Received: 4 July 2024 | Revised: 9 August 2024 | Accepted: 16 August 2024

DOI: 10.1111/ejss.13570

OPINION

European Journal of Soil Science WILEY

Five decades' experience of long-term soil monitoring, and key design principles, to assist the EU soil health mission

David A. Robinson[1] | **Laura Bentley**[1] | **Laurence Jones**[1] | **Chris Feeney**[1] | **Angus Garbutt**[1] | **Susan Tandy**[1] | **Inma Lebron**[1] | **Amy Thomas**[1] | **Sabine Reinsch**[1] | **Lisa Norton**[2] | **Lindsay Maskell**[2] | **Claire Wood**[2] | **Pete Henrys**[2] | **Susan Jarvis**[2] | **Simon Smart**[2] | **Aidan Keith**[2] | **Fiona Seaton**[2] | **James Skates**[3] | **Suzanne Higgins**[4] | **Giovanna Giuffrè**[5] | **Bridget A. Emmett**[1]

[1]UK Centre for Ecology & Hydrology, Bangor, UK

[2]UK Centre for Ecology & Hydrology, Lancaster Environment Centre, Lancaster, UK

[3]Welsh Government, Conwy, UK

[4]Agri-Food and Biosciences Institute (AFBI), Agri-Environment Branch, Belfast, UK

[5]ISINNOVA, Institute of Studies for the Integration of Systems, Rome, Italy

Correspondence
David A. Robinson, UK Centre for Ecology & Hydrology, ECW Deiniol Road, Bangor, UK.
Email: david.robinson@ceh.ac.uk

Funding information
HORIZON EUROPE Framework Programme, Grant/Award Number: 101086179; Natural Environment Research Council, Grant/Award Number: NE/R016429/1; Llywodraeth Cymru, Grant/Award Number: C147/2010/11; UK Research and Innovation, Grant/Award Number: 10053484

Abstract

The European Union has a long-term objective to achieve healthy soils by 2050. The European Commission has proposed a Directive of the European Parliament and of the Council on Soil Monitoring and Resilience (Soil Monitoring Law, SML), the first stage of which is to focus on setting up a soil monitoring framework and assessing soils throughout the EU. Situated in NW Europe, the UK has substantial experience in soil monitoring over the last half century which may usefully contribute to this wider EU effort. A set of overarching principles have and continue to guide design of national soil monitoring and may prove helpful as other European countries embark on similar monitoring programmes. Therefore, we present the principles of design from five decades of national soil monitoring. The monitoring discussed is based on a stratified-random design, has matured in support of policy questions, and operates over space and time scales relevant to the SML. The UK Centre for Ecology & Hydrology (UKCEH) Countryside Surveys (CS) of Great Britain and Northern Ireland, Welsh Government, Environment and Rural Affairs Monitoring and Modelling Programme (ERAMMP) and the England Ecosystem Survey (EES) monitoring programme are national programmes currently operating in the UK. Some important lessons learnt include: adopting a question-based approach; having a clear robust statistical design for the purpose; selecting indicators that address policy and underlying scientific questions; and selecting indicators that can detect change and use robust and well-tested methodologies across a wide range of soil and land use types, remaining valid over long time scales, supporting thinking long-term. Technical lessons learned include the proven cost effectiveness of a stratified-random design including replication, while adopting a common stratification layer of stable environmental attributes aids comparability between monitoring programmes. Common protocols are vital for future intercomparisons, but a full ecosystem approach that includes co-located soil and vegetation samples for interpreting a co-evolving system has proved hugely advantageous. UK monitoring programmes offer a range of experience that may prove valuable to future soil monitoring design to address the major societal challenges of our time, such as maintaining food production and addressing climate change and biodiversity loss.

KEYWORDS

countryside survey, ERAMMP, function, indicator, land class, mission, soil monitoring law, soil science, stratification, stratified random

1 | INTRODUCTION

The main goal of the Mission 'A Soil Deal for Europe' is to establish 100 Living Labs and Lighthouses to lead the transition towards healthy soils by 2030 and represents an investment of ~1 billion euros (European Commission Directorate-General for Research and Innovation et al., 2020; Panagos, Borrelli, et al., 2024). With the EU Soil Strategy for 2030 (COM/2021/699), (European_Commission, 2021), the European Commission (EC) has proposed a soil monitoring law (SML) laying down objectives for the protection, restoration and sustainable use of soil in view of achieving healthy soils in Europe by the year 2050 (COM/2023/416) (European_Commission, 2023). This requires action on the ground to improve soil health which can be supported by identifying soil degradation and implementing control measures or interventions to address and reverse this. While some soil degradation is visually obvious, namely, some forms of soil erosion, much is not, for example, soil carbon degradation, biodiversity loss, acidification and some forms of contamination. Measurements from a soil monitoring programme develop an evidence base that enables all stakeholders (land managers, policy makers, citizens) to determine the status of soils and see how they are changing through time. This ensures that the success of policy as well as private initiatives and actions can be tracked, and allows policy development and control measures to be co-designed and then implemented at an appropriate level.

The impact assessment for the SML (COM/2023/416) states that, 'policy options have been described by using five key building blocks: (1) definition of soil health and establishment of soil districts, (2) monitoring of soil health, (3) sustainable soil management, (4) identification, registration, investigation and assessment of contaminated sites, (5) restoration (regeneration) of soil health and remediation of contaminated sites'. The legal text is now subject to the legislative co-decision procedure, which requires negotiation and joint adoption by the Council of Ministers and the European Parliament. To date, amendments to the EC proposal have been advanced by the European Parliament resolution of April 2024 and by the Council (Environment) approval of the SML general approach in June 2024. It is likely that some of the terminology will change during this process. However, the focus of the SML is that a sampling scheme should be designed and implemented to assess soil health. As member states consider designs it is important to refer to the literature where the advantages and disadvantages of different designs are discussed in detail (D. Brus, 2014; D. J. Brus, 2022; Gruijter et al., 2006).

Highlights

- Countryside Survey offers 5 decades of soil monitoring experience.
- Principles supporting soil monitoring to inform policy are provided.
- Robust soil monitoring methodologies are described.
- An example of a soil indicator specification table is presented.

The purpose of this article is to reflect on the implementation and practical experience of conducting soil monitoring in support of policy gained over the past five decades in the UK. EU member states may draw on this experience as they move forward with long-term soil monitoring of their own. The UK has been operating monitoring programmes that have evolved over this period and continue to evolve providing clear evidence of policy success and areas requiring continued action. Here, we describe this effort, the overarching design principles, and lessons learned from five decades of practical experience that may assist those nations in the EU and further afield that are embarking on long-term soil monitoring. Monitoring is an important investment and can have substantial inertia once implemented, so choosing an appropriately flexible design is desirable. The principles outlined are not unique to the UK, but the five decades of implementation of co-located soil and vegetation monitoring provide a valuable lens through which to consider their relevance and importance.

2 | UK STRATIFIED RANDOM SOIL MONITORING

How the environment is changing at a national scale is of interest to many stakeholders across society. Government departments in the UK historically commissioned a variety of actors to provide advice on the state and change of the UK environment (Bunce, Barr, Clarke, et al., 1996). In the mid 1970s there was a need to produce a standardised procedure for ecological monitoring (Bunce & Shaw, 1973). From this, the UK Centre for Ecology & Hydrology (UKCEH) Countryside Survey (CS) evolved with its first sampling completed across Great Britain (GB) in 1978, which continues to the present day (Wood et al., 2018). The underpinning design of CS was based on a stratified sampling approach to ensure the sample was representative of the entire country (Wood, 2011). Strata were defined based on a multivariate analysis of key attributes that do not change quickly, for example, relief, parent material and climate; these strata were called land classes. The survey focussed on structured sampling of vegetation and soils that were co-located along with other variables. Co-location was used as the soil–plant system was considered to co-evolve, and the co-location offered both greater power for integration and cost-effectiveness. Units of 1 × 1 km squares were chosen as the fundamental basic sampling unit. Within a sampling unit, 5 randomly located soil and vegetation sampling sites were nested to provide replication. The initial survey conducted ecosystem monitoring (covering both soils and vegetation) in a sample of 256 1-km squares across GB, based on 8 samples per strata and 32 strata. Monitoring was subsequently undertaken in 1984 and 1990, but it was not until 1998 that soils were resampled. The number of soil samples dramatically increased in 2007, according to policy need, when 591 1-km squares were sampled. Part of the reason for this increase was the policy interest in Scotland and Wales. In 1978, the only soil metrics recorded were soil organic matter (SOM) content and pH. As interest grew, by 2007 a suite of soil indicators for pH and nutrients (nitrogen [N], carbon [C], phosphorus [P]) (Reynolds et al., 2013), pollutants (POP's, heavy metals) and biodiversity (mesofauna, metabarcoding) (Griffiths et al., 2011) were measured across GB, including process measurements such as the potential for N mineralisation (Rowe et al., 2012) and basal respiration (Simfukwe et al., 2011). With the development of devolved administrations, with authority over environmental management, new country-specific monitoring programmes have evolved in the past decade. Since 2007, the design was again modified into a rolling programme to help account for annual variability and assist with the logistics of conducting the monitoring. Since 2019, about 100 squares have been sampled each year with a complete roll comprising 500 squares over 5 years. This more flexible approach proved invaluable during the Covid disruption, and ensured that monitoring could be maintained during this time.

