-	-	C
Organic carbon	**greater than 20%**	**+**	**+**	**+**	**+**	**C**
Bioavailable nutrients	**greater than 20%**	**+**	**+**	**+**	**+**	**C**
pH	**greater than 20%**	**+**	**+**	**+**	**+**	**C**
Electrical conductivity	greater than 20%	+	+	+	-	C
Compound diversity	Proposed	-	+	-	+	C
Mobile nutrients	greater than 20%	-	+	-	+	C
Heavy-metal toxins	greater than 20%	+	+	+	-	C
Pore-size diversity	Proposed	-	+	-	+	P
Aggregation	**less than 20%**	**+**	**+**	**-**	**+**	**P**
Water storage	**less than 20%**	**+**	**+**	**+**	**+**	**P**
Penetration resistance	**greater than 20%**	**+**	**+**	**-**	**+**	**P**
Infiltration	**less than 20%**	**+**	**+**	**+**	**+**	**P**

It is generally agreed that the inclusion of biological properties to assess changes in ecosystem functions is accepted practice both at national and European scales (Griffiths *et al.* 2018). However, biological properties still lack standardised measurement procedures and validation against different types of land use (Griffiths *et al.* 2018). Different stakeholders also require different information and therefore different properties are required to answer specific questions.

5 Newly emerging ways to measure soils

It is claimed that some newly emerging methods to measure soil properties linked to soil health may allow more detailed assessment of soil processes and functions. They also could have the potential to be developed into high-throughput soil analyses, which would provide better understanding of spatial and temporal variability of soil properties that are linked to soil functioning and delivery of goods and services.

Rapid evolution and decreasing costs associated with emerging methodologies (such as soil spectroscopy and metagenomics sequencing of soil biology) will facilitate their inclusion in soil health indicators and assessments. However, current limitations include the absence of standard operating procedures (SOPs) and accepted threshold values, especially for molecular methods, making the comparison and the interpretation of the results challenging (Callahan *et al.* 2016). The interpretation of these novel soil quality properties is limited by the lack of functional linkages with soil processes and land management implications. Novel methodologies need to be related to existing indicators so that eventually the novel methods can replace them when performance (or cost-efficiency) improves.

Recently there has been an increase in capability relating to data mining. This has great potential providing that it is accompanied by metadata such as GPS location and cropping history (Griffiths *et al.* 2018). For most novel soil quality properties, technological, practical and interpretation related issues need to be overcome before they are ready to be deployed in monitoring programmes (Bünemann *et al.* 2018).

5.1 Soil biota

Soil biota are considered the most sensitive properties of soil quality due to their high responsiveness to changes in environmental conditions (Bastida *et al.* 2008; Bone *et al.* 2010; Kibblewhite *et al.* 2008a; Nielsen & Winding 2002). Soil biological properties are currently underrepresented in soil quality assessments and mostly limited to black-box measurements such as microbial biomass carbon and soil respiration (Amat *et al.* 2021). Recent rapid developments in soil biology have prompted the feasibility of properties based on genotypic and phenotypic community diversity (Hartmann *et al.* 2015; Kumari *et al.* 2017; Nielsen & Winding 2002; Ritz *et al.* 2009). Molecular methods focusing on DNA and RNA hold great potential to perform faster, cheaper and more informative measurements of soil biota and soil processes than conventional methods (Bouchez *et al.* 2016). However, these methods do not yet provide general information that can for the basis of soil management decisions (GREATSoils factsheet 18/17). Also, microbial diversity measure using nucleic acid analysis procedures is highly dependent on the method used to extract DNA from the soil. (GREATSoils factsheet 18/17). Consequently, they may yield novel properties that could substitute or complement existing biological and biochemical soil quality properties in regular monitoring programs (Hartmann *et al.* 2015; Hermans *et al.* 2017). Next generation sequencing (NGS) has facilitated a profound change in the analysis of soil and plant-associated fungal comm-unities, to species level (Frąc *et al.* 2018). There are multiple pipelines for the analysis of fungal NGS data available (Mint *et al.* 2014; Gweon *et al.* 2015), so it is not data-analysis that is problematic, but the interpretation of results. recently online resources for functional annotation of fungi have been made publicly available (Nguyen *et al.* 2016). Currently, no widely accepted genomic properties of agricultural soil health exist. Primarily due to a lack of readily available targeted amplicon and shotgun metagenomics sequence data from geographically diverse soils (Norris *et al.* 2020).

5.2 Analytical techniques

Recent data analysis approaches such as network analysis, structural equation modelling and machine learning could facilitate the establishment of links between properties and functions (Allan *et al.* 2015; Creamer *et al.* 2016). The prediction of process rates from the presence and quantity of genes and transcripts is yet to be clearly established (Rocca *et al.* 2015). Potential issues include biases introduced by sample contamination, PCR reaction, choice of primers and OTU definition and taxonomic assignment techniques (Abdelfattah *et al.* 2017; Hugerth & Andersson 2017; Schloter *et al.* 2018). Big data remains challenging in terms of time, computing capacities and interpretation, since a large proportion of soil organisms yet remains to be characterized in taxonomic and functional terms (Schloter *et al.* 2018; Bouchez *et al.* 2016).

Alternative molecular techniques such as Metabolomics (Vestergaard *et al.* 2017) and metaproteomics (Simon & Daniel 2011) may yield potentially suitable soil quality properties because the measurements are directly linked to ecosystem processes (Bouchez *et al.* 2016) difficulty to extract metabolites and proteins from soil and to choose representative samples (Bouchez *et al.* 2016). Stable Isotope Probing (SIP) in conjunction with phospholipid fatty acid analysis (PLFA) and DNA probing could also help to link soil biodiversity to soil processes (Wang *et al.* 2015; Watzinger 2015). Meaningful integration of properties based on molecular methods into soil quality assessments, standardized techniques and a reference system are still lacking and will have to be established (Bouchez *et al.* 2016).

