European Journal of Agronomy 168 (2025) 127537

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Putting numbers to a metaphor: A Bayesian Belief Network with which to infer Soil Quality and Health

Kirsty L. Hassall [a,2], Joanna Zawadzka [b], Alice E. Milne [a], Ronald Corstanje [b], James A. Harris [b], A. Gordon Dailey [a], Aidan M. Keith [c], Margaret J. Glendining [a], S.P. McGrath [a], Lindsay C. Todman [a,o,4], Paul Alexander [d,3], Philippa Arnold [e], Amanda J. Bennett [g], Anne Bhogal [f], Joanna M. The late Clark [h,1], Felicity V. Crotty [i], Claire Horrocks [j], Nicola Noble [k], Robert Rees [l], Matthew Shepherd [m], Elizabeth A. Stockdale [n], Edward W. Tipping [c], A.P. Whitmore [a,*]

[a] *Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK*
[b] *Cranfield University, Bedfordshire, MK45 0AL, UK*
[c] *UK Centre for Ecology & Hydrology, Lancaster LA1 4AP, UK*
[d] *Royal Horticultural Society, Wisley, Woking, Surrey GU23 6QB, UK*
[e] *National Farmer's Union, Stoneleigh, Warwickshire CV8 2LG, UK*
[f] *ADAS Gleadthorpe, Mansfield NG20 9PD, UK*
[g] *Agriculture and Horticulture Development Board, Middlemarch Business Park, Siskin Parkway East, Coventry CV3 4PE, UK*
[h] *Soil Research Centre, School of Archaeology, Geography and Environmental Science, University of Reading, Reading RG6 6D, UK*
[i] *Global Land Team, Ricardo, The Gemini Building, Fermi Avenue, Harwell, Didcot OX11 0QR, UK*
[j] *Rothamsted Research, North Wyke, Okehampton, Devon EX20 2SB, UK*
[k] *National Sheep Association. The Sheep Centre, Malvern, Worcestershire WR13 6PH, UK*
[l] *Scotland's Rural College, West Mains Road, Edinburgh EH9 3JG, UK*
[m] *Natural England, Sterling House, Dix's Field, Exeter EX1 1QA, UK*
[n] *NIAB, Cambridge CB3 0LE, UK*
[o] *University of Reading, Reading RG6 6D, UK*

A R T I C L E I N F O

Keywords:
Soil quality
Soil health
Bayesian networks
Long-term experiments
National data sets

A B S T R A C T

Soil Quality or Soil Health are terms adopted by the scientific community as metaphors for the effects of differing land management practices on the properties and functions of soil. Because they are metaphors, consistent quantitative definitions are lacking. We present here an approach based on expert elicitation in the field of soil function and management that offers a universal way of putting numbers to the metaphor. Like humans, soils differ and so do the ways in which they are understood to become unhealthy. Long-term experiments such as the Broadbalk Wheat experiment at Rothamsted provide unparalled sources of data with which to investigate the state and changes of soil quality and health that have developed from known management over timescales of one hundred years or more. Similarly, large-scale datasets such as the National Soils Inventory and Countryside Survey provide rich resources to explore the geographical variability of soil quality and health in different places against a background of different observed management practices. We structure experts' views of the extent to which soil delivers the functions expected of it within Bayesian Belief Networks anchored by measurable properties of soil. With these networks, we infer the likely state of soil (i) on Broadbalk, (ii) at locations throughout England & Wales as well as inferring (iii) the most straightforward ways of improving soil quality and

* Corresponding author.
E-mail address: andy.whitmore@rothamsted.ac.uk (A.P. Whitmore).

[1] It is with great sadness that we report the recent death of Professor Clark
[2] Present address: University of Warwick, Coventry, CV4 7AL
[3] Present address Bulrush Horticulture Ltd, Newferry Road, Bellaghy, Magherafelt, Co. Londonderry, N. Ireland, BT45 8ND
[4] 0000-0003-1232-249X

https://doi.org/10.1016/j.eja.2025.127537
Received 17 October 2023; Received in revised form 2 February 2025; Accepted 2 February 2025
Available online 10 March 2025

health at the locations in (ii). Our methodology has general applicability and could be deployed elsewhere or in other disciplines.

1. Introduction

Soil Quality and Soil Health (SQH) are concepts (Janzen et al., 2021, Powlson, 2020) that have found considerable traction amongst soil scientists, practitioners and policy-makers alike, but can mean little without context. Their usefulness depends on quantitative deployment. There are many other terms and metaphors that are in use in scientific discourse that defy neat quantification: human health, for example. The use and function of these metaphors has proven incredibly effective within scientific disciplines, and when communicating with policy makers and the general public. Our challenge is to ground the use of such metaphors in the underlying scientific understanding upon which they are based so that their meaning is unambiguous**.** SQH is hard to quantify in a consistent, objective fashion (Bünemann et al., 2018). This has its origin in the complexity and multiplicity of the functions that soil is expected to fulfil and the fact that soil harbours a great diversity of organisms (e.g. Bardgett and van der Putten, 2014). For this reason, SQH (and other metaphorical measures) is often expressed as an index which is used to suggest that this soil or that practice is qualitatively better than another (Karlen et al., 1997. Karlen and Stott, 1994). Such one dimensionality engenders spurious certainty, whereas multi-dimensionality seems comprehensive but can be vague. Hierarchical structures that lead to overall scores (Ros et al., 2022) and demonstrate how properties and functions combine to give rise to indices and thence SQH can go a long way to mitigating this difficulty.

It would undoubtedly be useful if the properties of soil that enable it to deliver substantial crop yields, for example, were the same as those that enable it to perform other important functions such as buffer water flow, resist erosion (e.g. Rickson et al., 2012), regulate greenhouse gases or provide habitat for biodiversity (Doran and Zeiss, 2000. Kleijn et al., 2019), because development of a simple scored value for SQH would then be straightforward. However, the many characteristics of soil affect its functions to different degrees or in different ways (e.g. Wade et al., 2022). Further, SQH has become a catch-all for the diverse aspirations of different stakeholders who may interpret a score or value of an index (Rutgers et al., 2012; Karlen and Stott, 1994) differently from one another. Given the need to answer the questions 'How much better is this soil than another?' or 'What should be done to improve this soil?', several intrinsic properties of soil are sometimes used to articulate SQH in a multi-dimensional index. An example is the well-known radar plot (e.g. Kleijn et al., 2019 or Rutgers et al., 2012) that compares several properties together in one diagram (see for an example and fuller explanation Fig E1 in the extended data section). However, such figures fail to capture interactions between the components that determine the dynamic nature of SQH (Rickson et al., 2012. Wagenet and Hutson, 1997), and joining unrelated data points can be misleading, confusing, and presumptuous. Too often, the components are chosen in an *ad hoc* fashion (usually, they are those that are available), without due consideration of their functional importance or relevance (see Baveye, 2020, for a critique and Harris et al., 2022 for a systematics approach). Formal visual assessment (VA) of soil according to some prescribed recipe overcomes the limitations of single or multiple unconnected indices (Schipper and Sparling, 2000. Guimarães et al., 2011) but is not fully objective and depends heavily on the expertise of the individual doing the assessing. Despite this, VA has the merit that it provides a number that is intended to represent SQH explicitly. As has been seen, other methods often supply a hotch-potch of data, indices or properties that are probably connected to SQH in some manner but without consistent recognition of the function a soil is performing. Logical sieves (Ritz, et al., 2009, Zwetsloot et al., 2022) have been used to great effect in the soil health arena, but come into their own when ranking or selecting which of many indicators to use before going on to articulate SQH. Benchmarking (Feeney et al., 2021) and cluster analysis (Seaton et al., 2021) have helpfully been used to establish thresholds of SQH and implicit interactions between determinants respectively. Machine learning techniques (El Behairy et al., 2024 or Wilhelm et al., 2022) have also been used to infer soil health, but such methods depend on large amounts of data. Training appears to be on prior SQH values obtained from other index-based assessments, although latent variable analysis (Wade et al., 2022) is a step forward and similar to what we propose below.

Expert-systems are computer programs that rely on two components: a knowledge base and an inference engine. A knowledge base is an organized collection of facts about the system's domain. An inference engine interprets and evaluates the facts in the knowledge base in order to provide information about e.g. SQH. Typically, expert systems combine expert knowledge and data into a system of very many rules and provide answers as probabilistic outcomes. Thus, expert systems overcome many of the shortcomings of the index-based or visual assessment approaches to SQH. However, they tend to be rule-based, somewhat inflexible as a result and it is not always straightforward to update or add to them.

Bayesian methods are among the most promising means to both structure and interpret a knowledge base (Bui et al., 1999) because of the sophistication of their inference and because of the natural manner in which expert opinion and data can be combined. To reflect this, they are sometimes known as Bayesian *Belief* Networks (or Bayes Nets, BN for short). They are graph-based, directional networks that can incorporate probability distributions of the component variables. They have had diverse applications in the biological and social sciences (Aalders, 2008. Corstanje et al., 2015. Levontin et al., 2011). Their directedness proceeds from multiple pieces of information or properties, such as soil organic matter content or texture, to a conclusion such as the extent of delivery of a function. BNs can be constructed either by data-mining or from knowledge-based approaches (Corner et al., 2002) or both. The second and third options mean that it is possible to apply BNs to study areas where there is a shortage of data by eliciting the views of experts (Taalab et al., 2015). Crucially the opinions of more than one expert can be incorporated alongside objective data thus overcoming the objections to visual assessment or the sometimes *ad hoc* nature of the data presented in, say, radar plots. BNs model the domain of interest (i.e. the soil) under considerations of uncertainty through probabilistic reasoning, whilst expert systems model what an expert would reason assuming a known, fully certain domain or support. A BN infers impacts on a soil property (a child node) in relation to multiple parent nodes thus explicitly taking account of interactions. Although BNs are not explicitly dynamic in the sense of a series of differential equations or a computer simulation model, the experts can be asked to keep in mind the changing nature of the processes that they are considering and they can also be mindful of the various purposes to which information on, for example, SQH will be put. A BN, therefore, gives us not so much an index for SQH as a quantified and interconnected network of all that experts think is vital.