Four soil monitoring schemes are currently in operation and collecting data across the UK that are based on a stratified random design (Table 1). These are the (1) Countryside Survey for GB (1978–present) and (2) CS

TABLE 1 UK devolved administrations and their respective soil monitoring programmes; where NUTS is Nomenclature of territorial units for statistics used by the European Union.

Administrative unity responsible for natural resources	NUTS level	Monitoring programme	Reference
Great Britain (UK)	NUTS 0	Countryside Survey (CS GB)	Countryside_Survey (2024)
Northern Ireland (UK)	NUTS 1	Countryside Survey (NICS)	Countryside_Survey_NI (2024)
Wales	NUTS 1	Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)	ERAMMP (2024)
England	NUTS 1, 9-units aggregated	England Ecosystem Survey (EES), part of Natural Capital and Ecosystem Assessment (NCEA)	NCEA (2024)
Scotland	NUTS 1	In development	

Northern Ireland (1986–present), with soils being collected for the first time in the current survey (2023–2024), (3) the Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP) in Wales (2013–present), and (4) the England Ecosystem Survey (EES) (2023–present); Scotland is currently in a new design phase of a soil monitoring scheme that encompasses specific soil, land use and climatic conditions found there that will be responsive to emerging policy and management priorities. Other monitoring programmes for soils have been implemented in the UK and include grid-based designs such as the National Soil Inventory (NSI) (Kirk et al., 2010) and the National Soil Inventory of Scotland (NSIS) (Chapman et al., 2013; Lilly et al., 2011); in Northern Ireland the Soil Nutrient Health Scheme (SNHS) is sampling every agricultural field in Northern Ireland over a 4 year period (2022–2026). In this article, the focus is on experiences with running long-term monitoring based on stratified design, due to it being highlighted as the preferred methodology in the impact assessment for the SML (COM/2023/416).

All of these monitoring programmes evolved for the same reason as the proposed EU programme, based on a need to understand the state and change of the environment subject to anthropogenic drivers and management pressures, and identify locations for intervention. Hence, they are all designed within administrative boundaries, where the administrative authority has the need to monitor performance and the power to implement sustainable land management practices including soil conservation measures. Figure 1 displays the four UK administrations, Wales, Scotland and Northern Ireland; these administrations have responsibility for agricultural and environmental management.

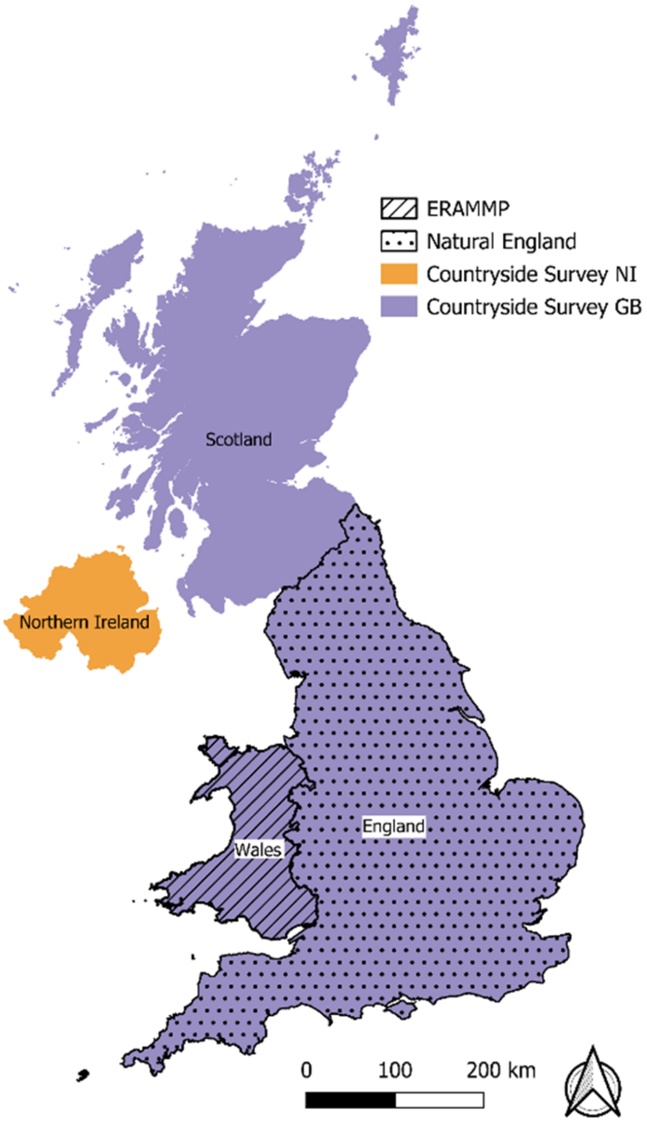

FIGURE 1 The extent of currently active UK Soil monitoring programmes. The UK was split into 12 NUTS1 regions prior to the UK leaving the EU. Wales, Northern Ireland and Scotland were 1 region each, whereas England comprised 9 regions. (1) CS GB covers England, Wales and Scotland (purple), (2) NICS in Northern Ireland, (3) ERAMMP in Wales (black hashes) and (4) England Ecosystem Survey (black dots).

3 | PRINCIPLES SUPPORTING STRATIFIED RANDOM SOIL MONITORING

Countryside Survey has always had a mixed purpose that provides both the underpinning data for scientific research regarding soil and vegetation change in response to drivers and answering policy questions. ERAMMP, the EES and NICS are a direct consequence of administrations requiring monitoring in support of policy. All the monitoring programmes were designed with a set of guiding principles in mind (Table 2). Establishment of principles aids in the development of integration, much like the F.A.I.R principles (findability, accessibility, interoperability and reusability) for scientific data management and stewardship (Wilkinson et al., 2016). The following sections describe how these principles have been applied in UK monitoring and may serve to inform the future design of soil or environmental monitoring programmes more widely.

3.1 | Purpose

Common to all monitoring was the general purpose to evaluate environmental change over time, in response to environmental and anthropogenic drivers and pressures. However, the more specific purposes of monitoring have evolved over time, creating an inevitable tension between the long-term utility of the programme and short-term objectives that must be balanced. The early goals of CS were to provide advice to policy on the state and change of the UK environment (Bunce, Barr, Clarke, et al., 1996; Bunce, Barr, Gillespie, et al., 1996), in a standardised way (Bunce &

TABLE 2 Overarching principles underpinning monitoring design.

Principles	Description
Objectives and ways of working	
Purpose	Clearly define the purpose of the monitoring system and the questions it will address.
Adaptability	Build flexibility into the monitoring system to accommodate changing circumstances, evolving priorities and new information. Periodically review and update indicators and methods as needed.
Accessibility and transparency	Make monitoring data accessible to relevant stakeholders and ensure transparency in the monitoring process, including data collection methods, analytical procedures and reporting mechanisms.
Ethical considerations	Consider ethical implications related to privacy, confidentiality and informed consent when collecting and using monitoring data. This is particularly important regarding the collection and handling of soil samples and information from private land.
Timeliness	Monitoring must capture essential time scales of the processes concerned to ensure decision making is supported and potential intervention feasible. Consider the needs of stakeholders regarding the frequency of reporting.
Sampling design	
Functional reporting units	Choose the units that will be used to report. In the UK, land cover or habitat are most common, soil types are feasible.
Design, stratification and allocation	Use an appropriate statistical design. The chosen design should enable the quantification of the relationship between the sample and the population of interest, and hence, allow robust, unbiassed inference. Static, dynamic or rotational designs offer different advantages and disadvantages (see e.g., Gruijter et al., 2006).
Cost effectiveness	Design monitoring systems that are cost-effective and efficient in terms of resource allocation, data collection methods, and analysis procedures. Balance the costs of monitoring with the benefits of improved decision making and accountability.
Indicator selection	
Sensitivity, specificity	Indicators must be sensitive to detect change and, yet, specific enough to capture the desired outcomes without being overly broad or ambiguous. Supporting metrics, such as soil texture, tend to be invariant to change but may also be required to aid in the interpretation of indicators or generate composite indicators.
Measurability	Identify indicators that are measurable and quantifiable. In the case of soils, a basket of indicators is often necessary to address the range of questions that sustainable land management poses.
Targeted indicator selection	Select indicators that will allow the questions to be addressed and can link through where possible to functions and ecosystem services.
Validity and reliability	Select indicators that are valid and reliable measures of both state and change that will be consistent regarding measurement over space and time, and across different observers.