5.3 Total soil organic matter

Currently, changes in response to land management and land use are difficult to detect since the total pool of soil organic matter is large (Haynes, 2005) and its relevance in soil processes is not unequivocal. However, pools of soil organic matter such as labile or active carbon are typically more sensitive to disturbance than total soil organic matter and can give a better indication about soil processes (Gregorich *et al.* 1994). Suggestions to measure this fraction include:

- particulate organic matter (Cambardella & Elliott 1992)
- permanganate-oxidizable carbon (Weil *et al.* 2003)
- hot water-extractable carbon (Ghani *et al.* 2003) and water-soluble carbon, also called dissolved organic carbon (Filep *et al.* 2015).

Their relationship with soil processes is not well understood, as it is not clear which part of the organic matter they represent. Thermal and spectroscopic methods are also rapidly developing (Clemente *et al.* 2012; Derenne & Quenea 2015; Mouazen *et al.* 2016) and hold promise for soil quality assessments.

5.4 Soil sensing approaches

VisNIR and MIR sensors are being developed into more portable forms. These could replace wet chemistry approaches that can be time-consuming, laborious, and costly for large scale sample analysis. Examples include near-infrared spectroscopy and remote sensing. Combining laboratory-based visible and near-infrared spectroscopy with in-situ measurements such as electrical conductivity and penetration resistance may be particularly useful (Veum *et al.* 2017).

VisNIR spectrometry and pressure sensors are being developed into more portable forms to measure soil carbon (Klein 2021), total nitrogen, β-glucosidase activity, active carbon, microbial biomass carbon, particulate organic matter carbon and soil respiration (Karlen 2020). It measures hydrogen and C bonding associated with silicate clays and organic and inorganic C, is strongly correlated with organic C and total N (Norris *et al.* 2020).

Limitations include that the information is gained only about the first millimetres of the soil. Second, sample characteristics such as moisture content, particle size distribution and roughness of the soil surface can influence the outcome of the analysis (Baveye & Laba 2015; Stenberg *et al.* 2010). Prediction is as good as the calibration data set.

The creation of freely available databases that can be used for proper calibration and prediction of soil properties is essential for realising the full potential of these techniques. These databases should involve both NIR spectra and results from wet chemistry and biological methods.

Advances in sensors to quantify apparent electroconductivity could also be used to assess soil texture, mineralogy, cation exchange capacity and water content, through calibration (Karlen 2020). Improvements in vertical penetrometers and mobile, horizontal sensors could help determine penetration resistance and provide information on compaction and soil bulk density (Karlen 2020).

5.5 Soil structure

Soil structure measurements could benefit from X-ray Computed Tomography, 3D imaging of soil. This is still some way from being a routine application for soil quality assessment. However, systems are getting much faster and automated, so throughput and replication are increasing and thus the technique has potential future reach.

5.6 Aggregate stability

Aggregate stability is currently measured through wet and dry sieving. However, in future, it may be possible to use integrated Aggregate Stability (IAS), interprets aggregate stability using a laser diffraction machine. Overall, IAS showed higher correlation with the wet sieving method (R2 = 0.49 to 0.59) than widely used median aggregate size (d50) (R2 = 0.09 to 0.27). IAS can also quantify stability of macro- and micro-sized aggregates, which d50 cannot. When comparing between IAS and wet sieving, IAS requires considerably less time and sample amounts (Gyawali 2019). The smartphone application SLAKES, the SoilInfo App (http://www.isric.org/explore/soilinfo), the LandPKS tool (www.landpotential.org), High-throughput soil analysis approaches, e.g. visual, and near-infrared spectroscope, and the soil quality assessment tool (SQAPP) being developed within EU iSQAPER project could assist with this.

5.7 Mapping soil health

Satellite data could be used to monitor tillage practice, cover crop planting and crop rotations, and soil residue dynamics (how surface material decays). These agricultural management practices are then mapped to greenhouse gas emission and nutrient cycling (University of New Hampshire).

5.8 Chemical analysis

Chemical analyses are currently undertaken using wet or dry analysis. However, in future, Nuclear magnetic resonance and pyrolysis-gas chromatography-mass spectrometry could be used (Arias *et al.* 2005).

6 Conclusions

For many soil properties, texture-dependent scoring curves need to be developed, which is possibly one of the greatest challenges (Bünemann *et al.* 2018).

The principle of identifying reference sites with acknowledged good soil quality (Rutgers *et al.* 2008, 2012) would be more suitable (Bünemann *et al.* 2018). However, monitoring changes of properties over time can reduce the importance of (subjective) reference values (Griffiths *et al.* 2018).

An overall soil quality indicator is often desired, but not very meaningful since soil quality is best assessed in relation to specific soil functions. Rather than calculating an overall indicator, a graphical representation of how well a given soil fulfils its various functions is much more effective in communicating with stakeholders, target users and the public (Bünemann *et al.* 2018).

According to Laishram *et al.* (2012), existing knowledge provides understanding of the current capacity of a soil to function but fails to predict the capacity of the soil to continue to function under a range of stresses and disturbances.

No current evidence suggests a single or generic set of properties that can represent all scenarios. It is important to remain open to inclusion of new methods and technologies that may bring advantages of cost saving and speed of analysis. Karlen (2020) suggests the way forward is to 1) improve understanding of soil biology; 2) develop better in-field and remote-sensing data collection techniques; and 3) aim to interpret soil biological, chemical, and physical data more holistically.

7 Application of the “Logical Sieve” approach to synthesize the findings of the Literature Review

Following the approach used by Black *et al.* (2008; Defra SP0534), Ritz *et al.* (2009) proposed a generic framework that supported a structured approach to the identification and subsequent selection of potential properties for monitoring both the current and changing status of soils. The conceptual framework, which the authors termed a ‘Logical Sieve’ (LS), was originally devised to provide an objective and quantitative means of ranking potential soil biological properties that are used as properties of soil quality / soil health. Run in an MS Excel based software tool, the Logical Sieve approach adopts a formalised method for assessing the relative strengths, weaknesses, and suitability of these candidate soil health measurements. The approach enables a consistent synthesis of available information and the semi-objective assessment of properties when tested against a series of scientific and technical criteria (known as ‘challenge criteria’).