Given the expertise-led focus on soil provided by BNs and the explicit way in which they define SQH and handle uncertainty and likelihood, we think BNs provide a more reliable, informative and explanatory means to articulate SQH and its nature than other methods currently in use. To date much published work with BNs has tended to trawl the literature for data to construct the BNs, only later linking the results using Bayesian methods coupled to expert opinion. Our approach is to quantify expert views alongside data from the start in order to capture the major interactions between variables and networks of variables that the experts believe to be important. In this way our nets act as a *definition* of what SQH is, and in principle the method should extend widely to other scientific metaphors.

Our objective therefore is to demonstrate that a BN integrates the essential components and interactions between components of SQH in a meaningful way that naturally links soil properties to their functions and thus how the whole integrates to quality and health. Importantly, we structure experts' views on exactly what SQH is via an inferential chain of reasoning through soil functions to measurable properties - something that index and other methods tend not to do.

Notwithstanding the need for universality, a single network for all soils that is relevant to all stakeholders would be very difficult to achieve. Context (which implies function), such as agriculture or nature is key. Therefore, we addressed three broad land-uses and elicited a BN for each: (i) arable, (ii) livestock agriculture or (iii) semi-natural land-use. Almost all land in England and Wales is managed to some degree, hence semi-natural. Further subdivisions are possible, but part of the appeal to different stakeholders is an index's multi-functionality, so we aimed to retain the ability of each broad land-use to express the multi-faceted nature of soil function – e.g., production as well as environmental quality.

We describe first in detail how we built the nets and how we evaluated their sensitivity to variations in measurable variables. We demonstrate the use of a BN against (i) long-term field data and (ii) national surveys of soil properties and function and then go on to infer the likely health or otherwise of soils regionally in England and Wales, bearing in mind what these soils are expected to deliver, before inferring which properties, if changed, would most readily bring about the greatest improvement in SQH. Approaches based on expert opinion to articulate soil health have been criticised on the basis that the structure and content of these expert systems are sometimes opaque (Wade et al., 2022). We seek to rectify these issues by (i) depicting both the structure and weight of connections within our nets fully and (ii) making explicit the elicited responses of the experts.

The Long-Term Experiments at Rothamsted and elsewhere have proven their worth again and again in the years since they began (e.g. Johnston and Poulton, 2018). Bearing in mind the different contexts of soils and their stakeholders, we look in detail at the quality and health of the soil in one of Rothamsted's long-term arable experiments: Broadbalk field, parts of which have grown wheat almost every year since 1843 (ERA). Data conservation and monitoring are also long-term activities from which environmental science benefits and in contrast to the field data from Rothamsted and in order to show the power of our methodology, we apply the nets to all three land-uses using data from Cranfield University's National Soils Inventory (NSRI, 2001) on a gridded network of sampling points throughout the whole of England & Wales and supplemented with observations from the UKCEH Countryside Survey (Carey et al., 2008).

2. Methods

2.1. Framing of a SQH BN

To develop meaningful models of SQH through an expert driven BN approach, we imposed a hierarchical structure to the networks. Specifically, three different types of nodes were considered:

1. The SQH node.
2. Functional and Process nodes
3. Measurable or Property nodes

SQH is to be defined in terms of the functional processes which a soil delivers, i.e. what is needed to ensure a good SQH, or what functional changes result in a bad SQH. These functional or process nodes are then connected through inferential chains to measurable nodes that can be directly connected to data. Thus, SQH is linked (via inference) to directly measurable quantities.

2.2. Expert selection

A schematic overview of the entire BN development, parameterisation and analysis process is given in Fig. 1. This starts with the careful selection of appropriate experts. We identified experts based on their knowledge and experience while seeking to cover different disciplines and land-management sectors. We invited 18 experts to participate in the elicitation process, mindful of a balance between genders and between experience and youthful enthusiasm. The experts invited ranged from soil scientists and soil surveyors to policy makers, farm advisers and managers. There were 3 agronomists, 2 biologists and 1 ecologist, 2 physicists, 2 chemists, 2 policy-makers both with interests in semi-natural, and 3 representatives of different sectors of the agricultural community. Five to eight experts is considered the optimal number (Clemen and Winkler, 1985) of practitioners for the elicitation process. In all, 16 attended to represent three sectorally-based SQH BNs for: semi-natural (5 persons), livestock management (6 persons) and arable land-uses (5 persons). A slight imbalance arises because not all invitees were able to attend but more than half of those who did have interests that span more than a single discipline. Participants attended a two-day workshop where day 1 was focussed on training and day 2 on developing the three land-use specific BNs.

2.3. Elicitation protocol

Developing the BNs comprised several steps:

(i) Identifying all relevant soil properties and processes,
(ii) reducing these to a manageable subset of the most important,
(iii) agreeing the relationships between this subset of properties and the ways that their interactions allow us to infer SQH,
(iv) quantifying these relationships using expert opinion

The protocol below evolved from a series of practice workshops that we held to hone our procedure.

Roughly one month prior to the two-day workshop we held 15–30 minute one-to-one interviews via video link with each of the invitees. These video meetings served to introduce the ideas behind the project and to focus attention on important aspects of SQH. We asked each expert two questions: (i*) How might you build up the quality or health of soil?* And (ii) *If you had a high-quality soil, what might you do to degrade it?* By means of shared screens, we captured the processes by which suggested actions might improve or degrade SQH in the expert's mind. These networks were displayed, anonymously, at the start of the workshop for experts to peruse and compare. Following all preliminary interviews, we collated information from across the networks by means of word clouds and bar charts showing the frequency of use of particular terms. From these we identified commonalities and shared views and were able to then seek out datasets and literature for experts to refer to if they wished during the workshop. Displayed around the workspace, these preliminary views formed a conceptual (naïve) model that is the basis for refinement in the process of developing a BN.

On day one of the workshop, experts were given an introductory session on BNs and the value of inference. Experts listed and prioritised the most important functions and processes in soils that are needed to infer SQH generally. Specific land-uses were introduced on day two. The prioritisation step was needed to gain a consensus on a net of manageable size and because one shortcoming in the BN approach is that the inferential reasoning required is quite hard for humans to grasp. Training the experts on a common network for all soil as a preliminary exercise helped to prevent divergence between the eventual, specialised nets subsequently.

Participants were asked "What functions or processes define whether a soil has good or bad SQH?"; these ideas were recorded on post-its,

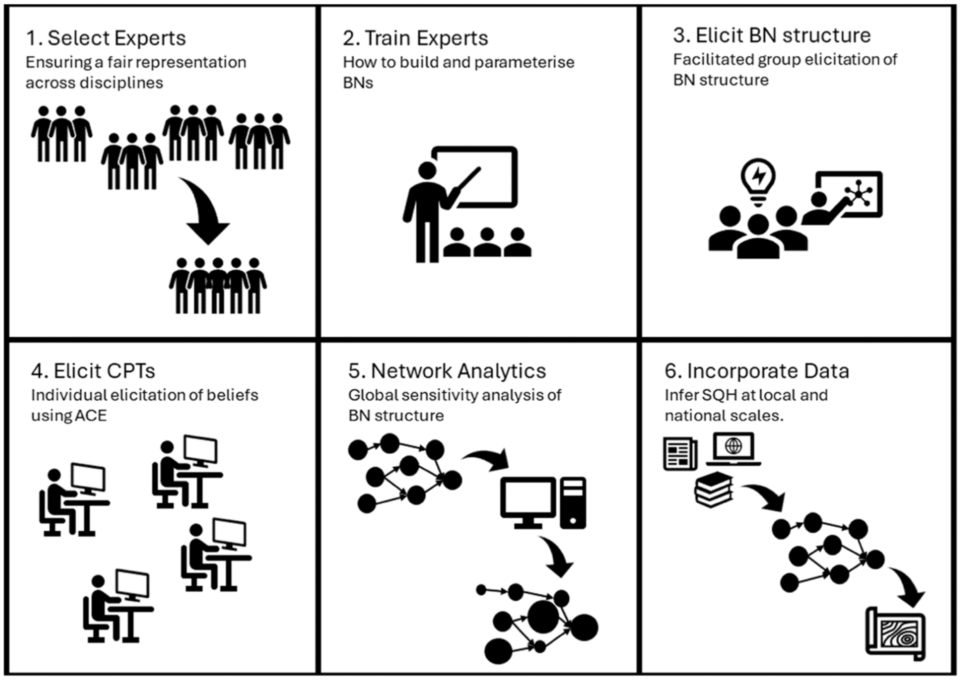


Fig. 1. Schematic representation of the elicitation, development and deployment of the Bayesian Nets. See 2 for full details.

considered by the group and refined to a minimum set.

Candidate parents that had a similar definition were then amalgamated into a single node. Groups were also encouraged to form intermediate nodes to avoid too many parents leading into SQH directly. For example, under the semi-natural network (Fig. 2c), the node "Water Regulation" was formed as an intermediary node from both "Water flow" and "Water quality".

A group consensus on the BN structure was achieved in most cases. For more contentious nodes, facilitators encouraged participants to keep the node present in the network, as the step of defining the conditional probability tables allowed individuals to essentially ignore its influence if they thought it was not important.

To link function and process nodes to measurable property nodes, participants were asked "What pieces of information do we need to be able to infer the identified functions delivered by the soil?". This then defined the parents of the nodes identified in step (i).

Occasionally, experts wished to define an Action node, for example, Tillage. In such cases, facilitators worked to identify what functions or properties such an action would affect: a heavily tilled soil will be of poor quality because it contains fewer earthworms or a less varied structure. In this way, abundance of macro-organisms or structure became the property node.

Once a preliminary consensus was reached on the structure and constituents of the net, we asked experts to think carefully what they meant by each node, and agree quantitative divisions into as few categories as possible whilst maintaining a reasonable resolution, ideally 2 or 3 reasonable categories, for example Light, Medium or Heavy, for a soil texture node corresponding to a combination of sand, silt or clay (Figure S1.1). Accordingly, we encouraged experts to introduce intermediate nodes rather than set up networks consisting of child nodes that possess a large number of parents. The reason for this parsimony is that the conditional probability table, that is the crux of the network and which is needed to elicit the nets, has a total number of categories of: $N_C \sum_{i=1}^{P} N_P(i)$, where N_C is the number of child categories, P the number of parent nodes and $N_P(i)$ is the number of categories in parent i. Thus, 3 parents each having 2 categories and feeding into a child with 3 categories implies 18 probabilities. The interrelationship of many more connections and categories than this is difficult to manage conceptually.