Shaw, 1973). This has evolved for subsequent monitoring programmes, such that the objective of ERAMMP is to, 'Collect data from across the Welsh landscape and employ models to assess change and future impact, while providing the evidence and insight the Welsh Government needs in order to advance effective policies that build social, economic and environmental resilience'. Hence, the following underpinning considerations are important:

3.1.1 | Underlying process time scales

When designing sampling intensity and frequency, consideration must be given to drivers and pressures and the space and time scales over which they operate. In addition, the time needed for subsequent responses to take effect and demonstrate impact should inform resurvey frequency. Essentially the monitoring should encapsulate what is now described by the Driver, Pressure, State, Impact, Response (DPSIR) cycle (OECD, 1993).

3.1.2 | Think long-term

It is important to select and maintain a suite of core indicators that can provide historic compatibility as technology progresses for 40–50 years plus. Thinking long-term is also vital to capture the DPSIR processes time scale. For example, degradation of soil organic matter (SOM) may be seen within years when grassland is converted to cropland, reaching a new equilibrium in ~20 years. However, SOM accumulation and restoration following reversion is a much slower process and may take up to a century (Or et al., 2021). CS has been running for nearly 50 years and SOM decline is apparent in cropland systems, but the impact of interventions over the last decade is only beginning to emerge in recent monitoring (Bentley et al., 2024). Such long-term processes can prove challenging to stakeholders thinking on policy cycles, so it is important to inform stakeholder expectations regarding change, something that can often be achieved with modelling.

3.1.3 | Question-based approach

Monitoring and its design should be specific enough to address distinct science or policy questions, but flexible enough to ensure new and emerging issues can be addressed. As an example, in the 2007 cycle of funding for CS, the UK government requested that the sampling design be checked to ensure a simple yes/no response would be provided regarding 10 specific policy questions. This approach helped focus resources most efficiently and ensured maximum impact within policy circles. Examples of some of those questions are:

- Can we confirm the loss of soil carbon (0–15 cm) as reported by Bellamy et al. (2005)?
- Has the recovery from acidification detected by Countryside Survey in 1998, that occurred between 1978 and 1998, continued?
- Can the trend of eutrophication of the countryside detected in the vegetation be detected in the soil using the mean total nitrogen concentration?
- Can the trend of increasing P status in intensive grasslands be confirmed and is it matched in other habitats?
- Is the decline in atmospheric deposition of heavy metals as reported by the Heavy Metals Monitoring Network reflected in soil metal concentrations measured in Countryside Survey?

There are efficiencies if a single monitoring programme can be designed with common indicators for both national reporting, but also more specific national or EU policy outcomes for example, the common agricultural policy payment scheme. This has been demonstrated as effective in Wales (ERAMMP), where a single common monitoring programme, using a common sampling design and set of indicators, allows for both national trend reporting and the impact of agri-environment schemes (Emmett et al., 2017; Emmett & Wales_AXIS_II_GMEP_Team, 2013); the NICS also seeks to achieve similar goals. Moreover, the platform should be capable of addressing a range of questions surrounding land use; management of productivity and farming, including the food chain and human and animal health; climate change including mitigation and weather extremes such as flood, heat and drought; pollution including contaminants, water quality and air quality; and ecosystem health and biodiversity. With regards to challenges such as Net Zero Plus, the question of what works where is fundamental to finding solutions.

3.2 | Adaptability and flexibility

New challenges and reporting requirements will come up during the lifetime of a monitoring scheme. This is certainly our experience with CS, where there were requests from devolved administrations to increase the number of reporting units. Fortunately, this could be addressed by the addition of extra squares within the sampling framework, maintaining statistical robustness. The repeat sampling of sampling units (or at least a subset of) is essential for time series analysis, enabling the separation of spatial variation from temporal variation. Moreover, it is important to review both the design and the indicators periodically. For example, with the new CS rolling programme design, some indicators are measured once (e.g., texture), most on the rolling cycle of 5 years (e.g., SOM and pH) and some on a decadal cycle or more (e.g., Total P).

3.3 | Accessibility and transparency

This is important to ensure the validity of the monitoring, such that all stakeholders can see, and challenge, if necessary, the results. The UK monitoring designs are all open and documented, and the methods and procedures are also documented, such as the soil methods (Emmett et al., 2008); however, sample locations are confidential as discussed later. Data gathering and analysis follows a rigorous quality assurance procedure. This includes barcoding samples in the field and the digital GPS collection of locations sampled. In the laboratories, all batches of samples are analysed with inclusion of standard soils to track accuracy and precision over time. Moreover, UKCEH labs are signed up to schemes like WEPAL (Van Vuuren et al., 2002) so that metrics such as particle size analysis can be compared across European laboratories to check performance. The use of standardised data processing and quality assurance scripts, in a programming language such as R (R_Core_Team, 2021), are helpful for keeping track of exactly how the data was processed ensuring repeatability and efficiency. We have found that it is best to avoid the use of spreadsheets where possible as copy-paste is not traceable. Consideration must also be given to how records are preserved, particularly where documents and processes may need to be revisited decades down the line, by different individuals.

3.3.1 | Accessibility of locations

All locations are kept confidential which serves two purposes: (i) to maintain the scientific rigour of the sampling. Even the landowners are not told the location and results from the sample points. It is possible that by knowing the locations and results, landowners might implement changes in management which would bias the survey; and (ii) to avoid undue burden on landowners dealing with requests from researchers seeking to experiment at sample locations. Feedback is provided to landowners, but only mean values are provided at a square level. This does produce an impediment on the sharing of data for scientific purposes, such that only obscured locations are released to the public; newer programmes are adopting alternative strategies, for example, limited licensing ERAMMP, or fully open, which is proposed for EES, but this does raise concerns regarding potential unknown bias.

3.4 | Ethical considerations

Privacy is a particularly important consideration when accessing private land. All monitoring is conducted through a permissions system and the consent of the landowner. Landowners are initially contacted by letter and then called on the phone in the day, or days, prior to a visit. As part of the permissions process, CS does not allow data to be used in any regulatory context. As a result, the programme has worked well with landowners, gaining permission from more than 90% of landowners.

3.5 | Timeliness

The monitoring must capture two aspects regarding timeliness: (i) that of the essential time scales of the processes concerned to ensure decision making is supported, and potential intervention responses captured; and (ii) consider the needs of stakeholders regarding the frequency of reporting. CS initially operated on a decadal time scale, but this represents logistical, consistency and training challenges, in addition to the data gap generated between reporting periods. Subsequent designs such as ERAMMP, EES and the latest CS adopted an annual rolling survey, designed to capture a full set of squares within a policy cycle that in addition allowed the capture of annual variability of process relevance. CS now runs on a 5-year return cycle with 100 squares measured annually and ~500 squares measured in a complete roll, which contains ~2500 soil samples in total.

3.5.1 | Return time and co-location

Monitoring designs can take account of this to optimise analytical power through the co-location of sampling and structuring of return time—the length between repeat visits to the same site. Return time is dependent on the rate of change in the process of interest and its respective indicator and what degree of change is policy relevant or detectable given practical limitations. Co-location of indicators can give greater power to detect effects across indicators that vary with spatial or temporal lags. The co-location reduces the variability due to differences in space or time and hence increases the power. This has proven vital to the unpicking of drivers of change and plant–soil feedbacks in a UK context, for example Seaton et al. (2023).

Soils change on a range of time scales (Richter & Markewitz, 2001). For monitoring the impact of human activity, change can be considered on short- (months), medium- (years) and long-term (decades) monitoring time scales. In the short-term (monthly), soil health changes in response to wetting and drying cycles as well as management activities such as tillage. On a seasonal timescale, soil health changes are likely to follow seasonal climate trends, affecting plant root carbon inputs, leaching, redox conditions and shrink swell. Moreover, management impacts of tillage, organic matter addition and fertilisation are likely to be evident in agricultural settings. On the longer-term (5+ years) soil change may be largely driven by a mixture of land use change, management, pollution and climate change.