Black *et al.* (2008) stated that the power of this approach had two clear advantages: it provides a clear record and audit trail on the decision-making process; and it can accommodate future inclusion of a) emerging candidate methods and b) new challenge criteria that determine the suitability of each candidate method as evaluated by the Logical Sieve. Also, the approach was designed to be adaptable to meet future changes in both application and any pertinent new data / knowledge or policy direction. The ‘sieving’ functions to select the properties are designed to be flexible, so that they can accommodate different user-defined scenarios. For example, criteria used to select appropriate properties for biodiversity may be different for those selected specifically to address carbon sequestration or food production. Equally, different land uses may be pertinent to different questions, for example salient SHIs may be different for arable land compared to moorland (e.g. soil nutrient levels). The Logical Sieve allows the criteria to be modified, emphasised (i.e. weighted) and/or omitted altogether in the scoring system, as considered appropriate. Ritz *et al.* (2009) considered the approach to be objective, realistic, able to integrate emergent knowledge and adaptable to changing end-user or policy requirements.

Later, Rickson *et al.* (2012; Defra SP1611) used the Logical Sieve approach to identify soil physical quality properties (SQIs), for potential use in national scale soil monitoring programmes. In SP1611, SQIs were scored across four separate categories (each of which included a number of subcategories):

- soils’ ability to deliver ecosystem goods and services (EG&Ss) – does the candidate SQI reflect all the EG&Ss that can be delivered by soil(s)? The sub-categories were: Provisioning; Regulating; Supporting and Cultural. The aim was to score each potential SQI as to whether it was ‘meaningful’ (i.e. that any change in its unit of measurement / value had a corresponding impact on the delivery of ecosystem goods and services).
- land use – does the candidate SQI apply to all land uses? The subcategories were based on the CEH’s land cover map (See Table 8 below)
- soil degradation – does the candidate SQI reflect soil degradation processes? Subcategories were: erosion; compaction; sealing; diffuse soil contamination; loss of organic matter; loss of soil biodiversity
- Does the SQI ‘pass’ the ‘challenge criteria’ used by Merrington *et al.* (2006) (see explanation below) (included 14 subcategories)

In Rickson *et al.* (2012), 49 candidate physical SQIs were scored by expert knowledge within the project team, augmented by a systematic literature review. A group of independent experts was also used to validate the process. This scoring process resulted in 7 SQIs being short listed (or ‘sieved’) as they:

a) could be related to soil processes, soil functions and delivery of ecosystem goods and services;
b) passed the technical challenge criteria (such as ‘ease of measurement’, etc.); and
c) had direct relevance to current and likely future soil and environmental policy.

The Logical Sieve approach has been applied in this project as the culmination of the systematic literature review process to identify the overall scientific consensus as to which are the most salient properties to measure. For this project, all categories (and subcategories) of the Logical Sieve were populated using the information extracted from the extensive sources reviewed in Work Package 1 (Literature Review). (It is recognised that it was not possible to review every global study of SHIs / SQIs given time limitations of the project, but the methodology of using the LS to extract pertinent information to devise a short list of properties has been demonstrated). Of the four categories used by Rickson et al (2012) (soil function/ delivery of EG&Ss; land use; soil degradation; and challenge criteria), the ability of the properties to indicate soil degradation processes was not included here. This is because current soils policy (e.g. as driven by the 25 Year Environment Plan) emphasises the beneficial role of soils in delivering EG&Ss to society, rather than a focus on the current state (and threats) to soil resources (as was previously the case under the policy driver of the EU Thematic Strategy on Soil Protection).

7.1 Structure of the Logical Sieve

The LS is in a matrix format and is operated in an Excel spreadsheet. The rows of the spreadsheet are the candidate SH properties; the columns are the categories (and their subcategories – see below) against which each soil health property will be scored. The ‘ReadMe’ tab describes the four categories used in the scoring process (soil function/ delivery of EG&Ss; land use; soil degradation; and challenge criteria). The scoring system is also described. The scores are generated from the information provided in the sources reviewed (as listed in the ‘Reviews’ tab).

The ‘JNCC Logical sieve’ tab shows the scoring based on literature reviewed for the present project. The ‘JNCC logical sieve PLUS’ tab shows the scoring based on the current literature review in combination with information already known to exist in literature (but not specifically reviewed in this project). The ‘Bio, Phys, Che’ tab shows how all of the individual soil properties mentioned in the literature have been classed as either a physical, biological, or chemical soil property. This tab also shows how similar terms used by authors for the same soil property have been grouped together. For example, the following terms have all been assigned the label ‘Aggregate stability’: “Aggregate stability” (16 papers); “Wet aggregate stability” (3 papers); “soil aggregate stability” (1 paper); “Aggregate strength and stability” (1 paper); “Soil aggregate stability (2 mm, 0.25 mm and 0.053 mm)” (1 paper).

7.2 Soil function / delivery of ecosystem goods and services

Properties were assessed in relation to their pertinence to defined soil functions / delivery of EG&Ss and ascribed a score for each ecosystem goods / service of interest (Table 7).

Table 7. Pertinence of soil health properties to the 'Soil Functions / Ecosystem Goods & Services' category of the Logical Sieve, with representative scoring system

Soil function / ecosystem goods and services (EG&Ss) category	Score used in the Logical Sieve
Provisioning (P) Food (including seafood and game), crops, wild foods, and spices fibre and fuel genetic resources volume and quality of water (domestic, industrial, agricultural use) pharmaceuticals, biochemicals, and industrial products energy (hydropower, biomass fuels) **Regulating (R)** climate regulation water regulation water purification/detoxification air purification/detoxification of waste carbon sequestration waste decomposition, bioremediation, and detoxification crop pollination pest and disease control **Cultural (C)** spiritual and religious value inspiration for art, folklore, architecture etc social relations aesthetic values cultural heritage cultural, intellectual, and spiritual inspiration recreational experiences (including ecotourism) scientific discovery **Supporting (S)** soil formation and retention nutrient cycling primary production water cycling provision of habitat nutrient dispersal and cycling seed dispersal	0 = Not pertinent 1 = Poorly related 2 = Pertinent 3 = Highly pertinent

7.3 Land use

Within a national survey, a diverse range of land uses will be encountered and ideally, any property should be meaningful on all land uses found at the national scale. Similarly, the methodology used to determine the property of choice should be applicable on every land use encountered. Therefore, assessing the applicability of each property to land use classes identifies which are most suitable and meaningful properties of soil health in the widest range of land use types. To a certain extent, land use will also reflect differences in land management (e.g. annual cultivations on arable land as opposed to pasture). The range of land uses considered was based on the CEH's land cover map (LCM 2000) designations (Table 8).