Our app ACE (Hassall et al., 2019), was designed to make completing these conditional probability tables as straightforward as possible and help avoid unlikely or impossible combinations. Experts were encouraged to think how these functions and processes link together, to infer SQH. In keeping with the need for nets to be of a manageable size, the experts developed generic nodes whose values were context dependent: *appropriate* biodiversity, *adequate* nutrient levels and so on (See SI- data sources). On day one the target was a general soil, but experts worked together in the groupings that we expected to keep together for day two. This introductory exercise took most of the first day, after which we gave workshop experts the opportunity to reflect, discuss and feed back their experiences of the elicitation process and procedure over dinner in advance of day two during which the three groups repeated the exercise but concentrated on developing nets within their core expertise restricted to one of three specific land-uses: (i) arable, (ii) livestock and (iii) semi-natural. With the structure in place, we set about eliciting the Conditional Probability Tables (CPT) using our bespoke ACE software (Hassall et al., 2019). A laptop, networked to the app, was provided to each expert, so that their individual choices could be captured throughout the elicitation process.

The CPTs were sense checked and because voice recorders were used to capture discussion it was possible to check the intentions of experts for certain issues. When constructing the nets, we queried odd-seeming views with experts by video link following the workshop using open questions so as to not to bias responses.

We held a video conference some weeks later to present the nets to all participants and ask them for comments on one another's nets and confirm that they were content with their own. As a result, we changed the name of the node that represented soil organic matter to SOM in all three cases and converted this SOM node into a data node (which previously was not the case) in the arable net. At this stage, too, we reduced (with the experts permission and help) the number of categories in the SQH node in the Livestock net from five to two to match the arable and semi-natural nets.

The nature of each node, the states which each can take are given in Tables 1.1a & b, Tables 2.1 a & b, Tables 3.1 a & b; the means by which we populated the data nodes is described briefly below (2.2) and fully in Supplementary Information – data sources, processing and mapping.

Table 1.1a
Data nodes (coloured) in the Arable Net (Fig. 2a).

Node specified by experts	Source of data	Thresholds or Categories
Texture	NSI: CLAY, SILT, VSAND, MSAND, CSAND	{Heavy, Medium, Light}
pH	NSI: PH	{Extreme, 5.5–7.5}
SOM	NSI: CARBON	{< 2 %, 2–4 %, > 4 %}
Nutrients	NSI: K_NITRATE, P_OLSEN	{Inadequate, Adequate}
Contamination	NSI, various. Thresholds: Charlton et al., 2016.	{No, Yes} mg kg^{-1}
Plough pan	Bradley (2002)	{Absent, Present}
Aggregate Stability	Defra 2004	{Unstable, Stable}
Excess Nutrients	NVZ maps	{Absent, Present}
Beneficial Biology	CS: Emmett et al., 2016a	{Low, High}
Pests & Pathogens	LandCover ® *plus: Crops*	{Below Threshold, Above Threshold}
Ecological Diversity	CS: Bunce et al. (2014)	{Low, High}

Table 1.1b
Intermediate nodes (grey) in the Arable net (Fig. 2a).

Node specified by experts	Parents	Thresholds or Categories
WHC Water Holding Capacity	Texture	{Low, High}
Agricultural Biology	Pests & Pathogens, Beneficial Biology	{Unsupportive, Supportive}
Chemistry	Nutrients, Contamination, pH	{Bad, Good}
Physical Condition	Aggregate Stability, Plough Pan, Storage	{Bad, Good}
Productivity Consistency	Agricultural Biology, Chemistry & Physical Condition	{Below Average, Average, Above Average}
Regulation	Excess Nutrients, Physical Condition, Storage	{Bad, Good}
Storage	SOM, WHC	{Low, High}

Table 2.1
Data nodes (coloured) in the Livestock Net (Fig. 2b).

Node specified by experts	Source of data	Thresholds or Categories
Compaction	Estimated risk of poaching or impeded drainage	{No, Yes}
SOM	NSI: SOM= 1.72 *Carbon	{Red, Amber, Green}
Texture	NSI: CLAY	{Light, Medium, Heavy}
Soil Depth	NSI + Soil series: DROCK	{< 40, > 40} cm
Slope	NSI: SLOPE	{< 2°, 2° - 10°, > 10°}
Soil P	NSI: P_OLSEN	{< 15, 15–45, > 45} mg/l.
Applied N	BSFP 2019	{< 100, 100–200, > 200} kg/ha
pH	NSI: pH	{0–5, 5–7, 7–14} Experts {0–6,6–7,7–14} AHDB
Micronutrients	NSI: CA_ACID and MG_ACID	{Insufficient, Sufficient}

These BNs elicited in this way became the networks displayed in Fig. 2.

In the literature there are two approaches to combining views from multiple experts. Opinions can be combined either (i) through allowing a group of experts to reach a consensus by repeated discussion and revision or (ii) through mathematical aggregation. There is evidence to suggest that (i) may induce a dependence between responses (Hanea et al., 2017) whilst Taalab et al. (2015) recommend that a model go through several iterations until experts agree the structure. We adopted a hybrid approach in which the structure of the net, including nodes and categories were agreed in consensus but the experts populated the conditional probability tables (CPT) independently of one another. In this way, we obtained a consensus structure representing features important to inferring SQH from a variety of different disciplines and stakeholders, whilst maintaining the individual's perspective on the importance of each feature.

Table 2.1b
Intermediate nodes (grey) in the Livestock net (Fig. 2b).

Node specified by experts	Parents	Thresholds or Categories
Air	Ammonia, GHG Balance	{Positive Outcome, Neutral Outcome, Negative Outcome}
Ammonia	pH, Nitrogen Surplus	{Low, High}
Animal Health	Forage Nutritional Quality, Risk of Excess Water	{Good, Bad}
Environment	Air, Water	{Positive Outcome, Neutral Outcome, Negative Outcome}
Erosion Risk	Slope, Soil Structure	{Low, High}
Forage Nutritional Quality	Micronutrients, Sward Diversity	{Meets Requirements, Does not Meet Requirements}
Forage Yield	pH, Soil Structure	{Close to Potential, Below Potential, Well Below Potential}
GHG Balance	Compaction, Nitrogen Surplus	{Sequestration, Neutral, Emitting}
Nitrogen Surplus	Applied N	{< 50,50–75,> 75}
Productivity	Animal Health, Forage Yield	{Meets Expectation, Below Expectation, Well Below Expectation}
P Surplus	Soil P Flag, Soil Structure	{Low below 15, Medium 15–45, High Over 45}
Risk of Excess Water	Compaction, WHC	{Low Risk, High Risk}
Soil Structure	SOM, Texture	{Good, Bad, Really Bad}
Sward Diversity	Applied N, Soil P Flag	{Grass Forbes adv spp5, Grass Forbes 25, Grass only}
Water	Risk of Excess Water, Water Quality	{Good, Bad}
Water Holding Capacity	Soil Depth, Soil Structure	
Water Quality	Erosion Risk, P Surplus	{Meets WFD[1], Does not meet WFD}

[1]Water Framework Directive

Table 3.1a
Data nodes (coloured) in the Semi Natural Net (Fig. 2c).

Node specified by experts	Source of data	Thresholds or Categories
Soil Nitrate	CS: Emmett et al., 2016b	{Low, High}
ΔpH[1]	Inferred from Pearsall (1952) & current values (NSI: PH)	{< 0.5 units below expected, expected, > 0.5 units expected}
Metals Contamination	NSI: acid extracted Cd, Cr, Cu, Ni, Pb, Zn	{No exceedance, Moderate exceedance, High exceedance}
Bare Soil	Observation	{< 10 % bare, > 10 % bare}
Soil Moisture	NATMAP	{Drained, Not drained}

[1]ΔpH is the difference between the currently observed and expected pH assuming no acidic deposition during the second half of the 20th Century.

For specific analytics (below) the individual CPTs were averaged (Sections 2.4 and 2.5). Nets and CPTs were incorporated into BNs using the Netica package (Norsys, 2016, Almond, 2022).

The methodology is updateable: views from other experts can be incorporated once elicited. The nets reflect current expert opinion. If other pressures become relevant, these can be added.

2.4. Network analytics

One of the merits of BNs compared with other approaches is that interactions and strength of interactions are made explicit. Specifically, the belief of a node is defined to be the posterior probability conditional on the status of the network. Within our derived BNs SQH is a two-state node, thus, this posterior probability (or belief) of SQH can be fully

Table 3.1b
Intermediate nodes (grey) in the Semi Natural Net (Fig. 2c).

Node specified by experts	Parents	Thresholds or Categories
Appropriate Chemistry	Contamination, ΔpH, Soil_Nitrate	{Bad, Good}
Appropriate Soil Biology	Appropriate Chemistry, Plant Community Type	{No, Yes}
Plant Community Type	ΔpH, Soil_Nitrate	{Dominated by Competitors, Dominated by Stress tolerators}[1]
SOM (C Storage)	Bare Soil, Plant Community Type, Soil_Moisture	{Decreasing, Stable, Increasing}
Water Flow	Bare Soil, Soil_Moisture	{Extreme, Normal}
Water Quality	Appropriate Chemistry	{Bad, Good}
Water Regulation	Water Flow, Water Quality	{Bad, Good}

[1]See Grime (1979)

characterised by the probability that SQH is "Good" conditional on the network $\mathcal{N}$ having status n,

$$\text{Belief}(\text{SQH}\mathcal{N}) = \text{Prob}(\text{SQH} = \text{Good}\mathcal{N} = n). \tag{1}$$

We analysed the derived BNs and what they meant in four different ways, we: i) investigated the local structures within the full BN, ii) assessed the importance of the different nodes in determining SQH, iii) assessed the relative importance of the observable data nodes in determining SQH and iv) examined combinations of node states in relation to SQH outcomes. The first three aspects are depicted in the network diagrams of Fig. 2 through arrow width, node size and SQH pie sector, respectively, with the fourth depicted in the associated regression trees (Figs. E2). These procedures are described below. Note, for these procedures all nodes had an additional parent to that depicted in Fig. 2, specifically, an elicitee (expert) node capturing each individual's CPT.

i) Local structure: Arrow Width in Fig. 2

The entropy of a node, X, is defined to be

$$H(X) = -\sum_x \text{Prob}(X = x)\ln\{\text{Prob}(X = x)\} \tag{2}$$

where x is a particular state e.g. "Good" in SQH. The entropy, $H(X)$, is a measure of the amount of information contained within a node. The entropy is maximised when the state of X is unknown, i.e. X has a uniform distribution. If X is known with certainty, the entropy is 0.

