Article (refereed)

Ritz, Karl; Black, Helaina I. J.; Campbell, Colin D.; Harris, Jim A.; Wood, Claire. 2009 Selecting biological indicators for monitoring soils: a framework for balancing scientific opinion to assist policy development. *Ecological Indicators*, 9 (6). 1212-1221. 10.1016/j.ecolind.2009.02.009.

MS for Ecological Indicators

REVISED: 2 FEBRUARY 2009

Selecting biological indicators for monitoring soils: a framework for balancing scientific opinion to assist policy development

Authors: Karl Ritz[1] , Helaina I.J. Black[2,3], Colin D. Campbell[2], Jim A. Harris[1], Claire Wood[3].

1. Natural Resources Department, School of Applied Sciences, Cranfield University Cranfield MK43 0AL;

2. The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH;

3. Centre for Ecology and Hydrology, Lancaster, Library Avenue, Bailrigg, Lancaster LA1 4AP, U.K.

Running title: Biological indicators of soil quality

Corresponding author:

Prof. Karl Ritz, National Soil Resources Institute, Natural Resources Department, School of Applied Sciences, Cranfield University, Cranfield MK43 0AL, UK

Email: k.ritz@cranfield.ac.uk

Direct line +44 (0)1234 758064

Pages: 35

Tables: 7

Supplementary tables (online only): 5

Figures: 0

Word count: 6157

ABSTRACT

Soils are one of the most important features of the natural capital of terrestrial ecosystems. There is a strong and increasing policy requirement for effective monitoring of soils at local, regional and national scales. However, it remains unclear which properties of soils are most appropriately monitored. This is partly due to the wide range of goods and services that soils provide, but also their inherent chemical, physical and biological complexity. Given that the biota plays such fundamental roles in the majority of ecosystem services provided by soils, biological properties are logical candidates as effective indicators, to complement other physico-chemical properties. A plethora of biological methods have been suggested as indicators for monitoring soils but few are used in national scale monitoring or are published as international standards. A framework for selecting ecologically-relevant biological indicators of soil quality for national-scale soil monitoring that covers the full range of ecological functions and services of soil was devised. The literature was surveyed to identify 183 candidate biological indicators which were then scored by experts and stakeholders against a wide range of scientific and technical criteria. The framework used the scores and weightings to then rank, prioritise and select the indicators. This semi-objective approach using a "logical-sieve" allowed repeated iterations to take account of end-user requirements and expert opinion. A ranked list of 21 indicators was produced that covered a range of genotypic-, phenotypic- and functional-based indicators for different trophic groups. Four of these were not deemed sufficiently robust for ready deployment in a national-scale monitoring scheme without further methodological development. The suite of indicators identified offers the strongest potential candidates for deployment in national-scale soil monitoring schemes provided standard operating procedures are defined and their

inherent sensitivity, ability to discriminate between soil:land-use combinations, and provide ecologically interpretable signals is confirmed. The power of the approach adopted here is that it provides a clear record and audit trail on the decision-making process, enables different priorities to be set contingent on the nature of the desired monitoring, and can direct and allow the inclusion of further methods or indicators into the framework.

Key-words: Biological indicators, Soil health, Soil quality, Monitoring.

1. Introduction

Human societies have always been highly dependent upon healthy soils for their nutritional, economic and social well-being and the associated requirement for delivery of ecosystem goods and services ((Millennium Ecosystem Assessment 2005). Consequently, requirements for monitoring soil quality are increasing and targeted at gathering information on soil functions, rather than soil properties *per se*. These functions include biomass production for food and fibre, protection of our environment through interactions between soils, air and water, support of habitats and biodiversity, protection of archaeological remains, provision of a platform for building, and provision of raw materials (Blum 2005).

Such is the acknowledged significance of these functions that they now form the basis for the proposed legislative protection of soils in the emerging European Communities Framework Directive (Commission of the European Communities 2006a;Commission of the European Communities 2006b) and have been used to identify requirements for soil protection and management within the UK policy framework (SEPA 2001;Defra 2004;Environment Agency 2004;Towers et al. 2006). These policies have to

accommodate challenging scientific issues since soils are amongst the most complex systems on the planet.

Soil functions represent aggregated properties and processes such as decomposition, nutrient cycling, water retention and release, and the regulation of populations. Whilst physico-chemical properties of soils provide the fundamental context in which such functions operate, and have acknowledged utility in assessing ecological status, the majority of soil processes are underpinned by the soil biota. However, a mechanistic understanding of the relationships between soil biodiversity and function, whether in relation to the soil compartment or entire ecosystem, are undeniably complex and remain elusive (Bardgett et al. 2005;Fitter et al. 2005;Hooper et al. 2005). Whilst the definition of pertinent biological indicators is a challenging task, it is reasonable to propose that biological indicators should be considered in any monitoring of soil quality (Francaviglia 2008). Biological indicators, by virtue of their involving complex *adaptive* systems (i.e. the biota) integrate multi-dimensional phenomena such as the delivery of key soil processes in ways that other indicators do not. Biodiversity is a soil attribute in itself, and therefore implicit within the ecosystem approach (Doran and Zeiss 2000). Numerous reviews and reports have been published on soil biological indicators, with much emphasis on ecotoxicological perspectives, but many of these have direct relevance to national-scale monitoring schemes (Buchs 2003;Gadzala-Kopciuch et al. 2004;Arias et al. 2005;Becaert and Deschenes 2006). Biological indicators have already been deployed in a number of schemes throughout Europe and elsewhere (Parris 1998;Ditzler and Tugel 2002;Black et al. 2003;Lilburne et al. 2004;Winding et al. 2005). Although comparability between different international schemes may be desirable, from a scientific and political perspective, consideration is required to ensure that biological indicators chosen for

deployment in a nationwide monitoring framework are scientifically and technically appropriate. They must also be capable of addressing national soil/environmental protection policy requirements and are therefore practicable for the environment they are being applied to.