Table 8. Pertinence of soil health properties to the 'Land Use' category of the Logical Sieve, with scoring system.

Land use category	Score used in the Logical Sieve
Arable & horticultural (A) Improved grassland (IG) Unimproved/rough grassland (UG) Woodland/forestry (W) Moorland (M) Bare ground (BG) Urban (U)	0 = Not pertinent 1 = Pertinent

7.4 Challenge criteria

The selected 'challenge criteria' were those previously derived and applied in the EA project 'The development and use of soil quality indicators for assessing the role of soil in environmental interactions' (Merrington *et al.* 2006). These were integrated with criteria used to identify the best soil degradation indicators in the ENVASSO project (Huber *et al.* 2008). These criteria and their scoring system were revisited and adapted to meet the requirements of the current project (Table 9).

Table 9. Proposed challenge criteria to evaluate soil health properties (combining Merrington *et al.* (2006) and Huber *et al.* (2008) recommendations), with representative scoring system.

Challenge criteria	Scores used in the Logical Sieve
Relevance/significance (Sig): 1. The property must be relevant to the function of environmental interaction, and it must be interpretable in quantitative terms as an indicator of soil health and the temporal changes in soil health. 2. Allied to this is the issue of clarity. It must be clear what interpretation can or cannot be placed on an indicator. 3. It may be useful to consider properties as direct or indirect indicators of a soil function. Thus, a catchment hydrograph is an indirect indicator of rainfall interception and storage by soils, but changes in soil water storage following rainfall is a direct indicator.	0 = not relevant 1 = relevant 2 = very relevant
Measurability, sensitivity, discrimination, and signal-to-noise ratio (Mes): Practicability of properties depends on efforts needed for monitoring, data gathering and for indicator calculation (requirement for further interpretation and/or information). For wide application of the properties the complexity as well as the effort and costs of data gathering, and calculation of the indicator values should be acceptable for decision makers. This criterion is linked strongly with data availability. In order to be operational, properties should be easily measurable and quantifiable. 1. Soil properties are notoriously spatially variable: 50% is not unusual as the standard errors of the mean of many typical soil parameters. Against this, many soil parameters change only slowly with time. Thus, long term monitoring must attempt to discriminate long term trends from "noisy" backgrounds.	1 = poor 2 = good 3 = very good

Challenge criteria	Scores used in the Logical Sieve
Measurability, sensitivity, discrimination, and signal-to-noise ratio (Mes): (continued) 2. In selecting properties, we need to consider the probability of detecting significant changes over the sampling intervals, Thus, for example, if a parameter is likely to change by 5% between samplings, and the 95% confidence limits of the measured mean are equivalent to 50% of the mean, it will be many years before a significant change is detected. 3. This leads to the idea of the undetected change. Properties should be evaluated against the time span over which significant changes will go undetected; and whether such changes, once detected, are already irreversible. Ideally the time over which a change is undetected is minimised. These aspects are easily determined using simple statistical procedures. 4. We should not adopt properties which, because of significant variability (due either to actual spatial or temporal variability or to sampling and measurement errors), are unlikely to detect change over reasonable time intervals.	See row above.
Measurable in the field (MF) To be able to measure in situ	0 = not practical 1 = practicable
Measurable in the lab (ML) All or part of the measurement must be undertaken in a laboratory (ex situ).	0 = not practical 1 = practicable
Practicability and analytical soundness (Sou): 1. How practicable is a potential indicator? Are there robust, proven methods for its measurement? Are such methods in the pipeline? Or will they need considerable development? In the latter case, there would need to very strong reasons to include an indicator which would require significant further development. 2. Where such reasons do exist, possibly because the indicator furnishes information unavailable in any other way, the project should support such further development. 3. The methodological approach to calculate the indicator has to be technically and scientifically sound, based on international standards and international consensus about its validity and its suitability for linkage to economic models, forecasting and information systems.	0 = not practical 1 = practicable
Efficiency and cost (E&C): 1. We should seek to maximise the use of automatic methods including sensors, remote sensing, and automatic data retrieval. Potential properties should be examined against the need to minimise cost and maximise efficiency. 2. Allied to this is the general consideration of cost. Potential properties must be assessed against the likely cost of populating them over 5, 10 and 20 years.	1 = high 2 = moderate 3 = low

Challenge criteria	Scores used in the Logical Sieve
Integrative properties (II): 1. Wherever possible we should be looking for integrative indicators. These are properties which effectively integrate the information from a number of subsidiary properties. One example is the catchment hydrograph, which reflects the average hydrology of the soils in the catchment. 2. However, integrative properties should only be adopted where they can be interpreted in terms of one of the key soil functions. In the case of catchment hydrographs, for example, it is still difficult to extract quantitative information on soil hydrology from what is a very smeared picture.	0 = no 1 = yes
Policy Relevance (Rel): Policy relevance of properties is expressed by their thematic coincidence with key topics such as the 25 Year Environment Plan (25 YEP). In order to be of value for policy decision-making, key issues and soil properties should be related to policy objectives for soil and to environmental or other policy agendas where soil management is a central issue (as required by Defra and other key stakeholders, including Natural England, Environment Agency, and Forest Research/Forestry England).	0 = not relevant 1 = relevant
Geographical coverage (Geo): Geographical coverage indicates the area where the indicator or the input parameters needed to calculate the indicator have already been monitored. For the selection of properties, special attention should be given to properties already implemented, especially if the coverage is extensive. The advantage is a high applicability and most likely a high acceptance. But this should not hinder new developments if another soil property is more suitable to illustrate the key issue.	1 = poor 2 = good 3 = very good
Availability of baseline and threshold data (Bas): This criterion indicates whether or not baseline and or threshold values have been established for the evaluated indicator. In order to have the possibility of relative comparison over time the availability of baseline and threshold data is important. Baselines and thresholds enable an assessment of a suitable use of soil and needs for effective measures to avoid a critical status of soil degradation. If no baseline or threshold values are available yet, their development should be possible with reasonable effort.	0 = not available 1 = available