Taking soil carbon as an example, grassland to cropland conversions typically remove 1–2 t C ha^{-1} $year^{-1}$. The average topsoil (0–15 cm) carbon density in English grasslands in 2007 was 64.6 t ha^{-1} $year^{-1}$ and 46.9 t ha^{-1} $year^{-1}$ in cropland soils (Emmett et al., 2010). Hence, rates of carbon loss of 1.5%–3% per year of the original carbon stock are to be expected for grassland to cropland conversion. This is consistent with the findings of others, for example those reviewed in Or et al. (2021) indicating a loss of ~2% $year^{-1}$ from grasslands, with a new cropland equilibrium reached after 10–20 years. This is also consistent with the reported change across EU + UK with the LUCAS survey (De Rosa et al., 2024). Current recommendations in the SML are consistent with this resampling approach, which is consistent with the time frames of soil change for informing policy.

4 | SAMPLING DESIGN

4.1 | Functional reporting units

Analysis of results can be carried out on many different reporting units if they have adequate coverage in the sample and the relationship between the sample structure and reporting unit is known. For example, if a survey of farms was stratified based on soil classification and we wanted to report by crop classes, then both sufficient coverage of the crop classes and knowledge of how the different soil classifications relate to crop classes are needed to produce unbiassed results. Land uses and habitats are the main reporting unit in the UK, with combinations of soil type and habitat now also being used to benchmark land (Feeney et al., 2023), the reason being that land is managed by habitat in general, rather than by soils per se. This also links to international approaches such as natural capital accounting at the U.N. System of Environmental and Economic Accounts (SEEA) (United-Nations, 2014) which the EU contributes to. All the monitoring programmes are capable of reporting by soil type if required. However, land cover and soil properties should typically not be included as high-level stratum in design as they change too much on decadal time scales. Having strata that are subject to change over time significantly affects the ability to achieve consistent balance in the sample and to produce robust estimates or change.

4.2 | Design and stratification

A robust, structured, statistical design is required to allow the reporting of state and change and to enable policy questions to be suitably addressed. In the Countryside Survey, the core principles of this have solidified into four themes: a stratified random approach; the selection of strata for monitoring;

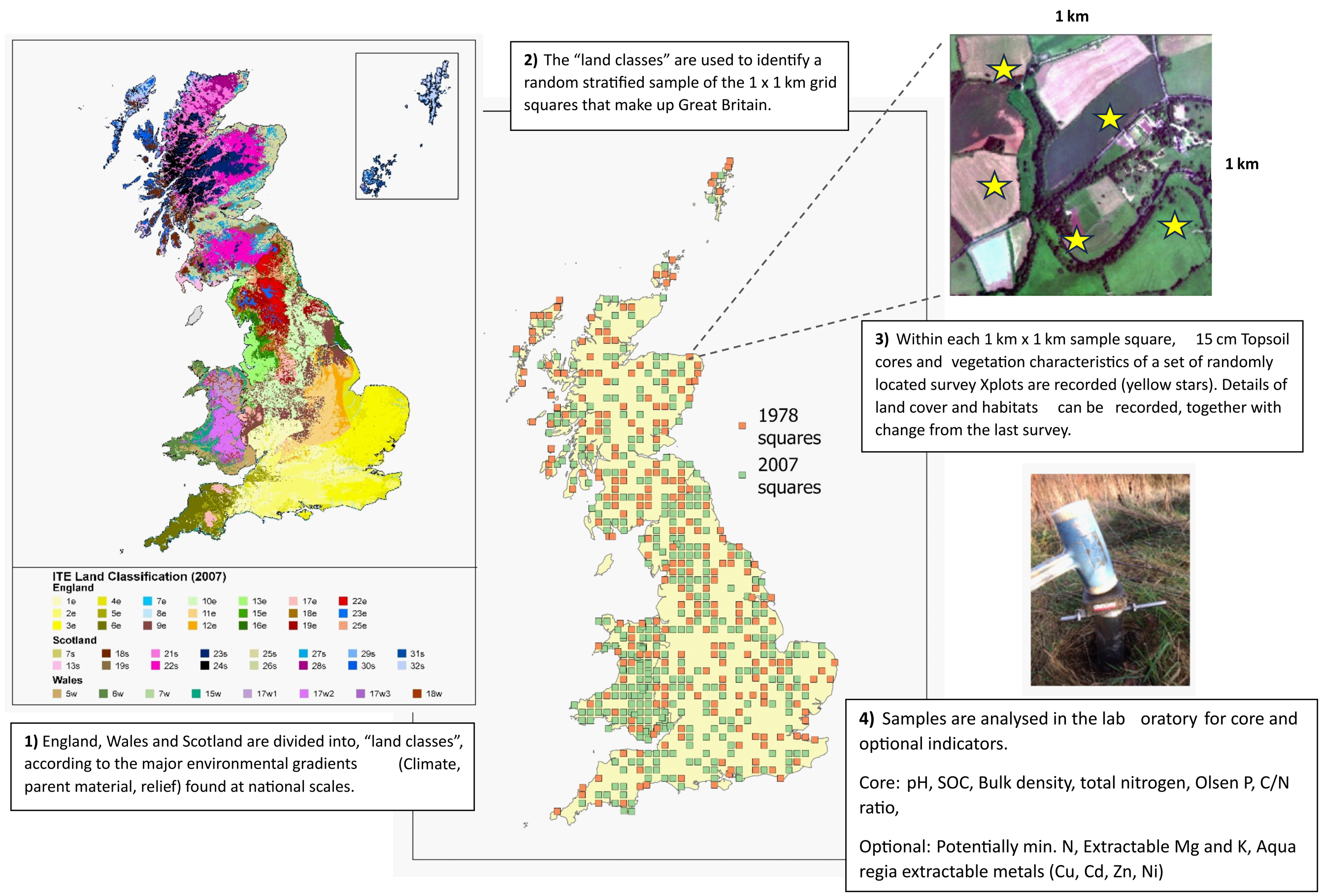

FIGURE 2 The overarching CS design, from strata to random stratified monitoring squares where samples are taken to have indicators measured. In 1978, 256 squares were selected for sampling, this increased in subsequent monitoring to a maximum of 591 in 2007, with the current rolling programme ~500.

indicator selection; and the selection of reporting units. The overarching design of CS GB is illustrated in Figure 2. The land classes (a classification of an ordination based on relief, parent material and climate) provide the strata for sampling. In the 1970s, Great Britain was overlain by a 15 × 15 km grid to generate strata using a multi-variate approach. The grid spaced samples but was used primarily to overcome the computational limits at the time in generating strata (Wood, 2011). Then 1 km sampling units sitting at the intersection points of this grid were randomly selected within the strata, thus the design is a gridded, stratified random sample (Wood, 2011), which we simply refer to as stratified random sampling throughout as the gridded element is not deemed to be essential and is not used in the Countryside Survey in Northern Ireland. The 1 km squares form the basic sampling unit. Within each square, 5 randomly located, co-located soil and vegetation samples are taken or recorded so that indicators can be measured; this aids replication.

4.2.1 | Stratified-random approach

All the described monitoring schemes use a stratified-random approach for the underlying design, rather than using simple random sampling, or a systematic grid approach. The key advantages of the stratified-random approach are, (1) ensuring sufficient coverage of the strata, which is not guaranteed with random sampling alone, hence, efficiently providing representation of the entire population of interest with the necessary reporting power, (2) enabling different allocation of resources across the different strata, (3) to be able to robustly scale results, accounting for the observed variation, and (4) the ability to compare like with like across administrative boundaries should it be required. To support this, there are helpful texts that can assist in the development of soil monitoring design that could be implemented within soil districts in addition to those described in this article, for example, D. Brus (2014); D. J. Brus (2022); and Gruijter et al. (2006).

Monitoring based on simple random sampling or a grid-based approach could allow sufficient power to detect the effects of specific actions or policies, but this would likely require more samples (Black et al., 2008). Where characteristics and responses of interest are distributed heterogeneously, stratified random designs may be more precise and efficient than random alone, with the added benefit over fixed grid sampling of being able to expand and contract with need, as CS has done. The strata must be:

- mutually exclusive, non-overlapping groups such that a sample can only belong to one group,
- jointly exhaustive, namely, sampling covers the entire area to be monitored.

Given the strata, a simple random sample can be extracted from each stratum.

4.2.2 | Selection of strata for monitoring

Ideally, the strata should be defined in such a way as to minimise within-strata variation and maximise between-strata variation. The following criteria should be kept in mind when selecting strata: (1) strata should capture sources of variation that are meaningful to soils for the system being monitored, (2) they should not be constrained to administrative boundaries, (3) they should be stable over time. Our experience in CS demonstrates a significant turnover in landcover between sampling periods, justifying the decision not to use this as a basis for stratification. For example, of the plots that were visited in 1978 and revisited in 2007, 41% had a different habitat assignment, that is out of 846 plots resampled, which is much higher than the 1% considered in Black et al. (2008). This poses a challenge when planning strata for long-term reporting whilst having a sample that retains a balanced representation of the population, hence the use of land class for UK monitoring.