Within the UK, the deployment of soil biological indicators in an extensive monitoring scheme was first carried out in the Countryside Survey of Great Britain (Black et al. 2003). The policy focus is now on establishing the most appropriate biological indicators for nationwide monitoring from an immense number of potential indicators and associated methods. These actions parallel other initiatives in the UK, Europe and North America (e.g. (Countryside Survey 2008;ENVASSO 2008;NEON 2008;Programme 3 2008). However the numbers and scope of published information on potential biological indicators of soil quality has expanded rapidly in recent years. For example a Web of Knowledge based search for journal papers using a suite of keywords associated with biological indicators and soils shows an essentially exponential increase in number of papers from the 1970's to date (totalling in excess of 17,500 by 1970-2008). The selection of biological indicators has thus gone beyond the reasonable scope of a conventional considered literature review or a standard meta-data analysis. Most reviews of biological indicators have a strong discipline bias, orientated for example to microbial, invertebrate or ecological processes. As (Niemeijer and de Groot 2008) point out, formal criteria relating to the utility of a particular indicator within the total collective set of indicators are rarely considered. We considered that any approach to the selection of biological indicators should be objective, realistic, sufficiently flexible to accommodate emergent knowledge and adaptable to changing end-user or policy requirements. To accommodate these issues, we devised a generic framework that supported a structured

approach to the identification of potential indicators for monitoring the status and change in soils, with the specific purpose of formulating a list of candidate biological indicators of soil quality that demonstrated most potential for application in a national-scale soil monitoring programme. As such, the decision-making process was informed by the requirements of the UK Soil Indicators Consortium (UKSIC 2008), a group of public stakeholders developing a set of soil indicators and a soil monitoring scheme for the UK, but could be applied for selecting indicators in other ecological contexts.

2. Methods

2.1. Basic approach

In considering a concept as broad as biological indicators of soil quality, we first had to define the boundaries of what we were attempting. Following the approach of (Doran and Zeiss 2000), we asserted that the quality of a soil relates to the provision of an appropriate set of soil properties and processes necessary for effective soil function i.e. to provide soils that are *fit for purpose*. Given this context, biological indicators can then be used to assess the status and change in ecological soil properties and processes within a physico-chemical context.

Accepting we had to constrain the scope of the study, our basic approach was as follows:

(a) A subset of three pertinent ecological soil functions was prescribed, *viz*. food and fibre production, environmental interactions, and habitats and biodiversity support. These functions have been adopted by a wide stakeholder community and hence have resonance with both science and policy communities. These functions are dependant upon a range of ecological processes and properties

(Table 1) and therefore most likely to be informed by the use of biological indicators of soil quality. This list of ecological processes and properties is not detailed, and serves to establish what primary information would be required from biological indicators to usefully inform on the individual functions.

(b) A comprehensive list of potential biological indicators of soil quality pertinent to the prescribed functions and ecological properties / processes was established by a trawl and review of the peer-reviewed literature in October 2004. An indicator was designated by either the ecological property or process that is assessed, and by the method used to measure this property or process.

(c) A simple database was constructed where each indicator was assigned a unique reference number (#) and then categorised to one of four primary categories, denoted *Genotypic, Phenotypic, Functional* or *Other*, and one of eight secondary categories denoted *Activity, Biomarker, Biomass, Enzyme, Fauna, Nucleic acid, Process* or *Other* (Table S1).

(d) Each potential indicator was then assessed by a 'factor score' (F) derived from the following three categories, each with defined scoring criteria:

    (i) *Pertinence* to the defined soil functions, denoted F_{SF} (Table 2).

    (ii) *Applicability* to the range of ecosystems under consideration, and ability to *discriminate* between soils that are intrinsically different in relation to the considered criteria, denoted F_{AD} (Table 3).

    (iii) *Technical* category, relevant to implementation in a national-scale soil monitoring scheme, denoted F_T (Table 4).

(e) For each criterion within the technical category, weighting factors were prescribed by a process of stakeholder consultation with the scientific community and likely end-user public-bodies. Respondents assigned weighting factors ranging from 0 (lowest) to 4 (highest) to each of the 13 technical criteria to reflect their views on the extent or temporal scale of any proposed application, their end-user requirements and actual or perceived budgetary constraints.

(f) Each potential indicator was then given an individual numerical score for each criterion in all categories; an indicator was only left unscored if its relevance was unknown.

The possible range of score values within each sub-set were set according to the nature of the corresponding criterion or question. Certain of the categories can only actually have a binary response, and are phrased as such. For example, for certain assays - such as some *in situ* activity methods - soil material simply cannot be archived; one-stop (single punctuated) sampling in the field is either tenable or not, and so on. We also elected to keep the majority of the other scores to three since it made the assessment procedure simpler and more rapid. Of course, more than three scores could be assigned to such non-binary categories, and this would increase the effective range (and precision) of the final score list, but we considered that given the nature of the categories such precision would actually be illusory, and at the expense of ease-of-application of the logical sieve. For the consultation phase, we provided a potential range of five scores, for just this purpose of expanding the final score range, since here the questions are more subjective and more amenable to such an expanded number of potential scores.

2.2. Framework to rank indicators

In order to structure the overall assessment according to our aims, a conceptual framework, termed a 'logical sieve', was devised to provide an objective and quantitative means of ranking the potential indicators. The framework was designed to be sufficiently flexible that it can be re-tuned according to the precise nature of the users' needs, and can be updated as new knowledge is accrued. The 'sieving' functions are also flexible and operate on the principle of achieving a ranking based on grading with respect to user-defined scenarios.