Challenge criteria	Scores used in the Logical Sieve
Comprehensibility and clarity (Com): Comprehensibility describes the level of expert knowledge needed to understand the information on the situation of a soil threat provided by an indicator. The properties should be generally understandable in order to facilitate communication of results provided by properties to the public and political decision-makers. The final information should be clear and easy to interpret. Behind it, complex functions/models can be used, but those have to be combined in a logical and clear structure. That is, a measure that can be used by policymakers and other stakeholders to understand soil health and if/how it is changing, requiring no further interpretation and/or information for an assessment of soil health. Properties should be interpretable by users, e.g. in terms of where soil health is currently good or bad.	0 = incomprehensible 1 = weakly comprehensible 2 = comprehensible 3 = highly comprehensible
Complementarity and/or conflict (Con) This criterion will consider potential complementarity and/or conflict between current and newly emerging SHIs.	0 = Conflicting 1 = Complementary
Sensitivity to change (Sen) This criterion will consider how sensitive the measure is at picking up change.	0 = not sensitive to change 1 = sensitive to change 2 = very sensitive to change
Soil texture Number of soil textures mentioned.	0 = not pertinent 1 = not all soil textures 2 = range of soil textures

7.5 Weighting factors

Each category / subcategory of the LS can be given a weighting factor to introduce flexibility and better sensitivity within the system. It is recognised that the importance of each of the 3 Categories (and their subcategories) used in the Logical Sieve can vary depending on the question being asked (and by whom). For example, the ability of a property to reflect the provisioning function of soil may be the most important consideration, if food production is the primary ecosystem service to be delivered by soil. On the other hand, the cost or practicality of measuring a particular property (i.e. one of the technical ‘challenge criteria’) may be of highest importance when commissioning and implementing a national soil monitoring programme. The weighting of each factor can reflect these different priorities. The weighting factors can be used to put different emphasis on criteria, depending on the question being asked at the time.

In the current use of the Logical Sieve (i.e. to demonstrate the methodology of using the LS to select salient properties as ‘proof of concept’), no weighting factors were imposed on the scores over the 3 Categories (and their subcategories).

7.6 Scoring

Each property was given an individual numerical score for each subcategory (columns in the LS spreadsheet) in all three categories. The scoring values (defined in Tables 6 to 8 above) and summed scores are designed so that higher aggregated scores reflect greater suitability for a given end use. The final scoring was based on the evidence extracted from the literature review.

For the present exercise, an additional score was included in the scoring system. The number of literature sources citing any given property was noted, on the assumption that frequently occurring properties reflected a consensus of salient properties within the scientific community. Thus, an additional score was given to each property, relating to the number of unique times an indicator was mentioned. This value was based on ranked order of frequency of unique mentions in the literature, divided by 100 (e.g. 36/100 = 0.36).

The resultant scores were collated and used to formulate a list of the most suitable properties. For this initial proof of concept, a simple ranking of cumulative (additive) scores (incorporating 0 scores, if any) was used, covering all three categories (soil function / delivery of EG&Ss), land use and challenge criteria, multiplied by the score based on the unique number of mentions (frequency of occurrence in the reviewed literature). No weightings have been applied. For JNCC Logical Sieve, scores were based on information from the reviewed literature. For JNCC Logical Sieve_PLUS, scores were based on information from the reviewed literature, plus information known to exist in literature (but not specifically reviewed in this project).

7.7 Results of the Logical Sieve

A total of 48 soil properties associated with soil health were identified from the reviewed literature. Following scoring and ranking (highest ranking scores assumed to be most suitable), the ranked indicator appeared in the order shown in Table 10 (shown in descending order of ranked value).

From the results of the JNCC Logical Sieve_PLUS, soil texture is ranked 10th, but this should be omitted because a) for a given location, this soil property is unlikely to change over time and as such, has limited use in monitoring changes in soil condition over time (as is a prerequisite of a SHI) and b) different types of soil texture (clay, sands, silts, peats etc.) and how these affect values is considered later in the project (Annex 3). Similarly, porosity (ranked 11th) was considered to be strongly related to and represented by bulk density (ranked 2nd), so both properties are not necessary for soil health measurement and monitoring. Mineralizable N (14th rank) may primarily reflect the fact that the vast majority of SHI/SQI studies are focused on the provisioning function of soils (i.e. agricultural production), where nutrient status is key. However, soil nutrient content may not necessarily be as critical if another outcome (e.g. biodiversity) is the desirable ecosystem service to be delivered by the soil.

Based on this approach, the properties highlighted in purple (representing the top 25%) were suggested as the most relevant to carry forward to the next phase of this project (i.e. the ecological modelling of soil properties affecting soil function and the delivery of different ecosystem goods and services).

Table 10. The 48 ranked properties based on simple additive scoring, based on: literature review only; and literature review plus known evidence in other publications.

Rank	JNCC Logical Sieve (reviewed literature only)	Rank	JNCC Logical Sieve_PLUS (reviewed literature + other sources)
1	pH	1	Soil organic matter / soil carbon
2	Bulk density	2	Bulk density
3	Soil organic matter / soil carbon	3	pH
4	Nitrogen	4	Infiltration/hydraulic conductivity
5	Infiltration/hydraulic conductivity	5	Water holding capacity
6	Soil texture	6	Nitrogen
7	Water holding capacity	7	Soil respiration
8	Soil respiration	8	Aggregate stability
9	Aggregate stability	9	Microbial/fungal diversity
10	Microbial/fungal diversity	10	Soil texture
11	Earthworms	11	Porosity
12	Porosity	12	Soil structure + aggregate distribution
13	Mineralizable N	13	Earthworms
14	Enzyme activity	14	Mineralizable N
15	Electrical conductivity	15	Microbial biomass carbon
16	Soil structure + aggregate distribution	16	Enzyme activity
17	Phosphorous/phosphate	17	Penetration resistance
18	Potassium (K)	18	Electrical conductivity
19	Microbial biomass carbon	19	Phosphorus/phosphate
20	Penetration resistance	20	CEC
21	C mineralisation	21	Potassium (K)
22	Compaction	22	Compaction
23	DNA/Genomics	23	Roots
24	Roots	24	C mineralisation
25	Erosion	25	Erosion
26	Soil depth	26	Nutrient status
27	Soil slaking and crusting	27	Soil depth
28	Nutrient status	28	DNA/Genomics
29	Microbial biomass N	29	Microbial biomass N
30	N fixation	30	Soil slaking and crusting
31	CEC	31	Soil moisture
32	Nematodes	32	Soil protein