In the case of soils, it makes sense to choose some of the soil-forming factors for strata following the approach of Jenny (1994) (Climate, Organisms, Relief, Parent material, Time) that can be considered stable for the purposes of monitoring. For example, relief and stable geomorphological units (e.g., based on the USGS land surface forms classification), parent material with consistent types, and climate metrics that can be analysed.

The land class used in GB uses the above factors and other pertinent information including, for example, the location of rivers and lakes and long-term infrastructure. Often layers selected are co-correlated and so principal components analysis (PCA) is used to compress the variance before running the data through a classification procedure. This approach underpins CS GB, ERAMMP and ESS through the use of Land Class strata (Bunce, Barr, Gillespie, et al., 1996). In Northern Ireland, a similar land classification is used (Cooper, 1986). Moreover, the stratification approach underpins how ‘environmental zones and strata’ were created, for example, in the EU (Metzger et al., 2012). The advantage of such an approach is that the number of strata can be chosen such that the sample selected meets the precision and reporting requirements desired.

4.2.3 | Sample size

The number of samples, to provide the required sensitivity to change, is an important aspect of any monitoring programme. It should provide enough samples to ensure statistical validity and enough power for hypothesis testing. Power analysis can be used to determine the number of samples that need to be collected in relation to (i) a policy action soil indicator threshold, or (ii) a level of change over time that is to be detected. Both scenarios were considered, and demonstrated, in Black et al. (2008) for the design of a UK soil monitoring scheme. The number of samples to detect change varies by indicator, for example, the number of samples needed to detect change in pH is less than the number required to detect change in SOM.

4.2.4 | Implications for statistical inference

The chosen monitoring programme design also influences the appropriate methods for statistical inference of any given property. This necessitates careful consideration of the monitoring design, and potentially how this has changed over the lifetime of the monitoring programme, when inferring change over time in soil indicators. Few monitoring designs lend themselves to simple statistical inference, and more complex methods (e.g., weighting of means, random effect structures) are usually required to achieve unbiased estimates at the population level. For a description of the modelling approach used in CS, see the Supplementary information (Reynolds et al., 2013).

4.3 | Cost-effectiveness

Due consideration should be given to costs. For UK-based monitoring, about half the cost is collecting the samples and the other half is the laboratory analysis, archiving and data interpretation. Selection of indicators should always consider cost/benefit, as adopting the latest measurement technique may prove of limited use when this is assessed. Indirect indicators should always be explored and used if appropriate as they may be cost effective. Empirical and mechanistic models are widely used as a way of contextualising results, or for making future prediction. The JRC uses some indicators for the EU based on modelling, such as erosion by water, that are currently intractable to obtain through observation (Panagos et al., 2020). Such approaches are now being explored as part of reporting for UK monitoring. Different designs were analysed using a travelling salesman algorithm in Black et al. (2008) to determine efficiency and random stratified worked well. Hence, the rationale of a 1 km square as a fundamental monitoring unit containing replicate samples in the square serves to balance costs and statistical power.

5 | INDICATOR SELECTION

5.1 | Sensitivity, specificity

A fundamental aspect of indicators is their ability to detect change over a relevant time scale, indicators that show significant variability due to spatial, temporal, sampling or measurement errors should be avoided. They must be sensitive enough to pick up change and specific regarding change detection in the soil attribute of interest, but they must not be prone to high levels of noise such that it obscures the change signal. Indicators that fluctuate on a daily or weekly basis are not helpful for long-term monitoring. Moreover, indicators that are prone to short-range variation present a challenge. Indicators must discriminate the long-term trends from 'noisy' backgrounds (Merrington et al., 2006).

5.2 | Measurability

Indicators must be easily measurable and quantifiable such that they are widely reproducible across observers or laboratories. Hence, in the UK, many monitoring schemes use a 0–15 cm topsoil core which is easily collected and processed. The EES also includes 15–30 and 30–40 cm samples; however, such measurements take substantially more time. Both NSI(S) have soil profile data, as this was fundamental to the purpose of mapping soils rather than monitoring. In the case of soils, a set of indicators is often necessary to address soil multifunctionality (Bünemann et al., 2018) and the range of questions that the SML poses. Current UK indicator selection evolved from a rigorous identification and selection procedure, the framework for which was developed by the UK Soil Indicator Consortium (UKSIC), a cross-community working group (Bhogal et al., 2008; Loveland & Thompson, 2002; Merrington et al., 2006; Nicholson et al., 2008). The work undertaken by the UKSIC provided a pool of 13 high-level indicators (pH, SOC, bulk density, total nitrogen, Olsen P, C/N ratio, Potentially mineralisable N, Extractable Mg and K, Aqua regia extractable metals [Cu, Cd, Zn, Ni]), selected using criteria such as Relevance; Sensitivity, Discrimination and Signal-to-Noise Ratio; Measurability and Practicality; and Efficiency and Cost. The ambition is to link all indicators to soil functions and ultimately ecosystem service delivery and wider policy questions (Emmett, Bell, et al., 2023). Whilst there is generally a core of indicators that are always measured (e.g., Soil Organic Matter, pH, bulk density, nutrients) not all indicators are monitored all the time or at all locations. The indicators are selected based on their ability to address the policy or science question. This enables consistency in core reporting through time and across administrative boundaries, whilst providing flexibility to changing policy and reporting objectives. Alternative selection procedures such as the logical sieve (Corstanje et al., 2017; Stone et al., 2016) have also been used and continue to develop selection frameworks.

5.3 | Targeted indicator selection

Indicators relevant to addressing the questions must be selected; in the case of soils this is often from a set of available indicators (Bünemann et al., 2018). It is ideal if the indicator is focused on the specific issue to be addressed and must be easy to interpret and quantitative in manner. More often, regarding soils, a requirement is that it links to a specific soil function, threat, or ecosystem service (Baritz et al., 2021; Bünemann et al., 2018). Moreover, it must offer information that gives the strategic insight required for effective planning and coherent decision making. Moreover, it is important to remember that in general it cannot be envisaged what the policy questions will be in the future or what the next 'big thing' in soils will be, so a monitoring scheme needs to be sampling a sufficiently wide base of indicators to cover all potential questions.

The choice of indicator should also be chosen carefully when quantifying ecosystem service delivery or functional recovery. In a recent evidence review commissioned to inform the English Environmental Land Management Scheme, monitored by the EES (Table 1), over 740 types of land management were reviewed for their impact against 53 ecosystem service indicators, six of which were focused on soil health (Bentley et al., 2023; Emmett, Cosby, et al., 2023; Newell Price et al., 2023). The review found that there are very few win-win solutions for land management across the range of indicators, with the most effective managements identified as a priority for only three ecosystem service themes. This lack of consensus was present within the indicators of a given ecosystem service as well, where the indicator chosen had a dramatic impact on the assessment result. This highlights the importance of using a range of indicators, each targeted to specific questions and processes. The effect of management was also frequently identified as being context dependent (15% of soil indicator assessments), reinforcing the need for a robust underpinning sampling design to capture this variation.

5.4 | Validity and reliability

This is a fundamental criterion when considering long-term measurements. Some measurements have a distinct lifetime due to changes in measurement technology, and as such, may not prove useful for long-term monitoring. Simple, tried-and-tested metrics may not be the most glamourous but may stand the test of time for monitoring purposes. These include measures such as pH, EC, SOM (loss on ignition) and bulk density, for example, that have proved reliable over decades of use. Selecting biological indicators has proved challenging, partly because the metrics are still undergoing development, and partly because methods are changing so rapidly considering the DNA revolution. Methods may be suitable to obtain the state of a specific metric at one point in time but may be obsolete in 5–10 years or less.

Beyond the choice of indicators, minimising variability and ensuring reliability requires rigorous field practice and laboratory quality assurance processes, especially for large scale and long-term monitoring programmes where some staff, surveyor and equipment turnover are unavoidable. Within the ERAMMP and CS programmes, which are directly managed by the authors, a 2-week programme of dedicated surveyor training prior to each annual field season is conducted, covering both practical and theoretical considerations of the survey, irrespective of surveyor experience. All surveyor teams are also visited during the field season by experts to ensure field protocols are being followed. When processing samples in the laboratory, every batch of soil measurements includes a repeat measurement (where one randomly selected sample is processed a second time) and a randomised standard soil (where a soil sample of known value is analysed to confirm sufficient accuracy) as routine. Building in methods of 'independent verification', whether for sample provenance or measurements in the laboratory, will increase confidence in data but also provides an invaluable opportunity to salvage or repair data when mistakes inevitably occur. For example, on the collection of a soil sample, a paired electronic record can be created by surveyors with IDs generated from a barcoding system without manual input, which provides an independently verifiable record of which physical samples should be expected. Should a soil sample or data be mislabelled, there is then a traceable record to easily identify the true identity and sample location of that sample.