As described above, the assessment was based on numerical scores assigned to each indicator with respect to the prescribed suite of criteria. Scores were assigned using the expert knowledge of the authors and their local peer-groups. Three instances of scoring were carried out, based upon partner organisations (i.e. Cranfield University, Centre for Ecology and Hydrology, Macaulay Institute), where a peer group of scientists within each organisation independently assigned scores to each potential biological indicator according to the project specification. A single score for each indicator:criterion combination was then determined using the arithmetic mean score where two or three of the scores were congruent between organisations; where the scores from each organisation were all different, the scores were debated verbally to reach a consensual value. The resultant scores were then transformed according to the following formulae:

$$F_{SF} = S_{FF} * S_{EI} * S_{HB} \quad (1)$$

where:

F_{SF} = soil function factor

S_{FF} = score for food fibre production

S_{EI} = score for environmental interactions

S_{HB} = score for habitats and biodiversity

$$F_{AD} = S_A * S_D * F_{SF} \quad (2)$$

where:

F_{AD} = applicability/discrimination factor

S_A = score for applicability

S_D = score for discrimination

$$F_T = \sum_{i=1}^{n} (S_{Ci} * W_{Ci}) \ldots + \ldots (S_{Cn} * W_{Cn}) \quad (3)$$

where:

F_T = technical factor

S_{Ci} = score for method category i

W_{Ci} = weighting value for technical category i

n = number of technical categories

$$F_A = F_{AD} * F_T \quad (4)$$

where:

F_A = aggregated factor

This structure operates such that if F_{SF} or F_{AD} is zero, then the indicator under consideration is excluded ("sieved") from subsequent tiers. Similarly, if technical weighting is zero, this criterion is excluded from contributing to F_T and F_A. F_T is a weighted sum since it contains many categories that are assigned different priorities according to the needs of the end-user. In principle, F_{SF} and F_{AD} could be similarly weighted but we did not deem this necessary for the purposes of this study. F_A is essentially a form of "integrated prioritisation" that aims to accommodate all information on an indicator. There is an implicit hierarchy in the tiers, i.e. pertinence to soil functions takes top priority, since if the indicator is not pertinent it should be excluded at the outset. The scoring values and aggregation functions were designed such that higher scores relate to positive attributes from the user's perspective, i.e. "more appropriate" or "more effective"

The data and algorithms that underwrite the logical sieve were incorporated into a database in Microsoft Excel that enabled a high degree of flexibility in terms of data population, exploring resultant scores and producing numerically ranked lists for any of the individual or aggregated factor scores.

2.3. Consultative processes

The overall approach was reviewed by an international peer-group of soil biologists who responded to a series of questions addressing the scientific and technical categories, the weightings and the full list of biological indicators included in the sieve. Following the

primary phase of scoring, a consultative workshop was carried out with a further peer-group of soil biologists from the UK and members of the UKSIC.

3. Results

3.1. Scoring of potential biological indicators

A total of 183 potential biological indicators were identified from the literature and assessed using the logical sieve framework. These indicators are listed alphabetically in Table S1 (Supplementary Material). The scoring assessment involved some 10,000 separate expert decisions. This was not as onerous as would initially appear since the way the sieve was structured, the assessors were able; to scan within and between aggregated indicators or criteria. In the first iteration, most scores were similar between the different scoring groups with completely dissimilar scores only occurring in 4% of the total indicator: criterion cases. The process ensured that, by the final iteration, all decisions were agreed within the core-group (*viz*. the authors of this paper). The consultative processes carried out also helped in refining the final framework.

3.2. Technical weighting factors

The final weights adopted, following respondents' responses to the weighting consultation, are shown in Table 5. The values reflect the relative importance of the individual criteria across the entire stakeholder group. For example, the cost of hardware involved in the measurement of the indicators was generally perceived as being of little significance, whilst reproducibility of measures was generally perceived as being of particular importance. There was a range in values assigned between respondents, with the final values arrived at by a process of debate and attention to ultimate end-user judgements – for example, ‘cost’ was strongly weighted by policy-advice based respondents. This was a pragmatic solution to the challenge of integrating views from a stakeholder consultation involving both research scientists and members of public-bodies.

Deployment status was considered a useful variable to potentially explore indicators which were currently available against those which might in the future be possible, and was weighted zero to prevent it from contributing initially to the F_T or F_A scores.

3.3. Consolidating the ranked lists

The framework enables the collation of indicators to be ranked according to the scores of any of the categories, or the derived criteria. As would be expected, the rank order of the indicators varied according to which category was considered. In terms of the overarching aim of the application of the framework for this study, the aggregated factor (F_A) is most pertinent. The full list of all assessed indicators ranked according to F_A is given in Table S2 (Supplementary Material). This first-order output highlights a degree of repetitiveness in the indicators, arising from different methodological approaches for assessing the same ecological properties or processes. This also highlights an area of inconsistency in the general scientific approach to indicators with the term *biological indicator* being applied to either the methods or the individual parameter of interest, but rarely both. We advocate that indicators are identified by both the parameter of interest and the method, since scientific comparability and consistency depend on having both pieces of information.

To produce a consolidated list fit for application, the primary output needed further refining. To this end, those indicators scored with deployment status of 2 (fully deployable) and an F_A value of greater than 100, were extracted and each indicator was considered in turn moving down the F_A rank. If the indicator was unique, in terms of the ecological property or process measured in comparison to any indicators thus far selected, it was transferred to a consolidated list, but if it was repetitive, i.e. the biological property was already covered by a previous (higher-ranking) indicator, it was passed

over. If there were secondary reasons why the indicator was deemed inappropriate, it was also rejected from the consolidated list. This procedure was repeated until all indicators with F_A scores of greater than 100 had been considered, producing a consolidated list of the top-ranking biological indicators with respect to an aggregate score across all soil functions, which was further aggregated according to deployment status (Table 6). The procedure resulted in thirteen indicators with 'ready for deployment' status, four deemed likely to be ready for deployment in the short-term, and a further four deemed not ready for deployment, with notionally some years development still needed. The logical sieve was further adapted to repeat the above consolidation exercise with respect to each of the three soil functions considered, by weighting to one (unity) the other two soils functions not being considered. These lists are given in the Tables S3-5 in Supplementary Material, and demonstrate that whilst there were often many indicators common to more than one function, others were more appropriate to specific functions.

4. Discussion

4.1. 'Logical sieve' approach

Although the framework is rigorous in its structure, pragmatism is needed to achieve a final useful output. It must be stressed that the logical sieve was designed to act as a structured decision-support tool to assist in formulating a ranked list of indicators, and is not an unequivocal and definitive list – the issues are far too complex for such rigidity to be appropriate. Genotypic, phenotypic and functional categories of indicator were all represented in the consolidated list, with genetic analyses predominating, partly reflecting recent advances in molecular techniques. It was notable that only one indicator (ammonium-oxidiser population structure by cloning and sequencing) was discarded on the basis of secondary considerations (that of throughput limitations, and that it was

effectively included by terminal restriction fragment length polymorphism- (TRFLP-) based approaches which ranked adjacent in the first output). This suggests the robustness of the logical sieve approach in that it also matched the "intuitive" consensus of both the project group and the consultees. Whilst there is an implicit ranking within these lists, the relative positions of the different indicators should not be over-interpreted. The point is that these indicators have been identified by an objective process.