Rank	JNCC Logical Sieve (reviewed literature only)	Rank	JNCC Logical Sieve_PLUS (reviewed literature + other sources)
33	Soil moisture	33	Soil colour
34	Soil protein	34	Nematodes
35	Phosphatase	35	N fixation
36	Arthropod	36	Heavy metals
37	Dehydrogenase	37	Arthropod
38	Heavy metals	38	Dehydrogenase
39	Enchytraeids	39	Phosphatase
40	Crop yield	40	Shear strength
41	Springtails	41	Enchytraeids
42	Mites	42	Crop yield
43	Lipid profiling	43	Bacteria
44	Shear strength	44	Soil temperature
45	Soil colour	45	Springtails
46	Bacteria	46	Mites
47	Soil temperature	47	Lipid profiling
48	Sorptivity (cm s-O.5)	48	Sorptivity (cm s-O.5)

References

Andrews, S.S., Karlen, D.L. & Mitchell, J.P., 2002. A comparison of soil quality indexing methods for vegetable production systems in Northern California *Agric. Ecosyst. Environ*. 90 (1), 25-45.

Arias, M.E., Gonzalez-Perez, J.A., Gonzalez-Vila, F.I. & Ball, A. (2005). Soil health-a new challenge for microbiologists and chemists. *International Microbiology*, 8, 13-21.

Andrews, S.S., Karlen, D.L. & Cambardella, C.A. 2004. The soil management assessment framework: a quantitative soil quality evaluation method. Soil Sci. Soc. Am. J. 68, 1945-1962. https://doi.org/10.2136/sssaj2004.1945.

Bünemann, E.K., Bongiorno, G., Bai, Z., Creamer, R.E., de Deyn, G., de Goede, R., Fleskens, L., Geissen, V., Kuyper, T.W., Mader, P., Pulleman, M., Sukkel, W., Van Groenigen, J.W. & Brussaard, L.T. (2018) Soil quality - a critical review. *Soil Biol. Biochem.* https://doi.org/10.1016/j.soilbio.2018.01.030.

de la Rosa, D. & Sobral, R. (2008) Soil Quality and Methods for its Assessment. In: Braimoh, A.K. & Vlek, P.L.G. (eds) Land Use and Soil Resources. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6778-5_9.

Ewing, S.A. & Singer, M.J. 2012. Soil quality. In: Huang, P.M., Li, Y., Sumner, M.E. (Eds.), Handbook of Soil Science: Resources, Management and Environment Impacts. Boca Raton, Taylor & Francis Group, pp. 1-26.

Fine, A.K., van Es, H.M. & Schindelbeck, R.R. 2017. Statistics, scoring functions, and regional analysis of a comprehensive soil health database. *Soil Sci. Soc. Am. J.* 81 (3), 589-601.

GREATSoils factsheet 18/17: Biological test for soil health. Produced by AHDB. (last accessed November 2021 at https://projectblue.blob.core.windows.net/media/Default/Imported%20Publication%20Docs/GREATsoils%20Factsheet%20-%20Biological%20Tests%20for%20Soil%20Health.pdf)

Gugino, B.K., Idowu, O.J., Schindelbeck, R.R., van Es, H.M., Moebius-Clune, B.N., Wolfe, D.W., Thies, J.E. & Abawi, G.S. 2009. Cornell Soil Health Assessment Training Manual, Edition 2.0. Cornell University, Geneva, NY Available from. https://www.canr.msu.edu/foodsystems/uploads/files/cornell_soilhealth.pdf.

Halpern, B.S., Longo, C., Hardy, D., McLeod, K.L, Samhouri, J.F., Katona, K, Kleisner, K, Lester, S.E., O'Leary, J., Ranelletti, M., Rosenberg, A.A., Scarborough, C., Selig, E.R., Best, B.D., Brumbaugh, D.R., Chapin, F.S., Crowder, L.B., Daly, K.L, Doney, S.C., Elfes, C., Fogarty, M.I., Gaines, S.D., Jacobsen, K.J., Karrer, L.B., Leslie, H.M., Neeley, E., Pauly, D., Polasky, S., Ris, B., St Martin, K, Stone, G.S., Sumaila, U.R. & Zeller, D. 2012. An index to assess the health and benefits of the global ocean. Nature 488,615-620. https://dotorg/10.1038/nature11397.

Haney, L.R. 2012. Soil Health Test https://www.midwestlabs.com/wp-content/uploads/2012/09/190Haney-Test-Explanation.pdf.

Idowu, O.J., Van Es, H.M., Abawi, G.S., Wolfe, D.N., Schindelbeck, R.R., Moebius-Clune, I.N. & Gugino, B.K. 2009. Use of an integrative soil health test for evaluation of soil management impacts. *Renewable Agric. Food Syst* 24 (3), 214-224.

Lilburne, L., Sparling, G. & Schipper, L. 2004. Soil quality monitoring in New Zealand: development of an interpretative framework Agric. *Ecosyst Environ*. 104 (3), 535-544.

Loveland, P.J. & Thompson, T.R.E. 41 (Environment Agency, Bristol, 2002).

Moebius-Clune, B.N., Moebius-Clune, D.J., Gugino, B.K., Idowu, O.J., Schindelbeck, R.R., Ristow, A.J., van Es, H.M., Thies, J.E., Shayler, H.A., McBride, M.B., Wolfe, D.W. & Abawi, G.S. 2016. Comprehensive Assessment of Soil Health – The Cornell Framework Manual, Edition 3.0, Cornell University, Geneva, NY.