A choice must also be made of whether to process soil measurements via an in-house laboratory or through an external laboratory. Processing samples in-house will not always be possible, and may confer additional costs, but will confer greater control over measurement methodology and capacity to run additional quality assurance checks. From the experience of managing the CS soil monitoring programme, the ability to measure indicators in-house has been extremely valuable. Should an external laboratory be used, care should be taken to ensure issues of methodological continuity and sample preservation are considered, and that sufficient quality reporting is provided.

5.5 | Indicator specification table

In order to communicate science to policy, indicator specification tables were produced by Black et al. (2008) that link the indicator to a particular function of interest. A set of such tables were created for the UK for 13 indicators and the following functions: food and fibre production, environmental interaction, support of ecological habitat and biodiversity and protection of cultural heritage. The example for topsoil pH is adapted from Black et al. (2008) for food and fibre production (Table 3). Perhaps the more important aspect of developing such tables is determining the action levels; across the EU these are likely to vary depending on pedo-climatic zones.

6 | WHAT DOES SUCCESS LOOK LIKE?

Success is the ability to measure the state and change of soils in an effective, consistent and scientifically robust way. As an example, acid rain caused a decline in soil pH causing socioeconomic harm across Europe. In 1979, the Convention for Long Range Transboundary Air Pollution (CLRTAP) established within the framework of the United Nations Economic Commission for Europe (UNECE) was set up to tackle this, entering into force in 1983. The UK response was to reduce emissions. Soil monitoring programmes have provided the evidence of the effectiveness of this policy intervention (Reynolds et al., 2013), demonstrating a steady recovery of soil pH up to the 2007 sampling, following reductions in UK sulphur emissions and consequent deposition to land surfaces (Kirk et al., 2010; Reynolds et al., 2013; Seaton et al., 2023). In Wales the ERAMMP monitoring has helped bring P fertiliser additions under control. The monitoring of SOC across programmes has highlighted losses in cropland systems, for example, NSI, CS and NSIS. However, CS and NSIS have observed no overall change to 2007. The overall picture regarding changes in soil indicators is often nuanced and careful interpretation is required.

7 | DEVELOPMENT OF MONITORING ACROSS THE EU

In the EU, the LUCAS monitoring programme (Orgiazzi et al., 2018) operates across the member states (MS). This monitoring programme provides a vital resource for understanding soil state and change across the EU using internally consistent methods. The SML proposes additional monitoring by MS that will support their efforts to sustainably manage soils. The SML identifies soil districts as the administrative unit, which could be considered equivalent to the devolved administrations in the UK who are responsible for the implementation of agri-environment schemes. The SML does not specify a design, but Panagos, Broothaerts, et al. (2024) says, 'the EUSO proposed a stratified sampling method that possibly meets these requirements, while minimising the cost of sampling. The minimum sampling size is calculated by implementing the Bethel algorithm (Bethel, 1989)'. Hence, whilst the SML allows flexibility, new monitoring designs may well follow the principles and concepts discussed herein.

7.1 | Soil descriptors

The SML proposes indicators and interpretive metrics that are named 'soil descriptors', or 'parameters describing a physical, chemical or biological characteristic of soil health'. The SML currently proposes a suite of descriptors to be included in monitoring. These are: texture, salinisation (electrical conductivity); soil erosion (soil erosion rate); loss of soil organic carbon (soil organic carbon concentration); compaction (bulk density in topsoil and subsoil); excess nutrient content (extractable phosphorus, nitrogen); soil contamination (concentration of heavy metals in soil: As, Sb, Cd, Co, Cr (total), Cr (VI), Cu, Hg, Pb, Ni, Tl, V, Zn); reduction of soil capacity to retain water (soil water holding capacity); acidification (pH); loss of soil biodiversity (soil basal respiration in dry soil); land take (total artificial land); and soil sealing. These are in alignment with the indicators used in UK monitoring for which there is also broad scientific consensus (Bünemann et al., 2018). Current and future EU Mission projects will continue to develop and refine these descriptors. In addition, the 'criteria for healthy soil condition of the soil descriptors are split into nonbinding sustainable target values (at EU level) and operational trigger values (set at Member State level)'. Hence, the operational trigger values are aimed at preventing degradation and implementing interventions in areas where they are activated to reverse loss. UK monitoring has found a number of approaches useful for conveying science to policy, these include critical thresholds, trigger values, change metrics (Black et al., 2008) and more recently, benchmarks (Feeney et al., 2023).

8 | CONCLUDING CONSIDERATIONS FOR SOIL MONITORING IN THE EU

This example of long-term stratified random monitoring from the UK demonstrates the general principles for the

TABLE 3 Indicator-specification table for soil pH for the food and fibre production function adapted to UKCEH CS from (Black et al., 2008).

Soil indicator: Topsoil pH	
Major function	Food, fibre and wood production
Policy objectives	Maintenance of soil pH for food and fibre production
Source(s)	Black et al. (2008)
Indicator assessments	To determine whether soil pH values fall above or below action levels, indicating that function may be compromised
Domain of interest:	Devolved administrations and land use type reporting units
Units (indicator variable)	Soil pH in water units
Units (measured variable)	Soil pH in water units
Indicator parameter	Mean, SD and upper and lower 95% CLs following transformation to normal distribution
Indicator quantity	• Values above or below trigger values • Mean status and change for specified reporting classes
Type of result (quantitative)	Is soil pH significantly different to previous estimates?
Type of result (qualitative)	Is soil pH lower than the action levels?
Tolerance level (critical limit, base value) (d: tolerance level)	(i) The width of a 95% CI for the true mean is 2d or less, or (ii) The width of a 95% CI for true change in mean is 2d or less
Land use type(s)	Arable and horticultural (AH); improved grass (IG), vegetables (V), forestry (F)
Action level required (mean pH in water) by land use type	AH, Mineral: <6.5 AH, Peaty: <5.8 IG, Mineral: <6 IG, Peaty: <5.3 V, Mineral: <6.5 V, Peaty: <5.8 F, Mineral: <3.5 F, Peaty: <3.5 F, Calcareous: >8.4
Soil depth	Topsoil 0–15 cm
Appropriate sampling procedure	UKCEH Countryside Survey
Analytical method(s)	The pH of fresh soil is measured using a modified version of the method employed by the Soil Survey of England and Wales (Avery & Bascomb, 1974) to give a ratio of soil to deionised water of 1:2.5 by weight. The suspension is stirred thoroughly and left to stand for 30 min after which time the pH electrode is inserted into the suspension and a reading taken after a further 30 s (Reynolds et al., 2013)
Archiving	Samples should be archived for future analysis. Air-dried 2 mm sieved soil sample.
Additional information	'Tolerance levels' (±d): e.g., the SE, CIs, error variance etc.

Abbreviation: UKCEH, UK centre for ecology & hydrology.

operational framework linking the status (achieved through monitoring), impact (determined through analysis of monitoring data) and response (policy measures developed and implemented) in accordance with the DPSIR framework. The EU is larger than the UK and has a greater diversity of soils, but as MS are encouraged to develop monitoring, the above examples might prove useful. Moreover, whilst soil monitoring is relatively cheap compared with other forms of environmental monitoring, such as water or earth observation platforms, cost and efficiency should always be a consideration in design and indicator selection; as previously stated the EUSO proposes the use of the Bethel algorithm in this respect (Panagos, Broothaerts, et al., 2024). If MS pursue a stratified random design, then due consideration should be given to whether MS produce strata, or whether a pan-EU body such as the JRC should be tasked with producing strata that MS can use in their monitoring design. In line with this, the European Parliament Resolution includes a SML amendment (article 8a) directing the European Commission to support MS in setting the monitoring framework and facilitating the exchange of best practices—building on initiatives such as Soil Biodiversity Observation Network (SoilBON).