The logical sieve allows the agreed database to be interrogated in a number of ways, reflecting the priorities of different users. It provides an auditable trail of how indicators were identified in terms of their respective priorities. It is also flexible and can be readily applied to scenarios other than those we adopted here; for example, it may be an individual, group or agency requires that any samples taken must be conducive to being archived, and could use this property as a primary sieve criterion. The sieve inputs may be updated as new techniques emerge, or by amendment as existing techniques are modified and improved. This basic framework could also be widely adopted for any class of environmental quality analysis reliant on a number of potential indicators for example plant species and dynamics.

4.2. Gap analysis

The robustness of the approach, and the appropriateness of the top-ranking indicators, can be checked by cross-referencing each of the three key soil functions prescribed at the outset. This simplified process confirms a generally comprehensive coverage of ecological properties and processes, and especially when the biological indicators are taken as a complete suite (Table 7). All functions, with the exception of N fixation, have at least one representation by an indicator, and the majority at least three. Tolerance/resistance to toxins and biodiversity are most represented. The latter would be

expected given the inherent nature of most biological indicators, and the former demonstrates why biotic indicators play key roles in ecotoxicology. The added value of this stage is that, if considered important, it is possible to go back to the framework and identify the most appropriate indicators to fill any gaps.

4.3. Candidate biological indicators

4.3.1. Soil microbial taxa and community structure using TRFLP techniques

Several nucleic-acid methods scored highly in the final rankings since they relate directly to microbial diversity and function. They also have many practical advantages, in terms of archivability, high-throughput and potential transferability of data. In this context, analysis via TRFLP was considered the most appropriate since it is currently deployable. This circumstance will undoubtedly change with the advent of ever higher-resolution, faster and cheaper sequencing technologies. This is why gene arrays are included in the final list but under Deployment Status 0. This notwithstanding, further work is required to identify the most suitable primers and optimise the polymerase chain reaction (PCR), restriction and fingerprinting steps for the identified organismal groups. It is also notable that the methods, including different primers, have not yet been coherently applied to a wide range of soil types and consequently there is no systematic understanding of discrimination potential and sensitivity to large-scale spatial and temporal variation.

4.3.2. Soil microbial community structure and biomass from PLFAs

The use of extracted lipids, in particular phospholipid fatty acids (PLFA), as signature lipid biomarkers, has become widely used to study soil microbial communities (Zelles 1999). The total PLFA content is indicative of the total viable biomass and individual PLFAs, or suites thereof, can be related to community structure as they are found predominantly (but not exclusively) in distinct groups (e.g. fungi, bacteria, Gram-

negative bacteria, actinomycetes). The main advantage of PLFA profiling is that it is a semi-quantitative method, does not rely on cultivability and provides wide coverage of the soil microbial community. Such biochemical phenotypic profiling is arguably particularly suited to monitoring contexts since by definition it integrates the genotype with the environment. It also has the virtue of not relying on a 'species concept' which confounds soil ecology due to the prevalent extreme levels of genetic diversity, and the lack of simple relationships between taxonomic status and functional traits in most soil microbial communities. Although PLFA analysis has been widely used, as with most methods, there has been no systematic study of the full range of soil types that might be covered by a soil monitoring exercise.

4.3.3. Soil respiration and C cycling from multiple substrate induced respiration

Carbon cycling is fundamental to soil function and the respiration of CO_2 from soils, arising from community-level biotic activity, is an intrinsic indicator of C cycling. Measurement of this property in isolation does not provide useful discrimination, and hence ranked low in the logical sieve. However, assays of C mineralisation that put the basal respiration rate into a wider context are considerably more powerful. The concept underlying the multiple substrate-induced respiration (MSIR) approach is to characterise how a soil community responds to exposure to a range of carbon substrates of differing chemical status (Degens and Harris 1997). Respiration determination by use of gas chromatograph is feasible to determine MSIR profiles but is laborious, restricting potential throughput of samples, and unlikely to be feasible for large-scale soil monitoring. There is potential to achieve high throughput systems, for example by application of the MicroResp^{TM} system (Campbell et al. 2003), but this has yet to be rigorously tested across a wide range of land uses and soil types.

4.3.4. Biochemical processes from multi-enzyme profiling

Biochemical reactions in soils are mediated by enzymes produced by the soil biota as part of their metabolic machinery. There is a plethora of enzymes that can be profiled, relating to virtually any defined biochemical transformation, with enzyme activity rates transferable between studies.(Burns and Dick 2002). Many individual enzymes were considered in the logical sieve framework, but ranked lower than the multiple enzyme fluorometric approach since this assay can inform on several ecological processes within a single assay. An increasingly wide range of fluorescently-labelled substrates are available which enable sensitive measurements to be made on small samples, permitting high-throughput assay systems (Marx et al. 2001). This method is suited primarily to enzymes involved in C-cycling, since the majority of fluorescently labelled substrates available target C-transforming enzymes. However, fluorescently labelled substrates that relate to phosphatase and sulphatase are also commercially available, and others may enter the market over time.

4.3.5. Nematodes

Nematodes are among the most abundant multi-cellular soil organisms and their potential as biological indicators of soil quality is already widely reported; several such indicators have been proposed from nematode taxa and community structures e.g. (Mulder et al. 2005). The most widely adopted is the Maturity Index (MI) which reflects the distribution of nematodes across functional groups (Bongers 1990). More amenable indicators are the total number of nematode taxa and abundance of individual functional groups. In contrast to the biochemically-orientated techniques for microbial characterisation, identification of nematodes, even to functional groups, relies on highly trained experts. Nucleic acid

techniques to characterise nematode community structure are emerging (Donn et al. 2008) and have potential to reduce this reliance as well as increase throughput.

4.3.6. Microarthropods

This group, in particular acari (mites) and collembola (springtails), are amongst the most numerous and widespread soil invertebrates in British soils. Both have been proposed as reliable biological indicators however some consideration is required to determine which metrics show the greatest discrimination between soil:land use combinations and sensitivity to environmental pressures and drivers (Parisi et al. 2005). With both groups, the enumeration from dry extraction is fairly straight-forward although higher levels of identification require expert skills and reliable keys for identification.