Recall, that for a node with parent(s) Y, the marginal probabilities are obtained by summing over the parental states, i.e.

$$\text{Prob}(X = x) = \sum_y \text{Prob}(X = x|Y = y)\text{Prob}(Y = y)$$

Thus, in all calculations of entropy, the distribution of the node X has been obtained by first summing over each elicitee node, assuming a uniform prior, in turn.

The conditional entropy of a child X in relation to one of its parents Y is defined by,

$$\begin{aligned}H(X|Y) &= \sum_y \text{Prob}(Y = y)H(X|Y = y)\\ &= -\sum_y \text{Prob}(Y = y)\sum_x \text{Prob}(X = x|Y = y)\log\{\text{Prob}(X\\ &= x|Y = y)\}\end{aligned}$$

and is a measure of the amount of information in X given knowledge about Y. If X and Y are independent, $H(X) = H(X|Y)$. However, if there is a dependence between X and Y, we would expect $H(X|Y) < H(X)$.

The mutual information between a variable X and variable Y is defined to be,

$$MI(X, Y) = H(X, Y) = H(X) - H(X|Y)$$

Thus, for each node in the network, we can look at the local structure between it and its direct parents. Specifically, the thickness of the arrows in Fig. 2 is given by the Mutual Information (*MI*) between each node and its connected parent as a fraction of the entropy of the child node *MI*(child, parent) / *H* (child). Thus, the thicker the arrow, the larger the mutual information and thus the greater the dependence between parent and child resulting in a greater reduction in entropy due to knowing the state of the parent.

ii) Importance of nodes: Node Sizes in Fig. 2

To investigate how different nodes affect the probability of SQH, we ran a large simulation study, or global sensitivity analysis, to calculate the belief of the SQH node for different statuses of the network. Specifically, a network status is defined by a particular combination of known node states (e.g. SOM is decreasing, Soil nitrate is low and metal contamination is in moderate exceedance. See Table 1.1a, Table 1.1b, Table 2.1, Table 2.1b, Table 3.1a, Table 3.1b for a definition of all node states). Given any status, the posterior probability of SQH can be calculated (Eq. 1). It was computationally infeasible to calculate the belief for every combination of node states (in the order of 10 million combinations for the semi-natural network, 7.8 billion for the arable network and 2.2×10^{15} for the livestock network). Thus, the beliefs have been obtained for every combination of node states where up to five nodes are known, whilst simultaneously fixing the elicitee node to be unknown and hence averaging over all experts. This study will be referred to as the 5-way study in what follows and generated 59,488 node combinations for the semi-natural, 478,040 for the arable and 286,512 for the livestock BNs. As with any model, not all scenarios for which the model is defined are equally likely in practice and here as below (iii), we have checked the nature of the node combinations. It is plausible that up to 10 combinations of the data nodes in the arable net are very unlikely or even impossible. Despite this, relative comparisons made using the global sensitivity analysis will be unbiased.

A main effects linear model was fitted to the belief of SQH over these node combinations for each landuse in turn. For example, the model:

$$\begin{aligned}\text{Belief(SQH)} = {}& \text{Soil\_Nitrate} + \text{pH} + \text{Plant\_Community\_type} + \text{Soil\_moisture} + \text{Bare\_soil} + \text{C\_Storage}\\ &+\text{Metals\_contamination} + \text{Appropriate\_chemistry} + \text{Appropriate\_Soil\_biology}\\ &+\text{Water\_quality} + \text{Water\_flow} + \text{Water\_regulation}\end{aligned} \tag{3}$$

describes the effect each term has on SQH in the semi-natural landuse. This effect can be summarised by the type II F-statistic of each term (F_{node}) in the fitted regression model and represents the amount of variation in SQH explained by the node. The diameter of each node in Fig. 2 is determined by $\ln(F_{node})$.

iii) Relative importance of data nodes: SQH sectors in Fig. 2

In practice, determining SQH comes down to observations on the measurable properties. These are what we have called the "data nodes" and which are depicted in colours other than light grey in Fig. 2. To investigate the relative importance of the data nodes in determining SQH, we ran a second simulation study. Here, the belief of SQH was calculated for every possible combination of node states but restricted to knowing only the data nodes. There are 5 data nodes in the semi-natural landuse, resulting in 432 node combinations, 11 data nodes in the arable network resulting in 1480,352 node combinations and 9 data nodes in the livestock network resulting in 110,592 node combinations. As above, the overall effect of each data node was assessed through a main effects linear regression analysis, using the SQH belief as a response, and summarised by the associated type II F-statistic for each term. For example, the model

$$\text{Belief}(\text{SQH}) = \text{Soil\_Nitrate} + \text{pH} + \text{Soil\_moisture} + \text{Bare\_soil} + \text{Metals\_contamination} \quad (4)$$

was used to describe the effect of each data node in the semi-natural landuse. The relative areas of the coloured sectors in the SQH node of Fig. 2 are determined by the F-statistics of each term from model (4) expressed as a proportion out of the sum of all F-statistics from model (4).

iv) Regression trees:

The analytics described in i)-iii) show which nodes are influential in determining SQH, but they do not identify the manner by which those nodes are important. Regression trees were fitted to the belief of SQH (e.g. probability SQH=Good) using results from the 5-way study and considering all non-SQH nodes as possible explanatory variables (Figure E2). A regression tree sequentially identifies the value of the explanatory variables (i.e. a node state) that results in the best partition of the response variable (the SQH belief). This partition is found through an assessment of the deviance. Explicitly, the deviance summarises the goodness of fit by assessing the difference between the observed SQH belief and the model prediction. For example, the belief of SQH in the semi-natural land-use is first partitioned into 2 categories based on whether SOM is decreasing or not. This partition of SQH belief is associated with the largest difference in deviance between any two possible partitions. The tree algorithm continues sequentially partitioning the response variable until the minimum difference in deviance between any two groups (set here at 0.01) is reached. The regression trees thus enable us to determine the combination of states that result in different SQH beliefs.

2.5. Interfacing the nets with data and mapping SQH

Inference of SQH in each net (Fig. 2) ends at measurable quantities (data nodes). These data are often spatially explicit, allowing a user to input local values to obtain a site-specific estimation of SQH. More generally, we aimed to estimate SQH in a spatially explicit way for mapping and validation against existing knowledge. To do so, we collated data on all input nodes such that either i) a spatially explicit value or ii) a distribution of values representative of a population was available.

To make the BNs operational, we need to have consistent, widespread sources of data from which to derive background probability distributions of the values of the measurable properties as default (prior) positions for the network (Fig. 2). Alongside these, we also need specific point data values where we seek to produce maps of SQH or the potential for amelioration (Figs. 4 and 5).

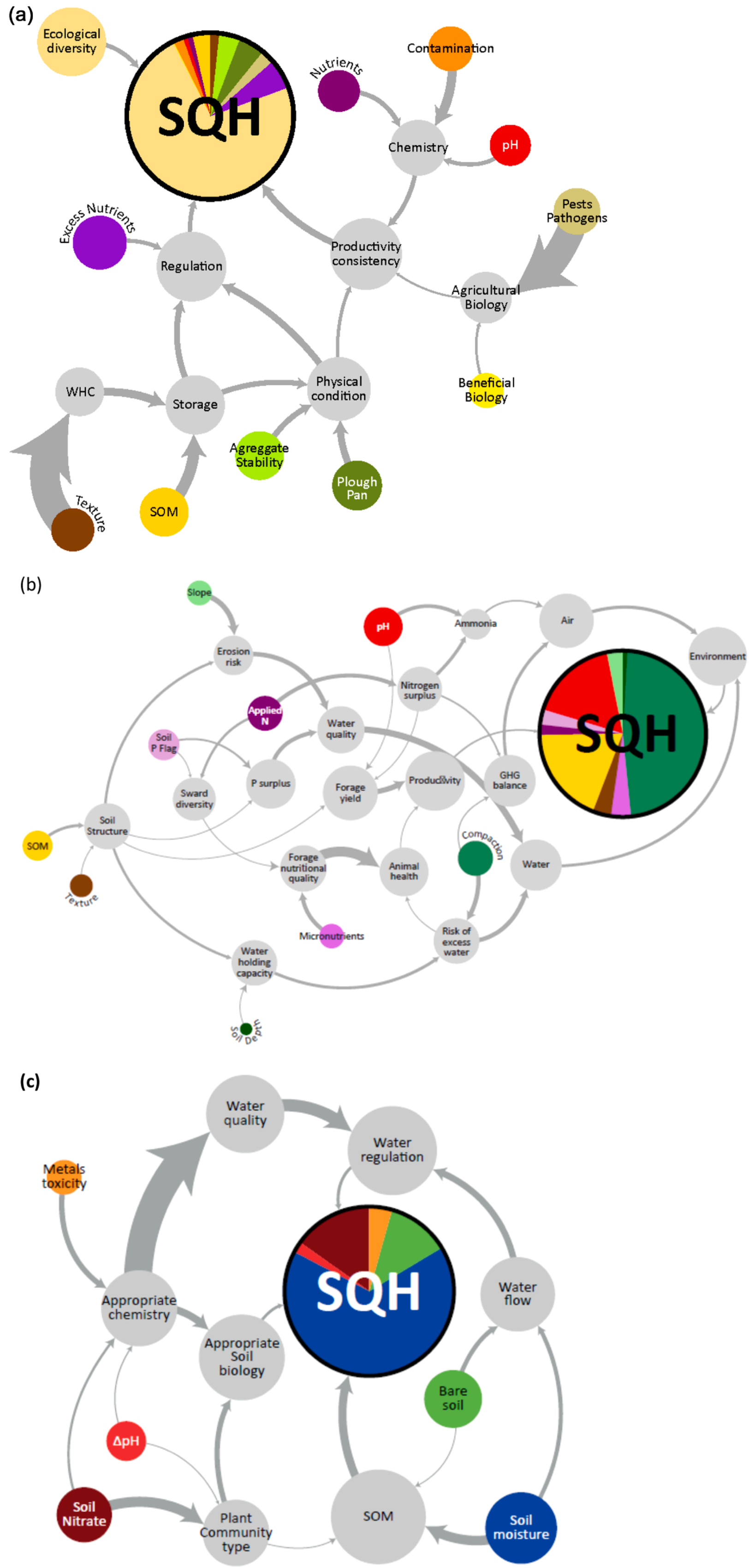
Major sources of data used to achieve this are; the National Soil Inventory (NSI) for England & Wales (NSRI, 2001), NATMAP (Cranfield University), the UKCEH Countryside Survey (Carey et al., 2008 Emmett et al., 2010, Maskell et al., 2008, Rowe et al., 2012) and the UKCEH LandCover Plus © maps (2017, 2016, 2007). Sources of data and the ways they have been used are given in full in the Supplementary Information – data sources.