Moncada, M.P., Gabriels, D. & Cornelis, W.M. 2014. Data-driven analysis of soil quality indicators using limited data. *Geoderma* 235,271-278.

Rinot, O., Levy, G.J., Steinberger, Y., Svoray, T. & Eshel, G. (2019) Soil health assessment: A critical review of current methodologies and a proposed new approach. *Science of the Total Environment*, 648:1484-1491.

Ritz, K., Black, H.I., Campbell, C.D., Harris, J.A., Wood, C.., 2009. Selecting biological indicators for monitoring soils: a framework for balancing scientific and technical opinion to assist policy development. *Ecol. Indic*. 9 (6), 1212-1221.

Roper, W.R., Osmond, D.L., Heitman, J.L., Wagger, K.G., & Reberg-Horton, S.C. 2017. Soil health indicators do not differentiate among agronomic management systems in North Carolina soils. *Soil Sci. Soc. Am. J*. 81 (4), 828-843.

Shepherd, T.G.; Ross, C.W.; Basher, L.R. & Saggar, S. (2000) Visual soil assessment, Volume 2.- Soil management guidelines for cropping and pastoral grazing on flat to rolling country. horizons. MW & Landcare Research, Palmerston North.

Svoray, T., Michailov, E., Cohen, A., Rokah, L. & Sturm A., 2012. Predicting gully initiation: comparing data mining techniques, analytical hierarchy processes and the topographic threshold. *Earth Surf. Process. Landf* 37 (6), 607-619.

USDA: National Resources Conservation Services (NRCS) of the US Dept. Agric. Soil Health, available at: https://www.nrcs.usda. gov/wps/portal/nrcs/main/soils/health/ (last access: October 2021), 2019.

USDA. (2006). Soil Quality Institute. Natural resources conservation service. Available at http://soils.usda.gov/sqi/.(Last access: October 2021)

Verheijen, E.G., Bellamy, P.H., Kibblewhite, M.G. & Gaunt, J.L. Organic carbon ranges in arable soils of England and Wales. *Soil Use Manage*. 21, 2-9 (2005).

Wood, M. & Litterick, A.L. (2017) Soil health - What should the doctor order? *Soil Use and Management* 33, 339-345.

Appendix 1: Methodology for the literature review

We reviewed the evidence contained in scientific data, reports and research papers related to SHIs and their derivation. These sources included information arising from government and regulators (e.g. Defra, EA, Natural England); extension / advisory organisations (e.g. ADAS); industry representatives (e.g. AHDB, agronomy companies, consultants); assurance schemes (e.g. Global Gap, LEAF); research groups; and other sources (e.g. producers and retailers (Waitrose Agronomy group) and environmental groups (Sustainable Soils Alliance)). We included the international literature on SHIs too.

Using a combination of bibliographic databases (e.g. Web of Knowledge and Scopus), search engines (e.g. Google Scholar and Bing), conference proceedings and sources available to experts within the team, the scientific and grey literature was searched on SHIs and the methodologies that have derived them. Boolean operators ‘AND’ and ‘OR’ were used to combine search terms into text strings to refine the search for specific words in the title, abstract and keyword of articles (Table 11). Preliminary investigation suggested that much of the work would be sourced from technical reports and academic journals (a search on Scopus of keywords ‘soil health indicators’ gave 175 papers).

Table 11. Search terms used in the literature review.

String	Google	Google Scholar*	Papers pulled off (assuming no repetition, ca. 186)
“soil health” AND review	6,900,000	57,200 20,400 between 2016 to 2021	19
“soil health” AND review AND method*	9,020,000	18,000	10
“soil health” AND measurement AND technique	6,890,000	22,900 17,600 between 2016 to 2021	37
selection of “soil health indicators”	43,000	1,970	9
“soil health index”	20,100	412	19
“soil health index” AND calculation?	69,900	207	3
“soil quality index” AND calculation?	122,000	3,530	8
“Soil quality” OR “soil health” AND monitoring AND indicators	1,100,000	17,000	32
“Soil quality” OR “soil health” AND monitoring AND indicators AND “ecosystem services”	238,000	17,000	22
Monitoring “soil health” OR “soil quality”	7,210,000	17,400	5
Soil health indices	24,500,000	517,000	7

String	Google	Google Scholar*	Papers pulled off (assuming no repetition, ca. 186)
Soil quality indices	51,900,000	1,150,000	2
Index of soil quality	134,000,000	3,670,000	0
Forest* AND soil AND “soil health” OR “quality indices” OR “quality index”	8,210,000	16,200	0
Urban AND soil AND “soil health” OR “quality indices” OR “quality index”	7,820,000	17,500	7
Natural soil AND “soil health” OR “quality indices” OR “quality index”	8,510,000	17,100	
Soil AND health OR quality AND threshold* AND critical limit*	29,300,000	393,000	6

The review covered:

- peer-reviewed literature
- grey literature where appropriate (e.g. websites)
- previous reports on SHIs (Table 12), many of which have involved the project team
- direct discussions with relevant organisations where key information is not easy to find using conventional academic databases or websites.

Table 12. Selected Defra reports related to soil health / soil quality properties and indicators (http://randd.defra.gov.uk/).