4.3.7. On-site visual recording of soil fauna and flora

This method (as a combination of different potential indicators) scored highly as it is one of the few approaches that could be used with relative ease to assess the presence of key groups of soil organisms that would otherwise be under-represented, namely ants, fungal fruiting bodies and earthworms (via casts). Truly reliable on-site recording does however require a consistent set of methodologies which have not yet been developed for the UK environment (Swift and Bignell 2001).

4.3.8. Pitfall traps for ground-dwelling and soil invertebrates

There is a substantial body of literature on the use of pitfall traps to assess ground-dwelling and soil invertebrates and the value of these invertebrates as biological indicators (Eyre 2006). Pitfall traps are a well-established technique and have been widely used for environmental surveillance. A disadvantage is this method requires return

visits to a sampling site (e.g. 2 weeks after deployment) which maybe impractical for a national-scale soil monitoring scheme.

4.4. Microbial biomass

The total quantity of life belowground obviously underpins any biologically-related property of soil systems, and is of great significance in many contexts. In almost all circumstances, the majority of the soil biomass is of microbial scale, and hence the microbial biomass intuitively represents a key property of soils in a monitoring context. It was notable that in all instances, biomass *per se* did not enter the upper ranks of the sieve. This was largely due to lower scores associated with the lack of discrimination that gross biomass measures provide in an ecological context. They are all-encompassing ('black-box') measures that have an undoubted direct pertinence to pool sizes in nutrient cycling, principally carbon, but are less informative in other contexts. The same argument applies to basal respiration. However, this is not to assert that these properties have no validity! In the broadest context, these concomitantly broad measures do have roles to play in monitoring contexts, arguably providing the fundamental baseline and of greatest value taken in combination with other higher-resolution measures. It is notable also that certain of the top-ranking indicators inherently include biomass measurements, *viz.* MSIR carries a zero-substrate and glucose-induced respiration measurement, encompassing both basal respiration and biomass; and PLFA can be used as a surrogate biomass measure if appropriate control measures are used in the assay.

5. Conclusions

The selection of a candidate suite of biological indicators is only the first stage towards deployment in national-scale soil monitoring. Following the principles established by (Doran and Zeiss 2000), we need to establish how these indicators are sensitive to

variations in management, correlated with soil functions and can be used to elucidate specific ecological processes. There remains considerable uncertainty over the reliability of these indicators over landscape spatial scales and within and between seasons for use across the UK environment. Method development is required to establish standard operating procedures (SOPs) that would ensure reproducibility of results and resolve practical issues, including cost-effectiveness, for monitoring purposes. There may also be surrogacy between indicators that is not yet apparent since sufficiently coherent datasets are not yet available. This lack of coherence in data and SOPs is a significant issue that thwarts attempts to integrate published data to elucidate discrimination and sensitivity potential for assays. Alongside these issues, effort is required to ensure that the information derived from national-scale soil monitoring will be comprehensible and useful to the end-users, such as land managers and policy-makers. If this information is to be used as an "early-warning", as proposed in so-called Tier 1 monitoring, then we need to identify the observed ranges and envelopes of acceptability for values from the individual indicators in different land uses or ecosystems, and establish whether deviations from these "envelopes" are a sign of beneficial or detrimental changes to ecological processes or properties and the likely consequences for soil functions.

Acknowledgements

The work was funded by the UK Department of Environment, Food and Regional Affairs (Defra). We thank the UK Soil Indicators Consortium, and the many members of the consultative groups we formally approached, for their valuable insights and advice given. We also acknowledge many useful conversations with colleagues and peers which have undoubtedly aided this work.

REFERENCES

Arias, M.E., Gonzalez-Perez, J.A., Gonzalez-Vila, F.J., Ball, A.S., 2005. Soil health - a new challenge for microbiologists and chemists. Int. Microbiol. 8, 13-21.

Bardgett, R.D., Usher, M.B., Hopkins, D.W. 2005. Biological Diversity and Function in Soils. Cambridge University Press, Cambridge

Becaert, V., Deschenes, L., 2006. Using soil health to assess ecotoxicological impacts of pollutants on soil microflora. Reviews of Environmental Contamination and Toxicology, Vol 188 188, 127-148.

Black, H.I.J., Parekh, N.R., Chaplow, J.S., Monson, F., Watkins, J., Creamer, R., Potter, E.D., Poskitt, J.M., Rowland, P., Ainsworth, G., Hornung, M., 2003. Assessing soil biodiversity across Great Britain: national trends in the occurrence of heterotrophic bacteria and invertebrates in soil. J. Env. Management 67, 255-266.

Blum, W.E.H., 2005. Functions of soil for society and the environment. Rev. Env. Sci. Technol. 4, 75-79.

Bongers, T., 1990. The maturity index: an ecological measure of environmental disturbance based on nematode species composition. Oecologia 83, 14-19.

Buchs, W., 2003. Biodiversity and agri-environmental indicators - general scopes and skills with special reference to the habitat level. Agric. Ecos. Env 98, 35-78.

Burns, R.G., Dick, R.P. 2002. Enzymes in the Environment. Marcekl Dekker, New York

Campbell, C.D., Chapman, S.J., Cameron, C.M., Davidson, M.S., Potts, J.M., 2003. A rapid microtiter plate method to measure carbon dioxide evolved from carbon substrate amendments so as to determine the physiological profiles of soil microbial communities by using whole soil. Appl. Environ. Microbiol. 69, 3593-3599.

Commission of the European Communities 2006a. Proposal for a Directive of the European Parliament and of the Council establishing a framework for the protection of soil and amending Directive 2004/35/EC. COM(2006) 232 final. Brussels, 30 pp.

Commission of the European Communities 2006b. Thematic Strategy for Soil Protection. COM(2006)231 final. Brussels

Countryside Survey 2008. The Countryside Survey. www.countrysidesurvey.org.uk, Accessed 1 February 2009

Defra 2004. The First Soil Action Plan for England. Department of Environment Food and Rural Affairs Publications, London.