Categories or thresholds for each node as defined by our experts are given in Tables (1a &b – arable net, 2a & b – livestock net and 3a & b semi-natural net) along with the sources of data where appropriate (data nodes Sections 1, 2 & 3 in SI). In some cases, measured or even measurable data do not map directly onto these categories and thresholds as conceived by our experts and so a degree of pre-processing was involved. We tabulated the instructions left by the experts, but in many cases populating the probability tables for the data nodes with specific values was not straightforward. Although some items can simply be read from one of the databases, others require interpolation from spatially mapped data, some themselves require inference from more than one source of information some of which itself might be available as discrete data and some on a spatially-mapped basis. In all cases we detail (in SI) the course of action taken, working with publicly available data. We took the NSI locations, on their 5 km grid, as the basis for our maps (Figs. 4 and 5) and inferred values for data at these locations either directly from other datasets if available or indirectly as spelt out in SI. For the intermediate (grey) nodes we state the categories agreed upon by the experts, what they mean and if appropriate, numerical thresholds and units. The effect of these data is best seen within the networks themselves.

Whilst the maps derive from national datasets (Figs. 4 and 5), values of SQH can be made at specific locations if a land-manager has local information or a regulatory body mandates measurement. Site specific data could be subjective, such as the relative productivities of the target field compared with any nearby. Where data are available for the intermediate nodes (grey discs in Fig. 2), these may be preferred because the structure of the net is such that upstream parent nodes become unnecessary. A number of pixels have *No data* land-use in Fig. 4. These are either urban pixels or ones where no land-use was recorded in the Countryside Survey (Carey et al., 2008). Where land-use was recorded but CS data was absent, the land-use given was: Deciduous (1 occurrence), horticultural crops (41), orchard (35), other (48), recreation (65), rough grazing (2), salt marsh (5), scrub (1), or blank (68).

For the Broadbalk Wheat experiment (Fig. 3) we used the arable net defaults supplemented with objective measurements as described in Supplementary Information (SI) 1a. Data was obtained from ERA and Watts et al. (2006)

To calculate opportunities for amelioration (Fig. 5), it should be noted that the BNs are not causal and therefore direct attribution cannot be given. Rather we calculated where the greatest opportunity for amelioration is by comparing the state of data nodes at each location to the state the data nodes should be in if SQH was in its "best possible" state. Note, this typically is not Prob(SQH = Good) = 1 due to the probabilistic relationships throughout the BN. Specifically, for each location, we identified the data node that was i) not in its optimal state and ii) was associated with the greatest improvement in SQH compared to all non-optimal node states at that location. This then identifies the data node that if it were changed through, for example, altering the underlying soil processes or functions, would be associated with the greatest change in SQH (hence "opportunities for" rather than direct amelioration).



(caption on next page)

Fig. 2. Bayesian Networks that define SQH. Networks of soils under (a) arable, (b) livestock, and (c) semi-natural land-uses. Networks display the lines of inference that lead to SQH; regression trees the importance of knowledge of the states of nodes. Discs are nodes in all BNs, black-ringed nodes are the SQH end-point. Distinct colours of the sectors in the SQH node correspond to those of the data nodes i.e. those which consist of measurable properties. Intermediate, potentially unobservable nodes are depicted in grey. For exact meanings and states of nodes see methods (SI). Where intermediate nodes are observable, such values may be preferred in practice. Nodes are connected by arrows representing the conditional dependence between the variables that the nodes signify. A sink/child node is conditionally independent from all other nodes in the network given the connection to the direct source/parent nodes. The size of each sector in the SQH node represents the proportion of variation explained by each data node as obtained from main effects regression analysis (only data nodes used as explanatory variables) of the probability of Good SQH using results obtained from network simulation runs (sensitivity: see methods). Node diameter (both data and intermediate) represents the amount of variation in SQH explained by the node as determined from a main effects linear regression (using all nodes as explanatory variables) of the probability of a Good SQH. The SQH nodes are arbitrarily large in order to make them and their sectors clearly visible. Arrow thickness represents the relative importance of parents in determining a connected child and can be compared within the same net. Relative node diameters and SQH sectors are compared on logarithmic scales; arrow thicknesses are compared linearly.

3. Results

3.1. Nets

We elicited the main determinants of SQH from our experts within three networks: one for each of arable, livestock and semi-natural land-uses. The three networks (Fig. 2a-c) reflect experts' preferences as well as the nature of the land-uses. Consequently, they differ in composition. In what follows, we present the findings from the BNs, which in turn are inferred from the views of the experts. Each node can take a value depending on two or three states (Table 1). Part of the value of the BN approach is that our nets contain information not only on the most important factors that contribute to the SQH, but also the most important *combinations* of factors (extended data, Fig E2a-c). Because we elicited SQH first and asked our experts to reason inferentially back to measurable data those nodes closest to SQH carry more weight.

3.1.1. Arable

In contrast to the other soils, SQH in arable soil (Fig. 2a) has a direct parent data node – *Ecological Diversity*. Consequently, this node assumes great importance in inferring SQH relative to the other measurable data (Fig. 2a, size of coloured sectors in SQH node). Its state, Good or Bad, is virtually binary in suggesting good or bad SQH based on the subset of observable data nodes. If known and because they are also immediate parents in the net in Fig. 2a, *Productivity Consistency* and *Regulation* are also highly important in signifying arable SQH. Thus, provisioning and regulating ecosystem services are key and of similar combined weight. The dependence can be seen in the main branches of the regression tree, Fig E2a. *Productivity Consistency* is one and a half times as important as the other two regulating factors; this importance is represented by the arrow thicknesses in Fig. 2a. (*Productivity Consistency*: *Regulation*: *Ecological Diversity* = 1: 0.65: 0.63). Finding above average *Productivity Consistency* is key if the best probability of Good SQH (0.78) is to be inferred in arable soils (arrow thickness, Fig. 2a). If *Productivity Consistency* is Unknown then the regression tree analysis suggests that either *Regulation* must be in a Good state or *Ecological Diversity* must be High to obtain a probability of at least 0.6 that SQH is Good (Fig E2a). In terms of other measurable nodes, the absence of a plough pan or managing nutrients well appears to be associated with best SQH

3.1.2. Livestock

In common with arable land SQH in livestock agriculture is chiefly inferred from production (*Productivity* in Fig. 2b, *Productivity Consistency* in Fig. 2a) and wider ecological considerations (*Environment* in Fig. 2b, *Ecological Diversity* in Fig. 2a). Good *Productivity* is essential for best SQH (0.72, Fig. 2b & E2b), followed by a positive score for *Environment* to obtain a probability of Good SQH equal to 0.63. However whilst Ecological Diversity was dominant among the measurable data feeding into SQH in the arable case, SQH can only be inferred reliably in livestock soils from good knowledge of several measurable factors, chief among these are the extent of *Compaction* or poaching, the *pH* of the soil and *SOM* – Soil Organic Matter content - (Fig. 2b). The complexity of livestock production is one reason why many interacting factors are needed to infer the state of the soil. Intermediate nodes generally contribute equally to their children and in similar proportion to other nodes (similar arrow thicknesses in Fig. 2b. Compare this with the variable importance of contributions of nodes in the arable and semi-natural nets (Fig. 2a and c).

3.1.3. Semi natural

In contrast to the agricultural soils the nodes closest to, and thus having greatest importance in inferring SQH are chiefly to do with regulating ecosystem services and less to do with provisioning: *Water Regulation*, *SOM* (where our experts focus on how carbon storage is changing) and *Appropriate Soil Biology*. *Soil Moisture* explains more variation in SQH in semi-natural soils (SQH node Fig. 2c) compared to other data nodes, with *Soil Nitrate* and whether or not the soil is *Bare* almost equally as important as one another. Good environmental *Water Regulation* and *SOM* are strongly associated with good SQH in these soils as seen in their key roles in the regression tree (Fig E2c). Note that the ΔpH node here represents the difference of the observed pH from the expected pH of a pre-acidic deposition soil (see SI). Although the impact of *Appropriate Chemistry* on *Water Quality* is very large (arrow size Fig. 2c) the eventual value of *Water Regulation* to inferring SQH is relatively small compared to *SOM* which is three times more likely to suggest Good SQH than the other parent nodes (*SOM: Appropriate Biology*: *Water Regulation* = 1: 0.32: 0.35). Increasing *SOM* and Good *Water Regulation* are required to achieve best values of SQH (probability equal to 0.81, Fig E2c).