9 | CONCLUSION

This work has articulated the principles that support the goal of measuring the state and change of soils in an effective, consistent, and scientifically robust way across administrative districts. In so doing, this will provide the evidence to underpin the main goal of the Mission 'A Soil Deal for Europe' to transition towards healthy soils by 2030 and for the EU to have sustainably managed soils by 2050. The principles set out a framework with objectives that are met through robust, statistical sampling design. Indicators are selected according to criteria in support of statistical rigour and presented in indicator-specification tables that address functions of importance to policy; the DPSIR approach serves as the causal framework. The key elements are: *Monitoring Objectives*: purpose, adaptability, accessibility and transparency, ethical considerations, and timeliness; *Sampling Design*, functional reporting units, design, cost effectiveness; and *Indicator Selection*, sensitivity, specificity, measurability, targeted indicator selection and validity and reliability. Principles are less rigid than a rules-based structure and allow for some flexibility in achieving the goals. This is important for integration, but also allows for innovation. Moreover, potential issues and solutions are provided regarding the implementation of soil monitoring in support of policy that might prove useful in informing the development of soil monitoring across the EU. Experience from the UK demonstrates the value of such monitoring, especially when part of a wider ecosystem approach that can offer integrated analysis to understand what-works-where in the context of sustainable land management. Ultimately, the soil monitoring effort in the EU is one part of the puzzle to help all actors across the continent of Europe transition towards sustainable soil management by 2050.

AUTHOR CONTRIBUTIONS

David A. Robinson: Conceptualization; writing – original draft; funding acquisition. **Laura Bentley:** Writing – original draft. **Laurence Jones:** Writing – review and editing. **Chris Feeney:** Writing – review and editing; visualization. **Angus Garbutt:** Writing – review and editing. **Susan Tandy:** Writing – review and editing. **Inma Lebron:** Writing – review and editing. **Amy Thomas:** Writing – review and editing. **Sabine Reinsch:** Writing – review and editing. **Lisa Norton:** Writing – review and editing. **Lindsay Maskell:** Writing – review and editing. **Claire Wood:** Writing – review and editing; visualization. **Pete Henrys:** Writing – review and editing. **Susan Jarvis:** Writing – review and editing. **Simon Smart:** Writing – review and editing. **Aidan Keith:** Writing – review and editing. **Fiona Seaton:** Writing – review and editing. **James Skates:** Writing – review and editing. **Suzanne Higgins:** Writing – review and editing. **Giovanna Giuffrè:** Writing – review and editing. **Bridget A. Emmett:** Writing – review and editing.

ACKNOWLEDGEMENTS

We would like to acknowledge both Allan Lilly from JHI, Panos Panogos from JRC and Jakub Olewski from Natural England for reviewing and commenting on this manuscript. The research was funded by the Natural Environment Research Council award number NE/R016429/1 as part of the UK–ScaPE Programme Delivering National Capability. Also supported in part by the Glastir Monitoring and Evaluation Programme (GMEP) funded by the Welsh Government as part of the Environment & Rural Affairs Monitoring and Modelling Programme (Contract reference: C147/2010/11) and NERC/Centre for Ecology & Hydrology (CEH Projects: NEC04780/NEC05371/NEC05782). This project acknowledges funding from the European Union's Horizon Europe research and innovation programme under Grant agreement No. 101086179 AI4SoilHealth; and funding from the UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (Grant number 10053484).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

David A. Robinson https://orcid.org/0000-0001-7290-4867
Laura Bentley https://orcid.org/0000-0001-5055-7673
Fiona Seaton https://orcid.org/0000-0002-2022-7451
Suzanne Higgins https://orcid.org/0000-0001-6116-4410

REFERENCES

Avery, B. W., & Bascomb, C. L. (1974). *Soil survey laboratory methods.*

Baritz, R., Amelung, W., Antoni, V., Boardman, J., Horn, R., Prokop, G., Römbke, J., Romkens, P., Steinhoff-Knopp, B., & Swartjes, F. (2021). *Soil monitoring in Europe*. Indicators and thresholds for soil quality assessments EEA Report No 08/2022.

Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M., & Kirk, G. J. D. (2005). Carbon losses from all soils across England and Wales 1978-2003. *Nature*, *437*, 245–248.

Bentley, L., Feeney, C., Matthews, R., Evans, C., Garbutt, R., Thomson, A., & Emmett, B. (2023). *Qualitative impact assessment of land management interventions on Ecosystem Services ("QEIA")*. Report-3 Theme-6: Carbon Sequestration (Defra ECM_62324/UKCEH 08044).

Bentley, L., Thomas, A., Garbutt, A., Williams, B., Reinsch, S., Lebron, I., Brentegani, M., Keenan, P., Wood, C., Smart, S., Cosby, B., Emmett, B., & Robinson, D. (2024). First signs that national cropland organic carbon loss is reversing in topsoils. *European Journal of Soil Science* (Submitted).

Bethel, J. (1989). Sample allocation in multivariate surveys. *Survey Methodology*, *15*, 47–57.

Bhogal, A., Boucard, T., Chambers, B., Nicholson, F., & Parkinson, R. (2008). *Road testing of 'trigger values' for assessing site specific soil quality. Phase 2–other soil quality indicators*. Science Report. SC050054SR2, 59. Environmental Agency.

Black, H. I. J., Bellamy, P. H., Creamer, R., Elston, D., Emmett, B. A., Frogbrook, Z. L., Hudson, G., Jordan, C., Lark, M., Lilly, A., Marchant, B., Plum, S., Potts, J. M., Reynolds, B., Thompson, R., & Booth, P. (2008). *Design and operation of a UK soil monitoring network*. Environment Agency. Bristol Science Project Number: SC060073.

Brus, D. (2014). Statistical sampling approaches for soil monitoring. *European Journal of Soil Science*, *65*, 779–791.

Brus, D. J. (2022). *Spatial sampling with R*. CRC Press.

Bunce, R., Barr, C., Clarke, R., Howard, D., & Lane, A. (1996). ITE Merlewood land classification of Great Britain. *Journal of Biogeography*, *23*, 625–634.

Bunce, R., Barr, C., Gillespie, M., & Howard, D. (1996). The ITE land classification: Providing an environmental stratification of Great Britain. *Environmental Monitoring and Assessment*, *39*, 39–46.

Bunce, R., & Shaw, M. (1973). A standardized procedure for ecological survey. *Journal of Environmental Management*, *1*, 239–258.

Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., De Deyn, G., de Goede, R., Fleskens, L., Geissen, V., Kuyper, T. W., & Mäder, P. (2018). Soil quality–a critical review. *Soil Biology and Biochemistry*, *120*, 105–125.

Chapman, S., Bell, J., Campbell, C., Hudson, G., Lilly, A., Nolan, A., Robertson, A., Potts, J., & Towers, W. (2013). Comparison of soil carbon stocks in Scottish soils between 1978 and 2009. *European Journal of Soil Science*, *64*, 455–465.

Cooper, A. (1986). *The Northern Ireland land classification*. University of Ulster.

Corstanje, R., Mercer, T. G., Rickson, J. R., Deeks, L. K., Newell-Price, P., Holman, I., Kechavarsi, C., & Waine, T. W. (2017). Physical soil quality indicators for monitoring British soils. *Solid Earth*, *8*, 1003–1016.

Countryside_Survey. (2024). UKCEH Countryside Survey. https://www.ceh.ac.uk/our-science/projects/countryside-survey

Countryside_Survey_NI. (2024). Northern Ireland Countryside Survey. https://www.daera-ni.gov.uk/articles/northern-ireland-countryside-survey

De Rosa, D., Ballabio, C., Lugato, E., Fasiolo, M., Jones, A., & Panagos, P. (2024). Soil organic carbon stocks in European croplands and grasslands: How much have we lost in the past decade? *Global Change Biology*, *30*, e16992.

Emmett, B., Abdalla, M., Anthony, S., Astbury, S., August, T., Barrett, G., Beckmann, B., Biggs, J., Botham, M., & Bradley, D. (2017). *Glastir monitoring & evaluation programme*. Final report.

Emmett, B., Bell, C., Bentley, L., Birnie, M., Botham, M., Bowes, M., Braban, C., Broughton, R., Burden, A., Carvell, C., Cosby, B., Costa Domingo, G., Drewer, J., Evans, C., Feeney, C., Fletcher, D., Garbutt, R., Gaskell, P., Goodenough, A., … Bell, C. (2023). *Qualitative impact assessment of land management interventions on ecosystem services ("QEIA")*. Report-1: Executive summary: QEIA evidence review & integrated assessment (Defra ECM_62324/UKCEH 08044).

Emmett, B., Cosby, B., Bentley, L., Birnie, M., Botham, M., Bowes, M., Braban, C., Broughton, R., Burden, A., Carvell, C., Costa Domingo, G., Drewer, J., Evans, C., Feeney, C., Fletcher, D., Garbutt, R., Gaskell, P., Goodenough, A., Hassin, A., … Bell, C. (2023). *Qualitative impact assessment of land management interventions on ecosystem services ('QEIA')*.