Degens, B.P., Harris, J.A., 1997. Development of a physiological approach to measuring the catabolic diversity of soil microbial communities. Soil Biol. Biochem. 29, 1309-1320.

Ditzler, C.A., Tugel, A.J., 2002. Soil quality field tools: Experiences of USDA-NRCS Soil Quality Institute. Agron. J. 94, 33-38.

Donn, S., Griffiths, B.S., Neilson, R., Daniell, T.J., 2008. DNA extraction from soil nematodes for multi-sample community studies. Appl. Soil Ecol. 38, 20-26.

Doran, J.W., Zeiss, M.R., 2000. Soil health and sustainability: managing the biotic component of soil quality. Appl. Soil Ecol. 15, 3-11.

ENVASSO 2008. Environmental Assessment of Soil for Monitoring. www.envasso.com Accessed 1 February 2009.

Environment Agency 2004. The State of Soils in England and Wales. Environment Agency, Bristol, UK

Eyre, M.D., 2006. A strategic interpretation of beetle (Coleoptera) assemblages, biotopes, habitats and distribution, and the conservation implications. J. Insect Cons. 10, 151-160.

Fitter, A.H., Gilligan, C.A., Hollingworth, K., Kleczkowski, A., Twyman, R.M., Pitchford, J.W., NERC Soil Biodiversity Program, 2005. Biodiversity and ecosystem function in soil. Funct. Ecol. 19, 369-377.

Francaviglia, R. 2008. Agricultural Soil Erosion and Soil Biodiversity: Developing Indicators for Policy Analyses. OECD, Paris

Gadzala-Kopciuch, R., Berecka, B., Bartoszewicz, J., Buszewski, B., 2004. Some considerations about bioindicators in environmental monitoring. Polish J. Env. Stud. 13, 453-462.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setala, H., Symstad, A.J., Vandermeer, J., Wardle, D.A., 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. Ecol. Mon. 75, 3-35.

Lilburne, L., Sparling, G., Schipper, L., 2004. Soil quality monitoring in New Zealand: development of an interpretative framework. Agric. Ecos. Env 104, 535-544.

Marx, M.C., Wood, M., Jarvis, S.C., 2001. A microplate fluorimetric assay for the study of enzyme diversity in soils. Soil Biol. Biochem. 33, 1633-1640.

Millennium Ecosystem Assessment 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, D.C.

Mulder, C., Schouten, A.J., Hund-Rinke, K., Breure, A.M., 2005. The use of nematodes in ecological soil classification and assessment concepts. Ecotoxicology and Env. Safety 62, 278-289.

NEON 2008. National Ecological Observation Network. www.neoninc.org Accessed 1 February 2009.

Niemeijer, D., de Groot, R.S., 2008. A conceptual framework for selecting environmental indicator sets. Ecol. Indic. 8, 14-25.

Parisi, V., Menta, C., Gardi, C., Jacomini, C., Mozzanica, E., 2005. Microarthropod communities as a tool to assess soil quality and biodiversity: a new approach in Italy. Agric. Ecos. Env 105, 323-333.

Parris, K., 1998. Agricultural nutrient balances as agri-environmental indicators: an OECD perspective. Env. Pollution 102, 219-225.

Programme 3 2008. Risk based methods to assess soil quality. www.programme3.net/soil/index.php, Accessed 1 February 2009.

SEPA 2001. State of the Environment: Soil Quality Report. Scottish Environmental Protection Agency, Stirling, UK

Swift, M.J., Bignell, D.E. 2001 Standard methods for the assessment of soil biodiversity and land-use practice.ASB-Lecture Note 6B International Centre for Research in Agroforestry, South East Asian Regional Research Programme, Bogor, Indonesia. pp. 323-333.

Towers, W., Grieve, I.C., Hudson, G., Campbell, C.D., Lilly, A., Davidson, D.A., Bacon, J.R., Langan, S.J., Hopkins, D.W. 2006. Report on the current state and threats to Scotland's soil resource. Environmental Research Report 2006/01. Scottish Government, Edinburgh

UKSIC 2008. www.defra.gov.uk/environment/land/soil/research/indicators/consortium/. Accessed 1 February 2009.

Winding, A., Hund-Rinke, K., Rutgers, M., 2005. The use of microorganisms in ecological soil classification and assessment concepts. Ecotoxicology and Env. Safety 62, 230-248.

Zelles, L., 1999. Fatty acid patterns of phospho-lipids and lipopolysaccharides in the characterisation of microbial communities in soil: a review. Biol. Fertil. Soils 29, 111-129.

Table 1 - Prescribed soil functions used to assess biological indicators, ecological processes and properties related to such functions* and examples of related soil biota.

Soil Functions	Ecological processes and properties	Examples of related soil biota
Food and fibre production	C cycling	microbial biomass, methanogens
	Decomposition of organic matter	microarthopods, saprotrophic fungi
	N cycling	nitrifiers, denitrifiers
	P cycling	phosphatase, mycorrhiza
	S cycling	sulphur-reducing bacteria
	N fixation	rhizobia
	Primary (microbial) activity	microbial community structure and activity
	Soil food web transfers	microbial community & food web structure
	Disease & pest transmission/suppression	predators, pathogens
	Nutrient supply from symbioses	mycorrhiza, N-fixers
	Redistribution by bioturbation	earthworms, ants
	Bio-aggregation of soil	fungi, worms
Environmental interactions	Degradation/immobilisation of pollutants	fungi, worms
	C retention/release	microbial biomass, methanogens
	N retention/release	nitrifiers, denitrifiers
	P retention/release	microbial activity, mycorrhiza
	Tolerance/Resistance (toxins)	soil community structure and activity
	S retention/release	sulphur-reducing bacteria
	Redistribution by bioturbation	earthworms, ants
	Bio-aggregation of soil	fungi, worms
Supporting habitats and biodiversity	Habitat for rare soil species	wax cap fungi, Southern Wood Ant
	Germination zone for plants	plant roots, mycorrhiza
	Nutrient supply from symbioses	mycorrhiza
	Food source (aboveground)	fungi, insects
	Reservoir for soil biodiversity (taxonomic)	soil species and diversity
	Reservoir for soil biodiversity (genetic)	community DNA and RNA
	Reservoir for soil biodiversity (functional)	nitrifiers, trophic structure, worms

*These functions correspond to UK soil policy documents and predate the recent communication on the European Commission Soil Framework Directive (2006). They can be considered amalga of the individual functions now being relayed in this communication.