In the livestock and to a lesser extent the arable net, most nodes are similar in size. In particular, several of the measurable data nodes are of a similar size to the intermediate function nodes. This is less true of the semi-natural net where the data modes are generally small. It perhaps suggests that it is the functioning rather than the state of soil that is thought (by our experts) to be responsible for delivering best SQH in semi-natural soils.

3.2. Broadbalk long-term experiment

The Broadbalk (Rothamsted) long-term experiment (see methods) has been running since 1843 and by applying data from this experiment to the arable BN, we derived estimates for SQH under a range of different crop management scenarios. Crop yield was used as an independent proxy for SQH for comparison because production is the main function for this arable field. This experiment was set up in the 19th Century with the specific aim of understanding how nutrients affect yield. The results are well known to us now, but at the time the results were of great importance. Broadbalk confirmed its originator's hypothesis that wheat obtains its nitrogen from the soil and not the air and demonstrated the importance of N over P (in the first instance) for yield. The arable SQH net (Fig. 2a) is able to distinguish between the many differences in soil that result from the long-term application of particular management practices (Fig. 3). Where N is lacking, SQH is clearly poor on Broadbalk (Fig. 3) but the worst values are found in the middle of the field where the amount of N applied is moderate but P and K are in very short supply. Best SQH in this field experiment appears to be associated with greater

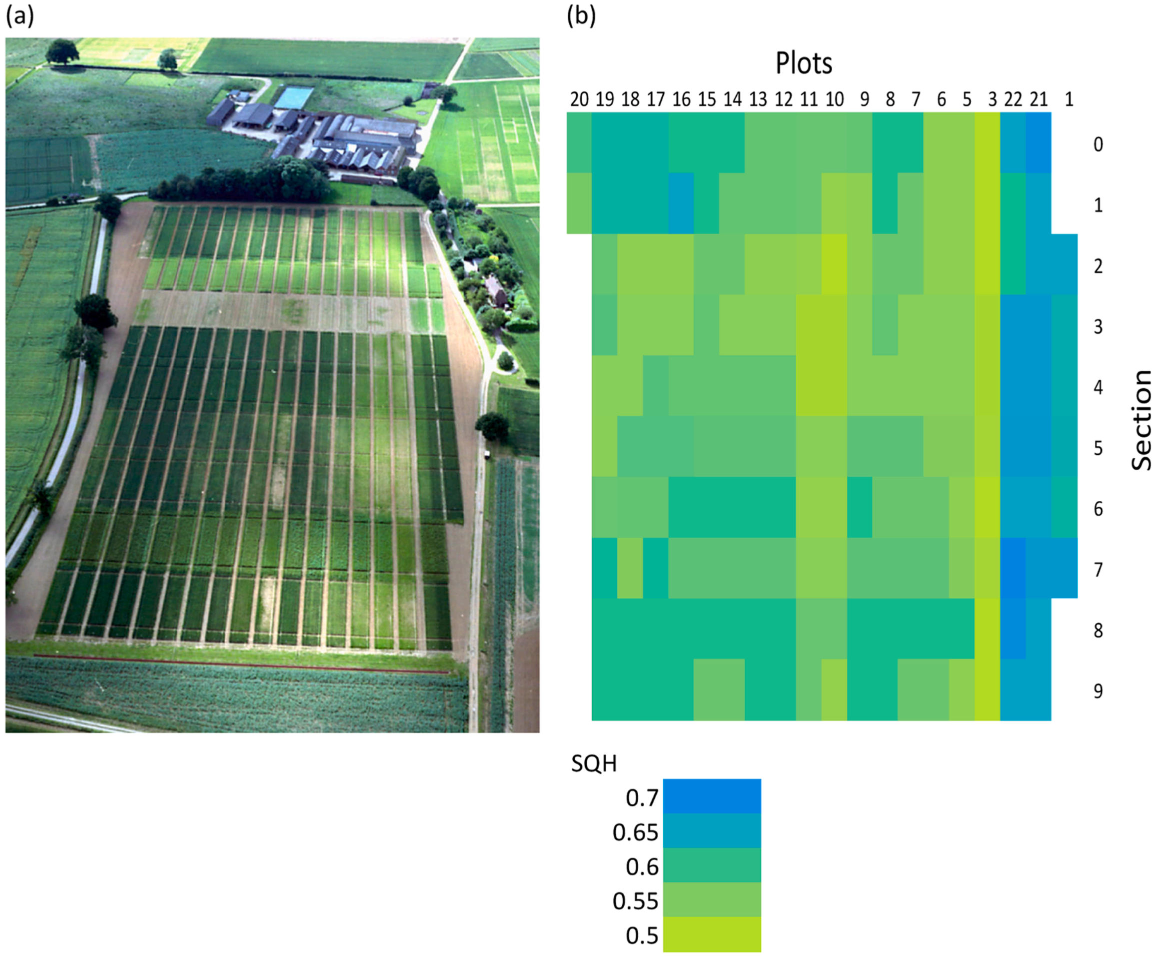


Fig. 3. The Broadbalk field experiment at Rothamsted has grown wheat on a silty clay loam soil under different but largely consistent management practices since it began in 1843. (a) Aerial photograph of the experiment oriented so that the sections (rows 0–9) descend from top to bottom. (b) Heatmap of SQH on Broadbalk field with the sections laid out in the same manner as Fig. 3a – section 0 is at the top, 9 at the base of the figure. The experiment tests the interaction between nutrient and manure applications (Plot treatments) against general crop management practices such as rotation or the use of biocidal chemicals (Sections). Sections 0, 1, 6, 8 and 9 grow wheat continuously and so risk soil-borne pathogens; other sections carry wheat in rotation such that second wheats carry a greater risk from soil-borne pathogens than first wheats but less than third. Layout is for 2021. Section 8 does not receive weedkillers and has a diverse flora as a result; plots differ in the particular regime of nutrients they receive (mainly K, P or N); the strips 2.2 and 2.1 (labelled 22 & 21) receive 35 tonnes farm-yard manure (fresh matter) each autumn. Although silt contents of the soils in the plots are predominantly above 55 %, clay varies from below 20 % in the bottom right corner in the figure to almost 40 % in the top left. These various factors influence SQH as shown; scale is identical to that used in Fig. 4a. Data to create the output using the arable net (SI: 1a) come from Watts et al. (2006) and from the Electronic Rothamsted Archive - ERA. Full details of the Broadbalk experiment can be found on ERA: or obtained from the curators on request. Author MJG is a curator.

amounts of either clay or organic matter. Poorest values of SQH are found where there is a nutrient deficiency, e.g plots 3 (all sections) which had received no fertiliser of any kind for 179 years, plots 5 and 6 which are deficient in N, plots 10 which are deficient in K and P and plots 11 which are deficient in K. Amounts of Soil organic matter (carbon) are relatively small in Sections 2–7, possibly because these sections included regular fallow years as part of the rotation during 1968–1996, with less carbon input to the soil overall as a consequence. As with the map of England & Wales (Fig. 5), managing nutrients is the one action that would most improve these arable soils – because their desired function is to produce food without loss of nutrients to air or water. Similarly, reducing the incidence of fallow and otherwise improving ground cover would be likely to increase organic matter inputs to soil and thus incorporation into SOM. Figure E1 gives a comparison between the arable BN and a radar plot.

3.3. National scale

The nets can be applied nationally. The National Soil Inventory (NSRI, 2001) records a consistent set of data from measurements made on a 5 km grid across England and Wales. Information held in the database includes land-use. In broad terms the climate is drier in the East and warmer in the South; the North and the West are more hilly. Thus, arable land is to be found predominantly in the East, livestock, fed from grassland is more abundant in the North and West. Semi-natural land is largely found in the uplands.

The probability of finding good SQH in any part of England and Wales is generally greater in semi-natural soil than soil under livestock agriculture which in turn is greater than soil under arable use (Fig. 4a). Upland soil is generally in semi-natural land-use whilst many of the livestock soils of intermediate quality are found in the West (Fig. 4b). Arable soils appear likely to be worse in the North than the South of the country which derives from a finding that there tend to be more non-agricultural species in field boundaries in the South of the country (Carey et al., 2008) and so potentially greater levels of *Ecological Diversity*. The risk of *Excess Nutrients* is deduced from whether or not the land is in a designated Nitrate Vulnerable Zone (NVZ). Even if farmers take action to control nitrate efflux to water systems, there is an

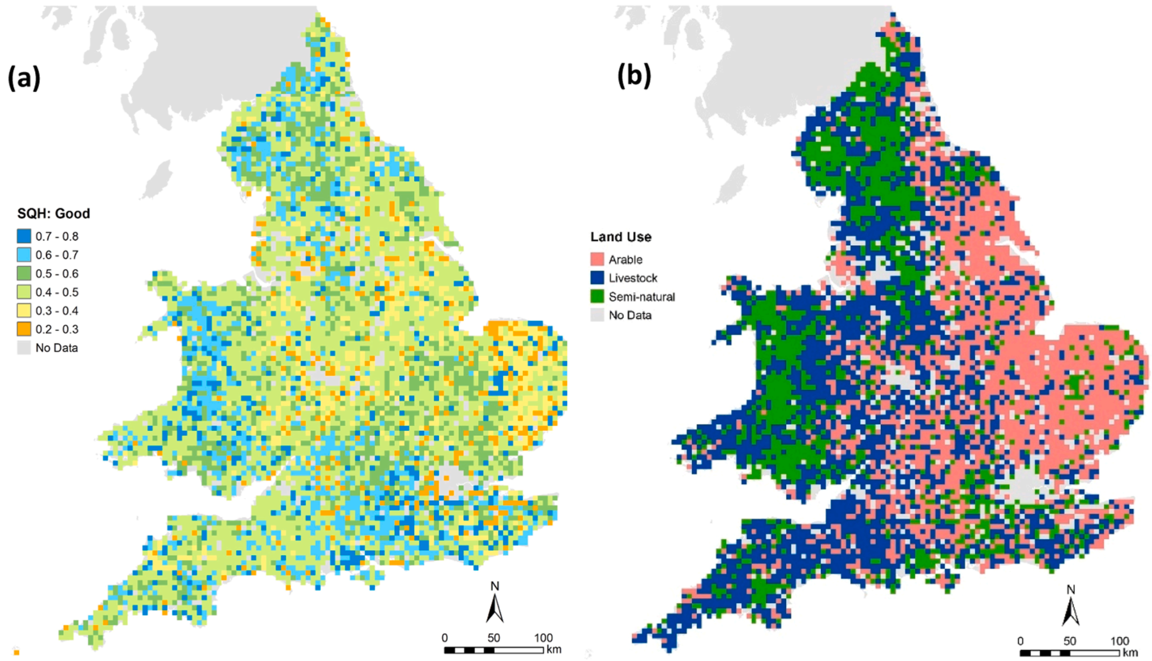


Fig. 4. State of SQH in England & Wales: (a) map of the probability of finding Good SQH in England and Wales based on data on land-use, texture and nutrient status in the National Soils Inventory on a 5 km grid and on inferred biological parameters derived from the Countryside Survey. All land-uses. The three respective land-use nets were applied at locations in (b) and have been used to compile the map in (a).

enhanced risk of poor SQH at locations within an NVZ all other things being equal. Soils in semi-natural habitats on the other hand can also be inferred to be in a poor state if the *Appropriate Soil Biology* is depleted (Fig. 2c). Both factors account for much of the geographical variation in values of SQH (Fig. 4a).