Emmett, B., Frogbrook, Z., Chamberlain, P., Griffiths, R., Pickup, R., Poskitt, J., Reynolds, B., Rowe, E., Rowland, P., & Wilson, J. (2008). *Countryside survey. Soils manual*.

Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S., Frogbrook, Z., Hughes, S., Lawlor, A., Poskitt, J., Potter, E., Robinson, D., Scott, A., Wood, C., & Woods, C. (2010). *Countryside survey: Soils report from 2007* (p. 192). Centre for Ecology and Hydrology.

Emmett, B. A., & Wales_AXIS_II_GMEP_Team. (2013). An integrated ecological, social and physical approach to monitoring environmental change and land management effects: The Wales Axis II Monitoring and Evaluation Programme. In N. Boatman, M. Green, J. Marshall, K. Musters, W. Peach, S. Peel, G. Siriwardena, & B. Smith (Eds.), *Environmental management on farmland*. Aspects of applied biology (Vol. 118, pp. 31–39). *Association of Applied Biologists*.

ERAMMP. (2024). Informing, Monitoring and Evaluating Sustainable Land Management. https://erammp.wales/en

European Commission Directorate-General for Research and Innovation, Veerman, C., Pinto Correia, T., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B., Frison, E., Grand, A., Hristov, L., Kriaučiūnienė, Z., Pogrzeba, M., Soussana, J., Vela, C., & Wittkowski, R. (2020). *Caring for soil is caring for life – Ensure 75% of soils are healthy by 2030 for food, people, nature and climate*. Report of the Mission board for soil health and food. Publications Office.

European_Commission. (2021). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, EU Soil Strategy for 2030 – Reaping the benefits of healthy soils for people, food, nature and climate, COM (2021) 699 final of 17.11.2021*.

European_Commission. (2023). *European Commission, Proposal for a Directive of the European Parliament and of the Council on Soil Monitoring and Resilience (Soil Monitoring Law) COM/2023/416 final*.

Feeney, C. J., Robinson, D. A., Keith, A. M., Vigier, A., Bentley, L., Smith, R. P., Garbutt, A., Maskell, L. C., Norton, L., & Wood, C. M. (2023). Development of soil health benchmarks for managed and semi-natural landscapes. *Science of the Total Environment*, *886*, 163973.

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental Microbiology*, *13*, 1642–1654.

Gruijter, J. J., Bierkens, M. F., Brus, D. J., & Knotters, M. (2006). *Sampling for natural resource monitoring*. Springer.

Jenny, H. (1994). *Factors of soil formation: A system of quantitative pedology*. Courier Corporation.

Kirk, G. J., Bellamy, P. H., & Lark, R. M. (2010). Changes in soil pH across England and Wales in response to decreased acid deposition. *Global Change Biology*, *16*, 3111–3119.

Lilly, A., Bell, J., Hudson, G., Nolan, A., & Towers, W. (2011). *National Soil Inventory of Scotland 2007–2009: Profile description and soil sampling protocols (NSIS_2)*. Technical Bulletin, James Hutton Institute.

Loveland, P., & Thompson, T. (2002). *Identification and development of a set of national indicators for soil quality, R&D Project*. Record P5–053/PR/02, Environment Agency UK.

Merrington, G., Fishwick, S., Barraclough, D., Morris, J., Preedy, N., Boucard, T., Reeve, M., Smith, P., & Fang, C. (2006). *The development and use of soil quality indicators for assessing the role of soil in environmental interactions*. Report to UK SIC, SC030265, Environment Agency.

Metzger, M., Shkaruba, A., Jongman, R., & Bunce, R. (2012). *Descriptions of the European environmental zones and strata*. Alterra.

NCEA. (2024). Natural Capital and Ecosystem Assessment Programme. https://www.gov.uk/government/publications/natural-capital-and-ecosystem-assessment-programme/natural-capital-and-ecosystem-assessment-programme

Newell Price, J., Williams, A., Bentley, L., & Williams, J. (2023). *Qualitative impact assessment of land management interventions on Ecosystem Services ("QEIA")*. Report-3 Theme-3: Soils. (Defra ECM_62324/UKCEH 08044).

Nicholson, F., Boucard, T., Chambers, B., & Merrington, G. (2008). *Road testing of 'Trigger values' for assessing site specific soil quality. Phase 1 – Metals*. Science Report-SC050054SR1. Environmental Agency.

OECD. (1993). OECD core set of indicators for environmental performance reviews. *Environmental Monographs*, *83*, 1–39.

Or, D., Keller, T., & Schlesinger, W. H. (2021). Natural and managed soil structure: On the fragile scaffolding for soil functioning. *Soil and Tillage Research*, *208*, 104912.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernández-Ugalde, O. (2018). LUCAS soil, the largest expandable soil dataset for Europe: A review. *European Journal of Soil Science*, *69*, 140–153.

Panagos, P., Ballabio, C., Poesen, J., Lugato, E., Scarpa, S., Montanarella, L., & Borrelli, P. (2020). A soil erosion indicator for supporting agricultural, environmental and climate policies in the European Union. *Remote Sensing*, *12*, 1365.

Panagos, P., Borrelli, P., Jones, A., & Robinson, D. A. (2024). A 1 billion euro mission: A soil Deal for Europe. *European Journal of Soil Science*, *75*, e13466.

Panagos, P., Broothaerts, N., Ballabio, C., Orgiazzi, A., De Rosa, D., Borrelli, P., Liakos, L., Vieira, D., Van Eynde, E., & Arias Navarro, C. (2024). How the EU soil observatory is providing solid science for healthy soils. *European Journal of Soil Science*, *75*, e13507.

R_Core_Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reynolds, B., Chamberlain, P., Poskitt, J., Woods, C., Scott, W., Rowe, E., Robinson, D., Frogbrook, Z., Keith, A., & Henrys, P. (2013). Countryside survey: National "soil change" 1978–2007 for Topsoils in Great Britain—Acidity, carbon, and total nitrogen status. *Vadose Zone Journal*, *12*, 1–15.

Richter, D.d., & Markewitz, D. (2001). *Understanding soil change: Soil sustainability over millennia, centuries, and decades*. Cambridge University Press.

Rowe, E., Emmett, B., Frogbrook, Z., Robinson, D., & Hughes, S. (2012). Nitrogen deposition and climate effects on soil nitrogen availability: Influences of habitat type and soil characteristics. *Science of the Total Environment*, *434*, 62–70.

Seaton, F. M., Robinson, D. A., Monteith, D., Lebron, I., Bürkner, P., Tomlinson, S., Emmett, B. A., & Smart, S. M. (2023). Fifty years of reduction in sulphur deposition drives recovery in soil pH and plant communities. *Journal of Ecology*, *111*, 464–478.

Simfukwe, P., Hill, P., Emmett, B., & Jones, D. (2011). Soil classification provides a poor indicator of carbon turnover rates in soil. *Soil Biology and Biochemistry*, *43*, 1688–1696.

Stone, D., Ritz, K., Griffiths, B., Orgiazzi, A., & Creamer, R. (2016). Selection of biological indicators appropriate for European soil monitoring. *Applied Soil Ecology*, *97*, 12–22.

United-Nations. (2014). *System of environmental-economic accounting (SEEA)*. Central framework. https://unstats.un.org/unsd/envaccounting/White_cover.pdf

Van Vuuren, J. A., van Dijk, D., & Claassens, A. S. (2002). Comparison and co-operation in quality control in soil analysis: WEPAL, AgriLASA. *Communications in Soil Science and Plant Analysis*, *33*, 2851–2859.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., & Bourne, P. E. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 1–9.

Wood, C. (2011). *The sampling strategy for countryside survey (up to 2007)*. https://nora.nerc.ac.uk/id/eprint/19487/1/CS_Sampling_Strategy_Revised.pdf

Wood, C. M., Bunce, R. G., Norton, L. R., Maskell, L. C., Smart, S. M., Scott, W. A., Henrys, P. A., Howard, D. C., Wright, S. M., & Brown, M. J. (2018). Ecological landscape elements: Long-term monitoring in Great Britain, the countryside survey 1978–2007 and beyond. *Earth System Science Data*, *10*, 745–763.

How to cite this article: Robinson, D. A., Bentley, L., Jones, L., Feeney, C., Garbutt, A., Tandy, S., Lebron, I., Thomas, A., Reinsch, S., Norton, L., Maskell, L., Wood, C., Henrys, P., Jarvis, S., Smart, S., Keith, A., Seaton, F., Skates, J., Higgins, S., … Emmett, B. A. (2024). Five decades' experience of long-term soil monitoring, and key design principles, to assist the EU soil health mission. *European Journal of Soil Science*, *75*(5), e13570. https://doi.org/10.1111/ejss.13570