Table 2 - Soil functions tier used in assessment framework, and associated scoring values used to assess potential biological indicators.

SOIL FUNCTION	SCORES
FOOD AND FIBRE PRODUCTION Maintaining soil in a suitable state for plant and animal biomass production [supplying nutrients and water, disease control, physical condition]	
ENVIRONMENTAL INTERACTIONS Protecting the capacity of soils to store, transform and regulate soil processes [gas exchanges, degradation and retention of solid materials e.g. pollutants and organic matter, water flow regulation] critical to environmental sustainability	0 = not pertinent 1 = pertinent 2 = highly pertinent
SUPPORT OF HABITATS AND BIODIVERSITY Maintaining the ecological, utilitarian and ethical value of soil biodiversity including maintenance of semi-natural habitats and biodiversity above-ground	

Table 3 - Applicability and discrimination tier used in assessment framework, and associated scoring values used to assess potential biological indicators.

CATEGORY	SCORES
APPLICABILITY Is the property, measured by this method, intrinsically applicable in all circumstances (e.g. ecosystems) under consideration?	0 = Not applicable, i.e. not ubiquitous 1 = Universally applicable
DISCRIMINATION What level of discrimination would method provide between, e.g. 5/10/20 samples from variety of contexts	0 = None 1 = Some discrimination 2 = Moderately high discrimination 3 = Very high discrimination 4 = Extremely high discrimination

Table 4 - Technical tier categories used in assessment framework, and associated potential scoring values used to assess potential biological indicators

CATEGORY	SCORES
THROUGHPUT: How many samples can be processed with optimised laboratory systems and dedicated staff? Assumes soils are in ready for method state (i.e. excludes post-sampling preparation time); rating is for one fully-trained operator.	1 = few per week 2 = dozens per week 3 = hundreds per week
STORAGE: Given appropriate preservation, how soon do post-sampling measures need to be applied?	0 = storage not possible 1 = soon (few days) 2 = can be delayed if suitably stored
ARCHIVABILITY: What is the potential for archiving soil samples (i.e. over decades) in order to accurately re-determine these properties?	0 = not archivable 1 = archivable by freezing, freeze-drying or pickling
SAMPLE COLLECTION: Is one-stop sampling in the field tenable?	0 = no 1 = yes
HOW MUCH SOIL: What mass of soil is needed for sampling and determination?	1 = large mass required (> 1 kg) 2 = relatively small mass (< 1 kg)
COST – HARDWARE: What are hardware costs to realise the method, assuming off-the-shelf technology?	1 = very expensive 2 = moderately expensive 3 = low cost
COST- LABOUR: What are the human resource costs to realise method and initial interpretation (including consideration of skill level required and associated salary)?	1 = very expensive 2 = moderately expensive 3 = low cost
EASE OF USE: What is the amenability of the method to ready application via a standard operating procedure when presented to a competent technician; includes training element?	1 = specialised 2 = moderate 3 = straightforward
POTENTIAL REFERENCE MATERIAL: Is the method amenable to the prescription and provision of such material?	0 = no 1 = yes
REPRODUCIBILITY OF RESULTS: What is the inherent ability for the method to generate reproducible results, given that full quality-control protocols are available and applied, including (assumed) availability of reference material	1 = inherently poor 2 = moderate 3 = high
DEPLOYMENT STATUS: Is the method "off-the-shelf" at the moment, with SOPs or ISO accreditation?	0 = not ready, years development needed 1 = likely to be ready for deployment with some months development 2 = fully deployable, in routine use
INTERNATIONAL COMPARISONS: Is the method used in soil monitoring schemes elsewhere?	0 = no 1 = yes
UK INFRASTRUCTURE: What is the state of the UK infrastructure to realise large-scale monitoring programmes using this method?	1 = none/few specialised labs 2 = moderate infrastructure 3 = ubiquitous infrastructure

Table 5 - Weighting factors adopted in the methodological tier in the logical sieve, with associated questions used in the consultation process to establish the final values adopted.

Criterion no.	CATEGORY: *Question*[1]	WEIGHT
1	THROUGHPUT: *How important is it to be able to have a high level of throughput (i.e. 100's per week) e.g.* • *0 = not important - dismiss* • *1 = relevant but not essential* • *2 = valuable but not essential* • *3 = valuable and preferred* • *4 = vital!*	3
2	STORAGE: *How important is it to be able to store samples until they can be analysed, for up to 2 weeks post sampling?*	3
3	ARCHIVABILITY: *How important do you consider archiving of samples (or analytical products) e.g. for future monitoring comparisons or for currently unknown analyses to answer new questions?*	2
4	SAMPLE COLLECTION: *Does it matter that the site would need to be visited more than once for a particular method to get the data?*	3
5	HOW MUCH SOIL: *Smaller soil samples cost less, easier to sample and handle etc; is a smaller sample preferred?*	2
6	COST – HARDWARE: *Does it matter how much it costs, in terms of hardware, to analyse the soil?*	1
7	COST- LABOUR: *Does it matter how much it costs, in terms of people, to analyse the soil?*	3
8	EASE OF USE: *Is it important that the method is relatively easy to carry out?*	2
9	POTENTIAL REFERENCE MATERIAL: *How important is quality control (QC) via reference material?*	2
10	REPRODUCIBILITY OF RESULTS: *How much do you care about being able to reproduce the same results time after time?*	4
11	READY-TO-USE DEPLOYMENT STATUS: *Is it important that the method is well established and has standard operating procedures?*	0
12	INTERNATIONAL COMPARISONS: *If the method is used in soil monitoring schemes elsewhere, is this important for UK soil monitoring?*	2
13	UK INFRASTRUCTURE: *Is it important that we have the capacity at present to deliver this method?*	3

[1]What *weighting* would you assign to the criterion when considering a trans-UK (cross-habitat) measuring and monitoring programme? Weight from 0 (i.e. dismiss entirely) to 4 (maximum relative weight).