3.4. Improving SQH

The nets can be used to prioritise actions that might improve SQH. Although our nets are not causal, it is possible to identify the states of nodes which if changed would be most likely to lead to widespread improvement in SQH. We cannot say for certain that at this or that point better SQH will follow from an action but we can say that country-wide and over many instances of a similar intervention, better SQH would be likely.

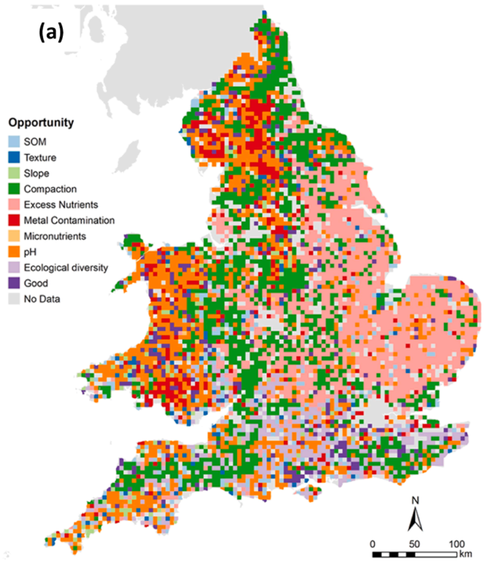
Where arable land-use predominates in England and Wales, it is likely that more soils could be improved by reducing the risk of *Excess Nutrients* (1246 instances that are within a Nitrate Vulnerable Zone out of a total of 1901) than by any other intervention (Fig. 5a & b). Farmers may, indeed should, already be doing so of course. Our national maps present risk based on probabilities of land state or position rather than detailed knowledge of current management. The Broadbalk net (Fig. 3) demonstrates the power of more specific local knowledge about nutrient management in arable soils. Although knowledge about *Ecological Diversity* is the most important directly measurable contributor to good SQH (Fig. 2a), *Productivity Consistency* and *Regulation* (into which *Excess Nutrients* feed, Fig. 2a) are together vital for the very best quality (Fig E2a). These results reflect both the structure of the elicited network and the input values for the data nodes of soils derived from two large-scale ongoing surveys as explained in the SI: the National Soils Inventory (NSRI, 2001) and the Countryside Survey (Carey et al., 2008). Managing nutrients more carefully, by whatever means is thus likely to lead to arable soils that are of better quality and in better health in the sense that they are less eutrophic or less likely to transmit pollution to the wider environment.

Many opportunities exist to ameliorate livestock soils that are likely to suffer from *Compaction* (1132 instances out of 2063) or where the *pH* is too low (Fig. 5). The pH of soils over basic rock is difficult to change. Direct interventions to improve *Productivity* or *Environment* (Fig. 2b, E2b), would help too, in so far as this is possible. Other states of nodes that are associated with poor SQH are slow (*SOM*) or impossible (*Texture, Slope*) to change. If a poor soil is difficult to ameliorate for these reasons, a change of land-use may be the only way to achieve Good SQH.

Despite historical exposure to acid rain, ΔpH has improved in recent years (Rose et al., 2016). Nonetheless, the discrepancy between actual and expected pH is the factor that most frequently suggests poor SQH in semi-natural soils. In this case, as is general in semi-natural soils, the opportunity is most often to modify industry or land management elsewhere. Information about *Soil Moisture* and its child nodes, *Water flow*, *Water regulation* and *SOM* is of critical importance to determine good SQH (Fig. 2c, E2c), with knowledge about *Water quality* and *Water regulation* key to inferring the very best quality. *Metals Contamination*, which we infer from heavy metal concentrations in soils that are above ambient background levels, affects *Appropriate Chemistry* (Fig. 2c) and is an issue in about 40 % of semi-natural soils. (Fig. 5, 286 instances in High exceedance and a further 87 in Moderate exceedance for at least one metal out of 1251 soil samples where metals data is available.) High metal content in soil may imply deposition from the air, mine spoils or some other historical misuse of land that reduces its function. It can also arise, however, from natural causes such as the underlying geology. In any case, it is difficult to improve SQH by removing metals. Poor values of soil moisture or soil organic matter may be more easily amended, eventually.

4. Discussion

There are perhaps three limitations to the BN modelling that should be made explicit. The first is that our implementation uses a discrete BN. All nodes are assumed to have a finite number of distinct categories. A greater level of nuance could be captured if some nodes were considered as a continuous variable instead. For example, when integrating the top-level data nodes with data, it is often the case that a threshold value was needed to define the distinct categories such as *Soil Nitrate* which must above or below 10 mg N kg^{-1}. A more realistic scenario would be to



(b)

Arable Nodes	Frequency	Livestock Nodes	Frequency	Semi-Natural Nodes	Frequency
Excess Nutrients	1251	Compaction	1132	(Δ)pH	533
Ecological diversity	375	pH	630	Metal Contamination	373
Plough Pan (Compaction)	214	SOM	233	Water Quality (Excess Nutrients)	54
		Slope	117		
SOM	41	Texture	82		
Aggregate stability (Compaction and SOM)	11	P, N Surpluses (Excess Nutrients)	23		
		Micronutrients (& Contamination)	13		
		Soil Depth	2		
		Good	23	Good	293

Fig. 5. Opportunity map for ameliorating SQH in England & Wales. (a) First-ranked factor which if improved would make most difference to SQH. Colours identify and locate factors by means of the key; one arable, 23 livestock and 293 semi-natural sites were in optimum condition and are coloured purple and labelled 'Good'. Nets used to compile this map were used at the locations given in Fig. 4b. Consequently, the range of factors evaluated differs from place to place. (b) Frequency with which each factor ranks first for the net indicated. If a factor such as texture is difficult to change, the opportunity is to change the land-use, e.g. convert arable to grassland under livestock. Text colour corresponds to the legend in Fig. 5a; similar factors such as those dependent on physical condition or place are grouped by colour for convenience. A factor and any amelioration depend on land-use context. Thus, pH and measures to adjust it in arable or livestock soils differs from ΔpH in semi-natural soils.

capture the continuity of nitrate values explicitly and this is a feature that should be investigated in future work. Eliciting such a distribution does, however, does make matters harder for the experts.

The second limitation is the diluting effect of intermediary nodes in certain subcases. Marcot (2017) explicitly states that i) summary intermediary variables should be used to help structure a BN and ii) outcomes are not necessarily most sensitive to nodes that are closest to them. However, in our sensitivity analysis when looking specifically at the impact of nodes for which known data is available, it is often the case that nodes closest to SQH have the most influence (Fig E2). For example, Ecological diversity in the arable network. This is not a limitation of the modelling *per se*, but rather in the interpretation, so one should be aware of the need to compare like for like, that is, if we can directly observe a process we should, as opposed to observing proxies.

A possible third limitation is that BNs do not capture the dynamic nature of many soil processes explicitly and there are circumstances (draining of water from poached land for example) where that dynamic nature is key. This said, we asked our experts to be mindful of the dynamic nature of some nodes or interactions when devising and completing their nets.

Existing methods of inferring SQH such as indices (radar plots, Figure E1) or even visual assessment (despite its merit of directness) tend to be descriptive or do not integrate the separate components of SQH explicitly. Furthermore, these indices are normally used locally, on site, by the land-owner or manager, and it is not easy to see how to scale them up or down. Our approach captures both intrinsic soils data as well as management decisions in a transparent and traceable way and can be used at different scales of interest. The same tool is capable of use at both national (Fig. 4 and 5) and local scales: the plans of the Broadbalk experiment (Fig. 3) illustrates how the components of the arable net (Fig. 2a) can be used to distinguish impacts of land management locally.

Most practitioners would emphasise the importance of soil organic matter (SOM) in determining SQH, but research has sometimes failed to find strong links (Corstanje et al., 2015). Our nets suggest why this might be so: many factors depend on SOM and these intermediate factors influence SQH in different ways and moreover interact in ways which differ because of still other factors such as land-use. Where a parent node (such as *SOM*) spawned many children, we encouraged experts to articulate these effects explicitly, the result of which is to push influential nodes such as *SOM* further away from the SQH node with multiple pathways between. In the semi-natural net SOM assumes far greater importance in obtaining good SQH (Fig. E2c) suggesting that its role is greater in regulating than provisioning services. Indeed, the experts originally conceived of this node as *Carbon Storage*.

It is interesting to pursue the structure of the nets further. *Texture* is important as a measurable node in the arable and livestock nets (Fig. 2a & b) but not in semi-natural (Fig. 2c). Our experts seem to be telling us that semi-natural land acquires the appropriate vegetation for the given texture. Under agriculture it appears to be that the productivity of the chosen vegetation (and hence inferred SQH) is restricted by texture. Where texture matters as in agricultural soils it has multiple, indirect effects rather like SOM. It interacts with SOM (and other things) to affect the physical condition of soil which impacts regulation and productivity in turn. Thus, texture in semi-natural soils will affect the appearance of the landscape but the quality less so, whilst under agriculture the landscape is imposed by man and so SQH is thought by our experts to be determined, at least in part, by how well the soil copes with its somewhat unnatural land-use. Thus, nodes such as *Texture* and *SOM* are important, crucial even, but affect SQH indirectly and sometimes in diverse or even conflicting ways.