Column header	Column header
SP1623	Developing an innovative approach to monitoring soil health in England and Wales
F0122	Potential for greater use of legumes in grassland systems for improved biodiversity, soil fertility and livestock health
F0370	Farm practice and soil health
F0401	Changes to soil quality indicators following conversion to organic vegetable production
OF0370	Farm practice and soil health
P0120	Impacts of heavy metals on soil quality with respect to microbial activity
P0310	To develop a robust indicator of soil organic matter status
SP0501	Effects of farm manure additions on soil quality and fertility
SP0504	Effects of fertiliser nitrogen additions on soil quality and fertility
SP0512	Identification and development of a set of national indicators for soil quality
SP0514	Sampling strategies and soil monitoring
SP0515	Comparability of soil properties derived from different data sources
SP0529	SQID: Soil quality indicators – developing biological indicators

Column header	Column header
SP0534	Scoping biological indicators of soil quality – phase II
SP0538	The impacts of climate change on soil functions
SP0544	Development of performance criteria for soil monitoring schemes
SP0546	Soil Organic matter as a headline indicator of soil health
SP0548	Soil indicator robustness testing: Food and fibre
SP0554	Soil indicator robustness testing: Foundation for the built environment
SP0558	Design of a UK Soil Monitoring Scheme
SP08006	Soil Quality Indicators Workshop – SQID
SP1611	Indicators of the quality of the physical property of soil
WQ0121	Upland agriculture – balancing productivity, water, and soil quality

Appendix 2: Desirable attributes of an indicator of soil quality/health by publication

Desirable attributes of an indicator of soil quality/health by publication:

- Lehmann *et al.* (2020)
  - relevant to soil health, its ecosystem functions, and services;
  - sensitive, by changing detectably and quickly without being reflective of merely short-term oscillations;
  - practical, by being conducted cheaply and with a short turnaround time for analysis;
  - informative for soil / land management.
- Rinot *et al.* (2019)
  - Measurement of relevant (physical, biological, and chemical), scientifically based data.
  - Open to sensitivity analysis to clarify variations in soil functions caused by soil management, land use, climate change, etc.
  - Manageable, available, accurate and cost-effective measurements, which can be conducted at a relevant time scale for decision making.
  - Reflection of the connection between soil functions and management targets (e.g. agricultural production and ecosystem services).
  - Target values need to be precisely defined and determined for selection and integration of the measured soil attributes.
  - Correlation between properties should be examined to minimize the number of required measurements (i.e. identifying and justifying a ‘minimum data set’). Alternatively, use of analysis of variance (ANOVA) could be used to exclude attributes that do not change significantly in response to different management regimes or crops (D’Hose *et al.* 2014). Principal component analysis (PCA) can assess most relevant attributes under different crop and soil conditions, assessing sensitivity of the attributes rather than significance.
- Stott (2019) [USDA]
  - Soil health indicator effectiveness.
    - Management-Sensitive – The indicator is sensitive to changes in soil and crop management systems.
    - Short-term Sensitivity – The indicator is generally able to detect changes within 1 to 3 years in subhumid to humid climates when significant changes in management are made. Changes are likely to take longer in semiarid to arid climates, or with minor changes in management.
    - Interpretable.
      - The indicator (by itself) represents specific physical, chemical, or biological soil processes or conditions relevant to agricultural production and environmental outcomes.
      - Interpretation with other tests: If not by itself, then the indicator's representation of specific processes/conditions can be interpreted if measured in conjunction with one to two other tests.
    - Useful – The indicator provides useful information towards assessing the SH status of an area and towards addressing specific resource concerns.
    - Production Readiness – Readiness for use in commercial production laboratories in terms of:

  - Ease of Use:
    - Sampling (for field conservation planners, consultants, other agricultural service providers, and producers).
    - Sample submission.
    - Sample preparation (for laboratories).
    - Measurements (for laboratories).
  - Cost effectiveness for producers on a per-sample basis:
    - Labour and supply expenses.
    - Specialized equipment cost.
    - Laboratory space and time requirements/overhead.
  - Measurement repeatability. The level of precision of the method is within acceptable limits.
  - Interpretable for agricultural management decisions.
    - Measured values are "directionally understood" (i.e. more is better, less is better, optimum).
    - Some management practices that improve the measure are known.
    - Regional potential ranges to define relative poor/good functioning are known.
    - Outcome based (yield, resilience, risk, environmental) thresholds are known.
- Bünemann *et al.* (2018)
  - Must be related to a given soil threat, function or ecosystem service and be relevant.
  - Easy to sample/measure.
  - Reliable.
  - Cost effective: Pedotransfer functions can be used to provide a proxy value of a specific soil indicator if it is considered too expensive, too difficult, or not possible to measure. However, the inaccuracy of pedotransfer functions need to be clearly stated.
  - Sensitivity to changes in management.
  - Sensitivity to changes in the environment.
  - Comparable to data from other sampling campaigns.
  - Clear (absolute) interpretation schemes for a given indicator.
  - Indication as to what extent soil quality properties fulfil the requirements of conceptual, practical, sensitivity and interpretation considerations.
- Sharman (2017)
  - Meaningful to business and investor communities so it can be used to drive decision making.
  - Methodology should be clearly understood.
  - Measurable and comparable, allowing for comparison across geographies and time.
  - Possible to aggregate from site level to regional and global scales.
  - Practical data that is accessible, measurable by a company or uses free, globally available data.
  - Ability to substitute better information where available.
  - Replicable and credible, based on a reputable scientific method.

  - Context based: considers local conditions/levels to reflect 'impact' (beyond 'usage').
  - Responsive: Responds to changes in company activities, both short and long term.
- Laishram *et al.* (2012)
  - Soil quality properties would be useful to farmers and planners, only if we know the properties’ critical limits (i.e. the desirable range of values of a given indicator that must be maintained for normal functioning of the soil).
  - Critical limits would vary depending on the goal of management within an ecoregion.
- Parisi *et al.* (2005)
  - sensitivity to variations in soil management;
  - good correlation with the beneficial soil functions and other variables which are difficult to access or measure;
  - helpfulness in revealing ecosystem processes;
  - comprehensibility and utility for land managers (i.e. utility of the indicator as a benchmark in land use decision making);
  - cheap and easy to measure.
- Nortcliff (2002)
  - land use;
  - soil function;
  - reliability of measurement;
  - spatial and temporal variability;
  - sensitivity to changes in soil management; (vi) comparability in monitoring systems;
  - skills required for the use and interpretation.
- Dumanski and Pieri (2000)
  - measurable in space (i.e. over the landscape and in all countries);
  - reflect change over recognizable time periods (5 to 10 years);
  - showing relationships with independent variables;
  - quantifiable and usually dimensionless.
- Karlen *et al.* (1997)
  - easy to measure parameters;
  - rapid/less time-consuming methods;
  - high sensitivity of parameters to detect differences on a temporal and spatial scales.