Table 6 - Consolidated listing of distinct indicators using combined F_{SF}, ranked according to F_A, categorised according to deployment status[1].

Indicator	Indicator descriptor	F_A	Sub-cat. 1	Sub-cat. 2	Ref #
(a) Deployment status = 2. Cut off point $F_A > 100$.					
TRFLP - Ammonia oxidisers/denitrifiers	Genetic profile - specific group	769	Genotype	Nucleic acid	115
PLFA profiles	Composition -total community	615	Phenotype	Biomarker	18
TRFLP - ITS fungal	Genetic profile - specific group	437	Genotype	Nucleic acid	118
Multiple substrate induced respiration (MSIR) GC	Activity capability profile - total community	311	Function	Activity	158
Nematode Baermann extraction procedure	Numbers, composition and size of nematode community	302	Phenotype	Fauna	52
TRFLP - Bacteria	Genetic profile - specific group	295	Genotype	Nucleic acid	117
Microarthropods Tullgren dry extraction	Numbers, composition and size of invertebrates community within soil	188	Phenotype	Fauna	50
On site visual recording - flora and fauna	Numbers estimate of animals	173	Phenotype	Other	162
Microplate fluorometric assay - multi-enzyme	Enzyme potential activity - wide range	172	Function	Enzyme	30
TRFLP - Archaea	Genetic profile - specific group	146	Genotype	Nucleic acid	116
TRFLP - Methanogens/ methanotrophs	Genetic profile - specific group	123	Genotype	Nucleic acid	122
Invertebrates Pitfall traps	Numbers, composition and size of invertebrates motile aboveground	123	Phenotype	Fauna	46
TRFLP - Actinomycetes	Genetic profile - specific group	121	Genotype	Nucleic acid	113
(b) Deployment status = 1. Cutoff point $F_A>100$					
TRFLP - Nematodes	Genetic profile	437	Genotype	Nucleic acid	119
Multiple substrate induced respiration (MSIR) MicroResp	Activity capability profile	313	Function	Activity	160
TRFLP - Protozoa	Genetic profile	291	Genotype	Nucleic acid	120
qPCR AM Fungi	Genetic profile	111	Genotype	Nucleic acid	92
(c) Deployment status = 0. Cutoff $F_A>50$					
Functional gene arrays	Genetic profile	788	Genotype	Nucleic acid	84
Phylogentic gene arrays	Genetic profile	511	Genotype	Nucleic acid	91
FISH - keystone species	Genetic profile	138	Genotype	Nucleic acid	83
Soil proteomics	Phenotypic profile	51	Phenotype	Other	108

[1]Deployment status defined, as at mid-2005, as follows: 2 = fully deployable; 1 = likely to be ready for deployment in the short-term; 0 = not ready, some years development still needed.

Table 7 - Cross-reference matrix of consolidated list of top-ranking biological indicators against ecological processes and properties associated with each soil function.

	FOOD AND FIBRE PRODUCTION													ENVIRONMENTAL INTERACTIONS								HABITAT AND BIODIVERSITY SUPPORT						
Indicator	C cycling	Organic matter decomposition	N cycling	P cycling	S cycling	N fixation	Primary (microbial) activity	Soil food web transfers	Pest or disease transmission or suppression	Nutrient supply from symbioses	Redistribution by bioturbation	Bio-aggregation of soil	Degradation/ immobilisation pollutants	C retention/release	N retention/release	P retention/release	Tolerance/Resistance (toxins)	S retention/release	Redistribution by bioturbation	Bio-aggregation of soil	Habitat for rare soil species	Germination zone for plants	Nutrient supply from symbioses	Food source (aboveground)	Biodiversity reservoir (taxonomic)	Biodiversity reservoir (genetic)	Biodiversity reservoir (functional)
DEPLOYMENT STATUS = 2																											
TRFLP - Ammonia oxidisers/denitrifiers			x												x										x	x	x
PLFA profiles	x		x				x			x				x	x							x	x		x		x
TRFLP - ITS fungal	x	x		x					x	x		x	x	x		x	x			x	x	x	x	x	x	x	x
Multiple substrate induced respiration (MSIR) GC	x	x					x						x				x										x
Nematode Baermann extraction procedure							x	x	x								x					x			x	x	x
TRFLP - Bacteria	x	x	x	x	x							x	x	x	x	x	x	x		x		x			x	x	x
Microarthropods Tullgren dry extraction		x						x						x			x				x			x	x		x
On site visual recording - flora and fauna											x	x							x	x	x			x	x		x
Microplate fluorometric assay - multi-enzyme	x	x	x	x	x								x	x	x	x	x	x				x					x
TRFLP - Archaea														x											x	x	x
TRFLP - Methanogens/methanotrophs														x												x	x
Invertebrates pitfall traps		x						x	x								x				x			x	x		x
TRFLP - Actinomycetes	x	x											x	x			x								x	x	x
DEPLOYMENT STATUS = 1																											
TRFLP - Nematodes							x	x	x								x								x	x	x
Multiple substrate induced respiration (MSIR) MicroResp	x	x					x						x	x			x										x
TRFLP - Protozoa							x	x	x								x								x	x	x
qPCR AM Fungi				x					x	x		x				x	x			x		x	x		x	x	x

Supplementary material

Table S1. Collated list of potential biological indicators of soil quality considered in the ranking exercise, listed alphabetically by name of indicator, with associated descriptors and assigned categories. Ref# is unique code for reference purposes and database management only.

Table S2. Potential biological indicators ranked according to F_A (aggregated factor) score. Where equal scores are manifest, grouped by Sub-category 1, then alphabetical by indicator.

Table S3. Consolidated listing of distinct indicators for 'food and fibre production' function (F_{FF}), ranked according to aggregated factor (F_A), sieved for deployment status =2, cutoff point $F_A > 100$.

Table S4. Consolidated listing of distinct indicators for 'habitat and biodiversity' function (F_{HB}), ranked according to aggregated factor (F_A), sieved for deployment status =2, cutoff point $F_A > 100$.

Table S5. Consolidated listing of distinct indicators for 'environmental interactions' function (F_{EI}), ranked according to aggregated factor (F_A), sieved for deployment status =2, cutoff point $F_A > 100$.