If a policy aim is to improve land significantly (Defra, 2018) then confidence in metrics that detect such change would be improved by better and finer scale monitoring and measurement. For example, our use of GB Nitrate Vulnerable Zones to assess *Excess Nutrients* in the arable net is rather imprecise. Likewise, and provided data of sufficient quality and resolution is available, it may be possible to detect and even warn of unsustainable practices before lasting damage occurs to soil. In this context, absolute precision in measurement is not as important as credibly detecting signs of change, positive, negative or none, enabling simpler measurement and the possibility of pre-emptive action.

Few soils could not be improved by one measure or another (Fig. 5) or by a change of land-use. This is not to say that all UK soils are in a parlous state but that most could be made at least slightly better and some, in agricultural use, substantially better (Fig. 4). Only one, 5×5 km arable pixel is suggested to be of *Good* quality. Fig. 5 makes it clear that managing nutrients effectively is vital for the quality of the majority, and ought to prevent off-site pollution.

5. Conclusions

The explicit visualisation of the strength of interactions between components of soil that lead to good SQH is a powerful improvement on a simple index. The amalgamation of important subjective and objective determinants in BNs is an elegant solution to what is otherwise an intractable but vital problem in the international scientific literature and in soil management. It is not only the value of a property such as pH that determines SQH, but rather its context (e.g. Wade et al., 2022); here not only arable, livestock farming or semi-natural land-use but also whether the same soil is naturally acidic or rich in organic matter, for example.

The power of our approach derives from quantitative representation of the influence of objective, measurable information on the value of SQH (represented by the area of the SQH node and size of data node in Fig. 2) and the strength of interactions between components (thickness of arrows).

Directly addressable pressures, such as compaction by livestock, or the application of excess nutrients to arable land, are key to avoiding degradation or improving the quality of soil over large parts of the UK. This is widely acknowledged of course, but our analysis stresses the other issues that need to be addressed alongside the obvious interventions, or the role of inherent properties such as texture in their context, if amelioration is to succeed. Measures that are often thought to be obvious steps to improving SQH, such as increasing soil organic matter, may improve soil in an indirect fashion or in different ways depending on factors such as management or land-use.

These arguments apply to most soils under similar land use in the temperate regions, so could form the basis for monitoring of the effects of management practices driven by legislative-regulatory compliance processes in support of Global initiatives to improve soil. Importantly, *direction of travel* to better or worse condition is easily identified and may offer the possibility of *triage* to identify which areas or practices need to be tackled first (Fig. 5). More widely, nets for use in tropical, boreal or montane environments, among others, would require re-elicitation because of the different stakeholders and conditions. Beyond SQH, the method is general and should be applicable to many other metaphors in need of quantification such as sustainability, resilience, air and water quality, human health or well-being.

Author contributions

APW and RC conceived, wrote the proposal and led the research. APW wrote the manuscript with major inputs from KLH & JAH. All senior authors contributed. KLH developed the BNs and modelled the interactions. Together with AGD and JZ she populated the nodes with the data. All senior authors apart from AMK, MJG and SPM contributed to the design of the initial survey and workshop protocols. MJG provided that data for the evaluation of the net at field-scale. SPM contributed a critical analysis of the sources and effects of metals. AGD carried out the initial interviews, devised the naïve nets and organised the workshop. JZ and KLH produced the maps from the BNs. AMK, RC, JAH & AEM suggested, provided or pre-processed much of the data. APW, JZ, JAH, AEM, RC, AGD led elicitations during the workshop. KLH superintended the acquisition and processing of elicited data during the workshop. All junior authors together with AMK, MJG and SPM were invited experts who expressed their beliefs from which the BNs were constructed.

CRediT authorship contribution statement

Glendining Margaret J.: Methodology, Data curation, Invited expert. **Rees Robert M:** Resources. **McGrath Steve P:** Methodology, Data curation. **Matthew Shepherd J:** invited expert. **Todman Lindsay C.:** Methodology. **Noble Nicola L:** invited expert. **Arnold Philippa:** invited expert. **Zawadzka Joanna:** Software, Methodology, Investigation. **Bennett Amanda:** invited expert. **Milne Alice E:** Methodology, Investigation. **Stockdale Elizabeth A:** invited expert. **Whitmore Andrew P:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Tipping Edward W:** invited expert. **Alexander Paul:** invited expert. **Hassall Kirsty L:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation. **Crotty Felicity V:** invited expert. **Dailey Gordon A:** Methodology, Investigation, Data curation. **Horrocks Claire A:** invited expert. **Keith Aidan M.:** Methodology, invited expert. **Bhogal Anne:** Resources. **Corstanje Ron:** Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Clark Joanna M:** Resources. **Harris James A:** Writing – original draft, Methodology, Investigation, Formal analysis.

Declaration of Competing Interest

The authors have no interests to declare

Acknowledgements

This research was funded from the Soil Security Programme grant no NE/P014313/1. The contributions by APW, KLH, AEM AGD & LCT were part funded by (i) the Natural Environment Research Council (NERC) under research programme NE/N018125/1 ASSIST – Achieving Sustainable Agricultural Systems www.assist.ceh.ac.uk. ASSIST is an initiative jointly supported by NERC and the Biotechnology and Biological Sciences Research Council (BBSRC) and (ii) S2N (BBS/E/C/000I0330). Countryside Survey data is owned by UK – Centre for Ecology & Hydrology. APW, LCT, RC and JAH were also part supported by BBSRC responsive mode grant and BB/J000671/1. APW acknowledges part support from the NERC Landscape Decisions 'New Science to Enable the Design of Agricultural Landscapes that Deliver Multiple Functions—AgLand', NE/T001178/1. Rothamsted Research receives strategic funding from the Biotechnological and Biological Sciences Research Council of the United Kingdom. KLH & AEM acknowledge part support from the Growing Health (BB/X010953/1) Institute Strategic Programme. The Rothamsted Long-Term Experiments - National Biosciences Research Infrastructure (RLTE-NBRI) is funded by the UK Research and Innovation – Biotechnology and Biological Sciences Research Council (UKRI-BBSRC) under award BBS/E/RH/23NB0007 (2023–2028). The RLTE-NBRI is also supported by the Lawes Agricultural Trust. We thank Bruce Pearce of the Organic Research Centre who contributed to the arable discussions but whose scores are not included in the final net.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eja.2025.127537.

Data availability

Data will be made available on request.

References

Aalders, I., 2008. Modeling land-use decision behaviour with Bayesian Belief Networks. Ecol. Soc. 13.

Almond, R., 2022. RNetica: R interface to Netica(R) Bayesian Network Engine. R package version 0.5-1. ⟨http://pluto.coe.fsu.edu/RNetica⟩.

Bardgett, R.D., van der Putten, W.H., 2014. Belowground biodiversity and ecosystem functioning. Nature 515, 505–511. https://doi.org/10.1038/nature13855.

Baveye, P.C., 2020. Soil at a crossroad. Soil Use Manag. 37, 215–219. https://doi.org/10.1111/sum.12703.

Bui, E.N., Lougldead, A., Corner, R., 1999. Extracting soil-landscape rules from previous soil surveys. Aust. J. Soil Res. 37, 495–508.

Bünemann, E.K., et al., 2018. Soil quality – a critical review. Soil Biol. Biochem. 120, 105–125. https://doi.org/10.1016/j.soilbio.2018.01.03.

Carey, P.D., Wallis, S., Chamberlain, P.M., Cooper, A., Emmett, B.A., Maskell, L.C., McCann, T., Murphy, J., Norton, L.R., Reynolds, B., Scott, W.A., Simpson, I.C., Smart, S.M., Ullyett, J.M., 2008. Countryside Survey: UK Results from 2007. NERC/Centre for Ecology & Hydrology, 105pp. (CEH Project Number: C03259).

Clemen, R.T., Winkler, R.L., 1985. Limits for the precision and value of information from dependent sources. Oper. Res. 33, 427–442.

Corner, R.J., Hickey, R.J., Cook, S.E., 2002. Knowledge based soil attribute mapping in GIS: the expector method. Trans. GIS 6, 383–402.

Corstanje, R., Deeks, L., Whitmore, A.P., Gregory, A.S., Ritz, K., 2015. Probing the basis of Soil Resilience. Soil Use Manag. 31, 72–81.

Defra 2018 ⟨https://www.gov.uk/government/publications/25-year-environment-plan⟩.

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

El Behairy, R.A., El Arwash, H.M., El Baroudy, A.A., Ibrahim, M.M., Mohamed, E.S., Kucher, D.E., Shokr, M.S., 2024. How can soil quality be accurately and quickly studied? A review. Agronomy 14 (8), 1682. https://doi.org/10.3390/agronomy14081682.

Emmett, B.A., Reynolds, B., Chamberlain, P.M., Rowe, E., Spurgeon, D., Brittain, S.A., Frogbrook, Z., Hughes, S., Lawlor, A.J., Poskitt, J., Potter, E., Robinson, D.A., Scott, A., Wood, C.M., Woods, C., 2016b. Topsoil mineralisable nitrogen (mineral-N) data 2007 [Countryside Survey]. NERC Environmental Information Data Centre. https://doi.org/10.5285/3bafb72b-9f2a-4cbc-a7b8-46e3731c6759. Electronic Rothamsted Archive http://www.era.rothamsted.ac.uk/.

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., Cosby, J., Emmett, B.A., 2021. Development of soil health benchmarks for managed and semi-natural landscapes. Sci. Total Environ. 886, 163973. https://doi.org/10.1016/j.scitotenv.2023.163973.

Grime, J.P., 1979. Plant Strategies and Vegetation Processes. Wiley.

Guimarães, R., Ball, B., Tormena, C., 2011. Improvements in visual evaluation of soil structure. Soil Use Manag. 27, 395–403. https://doi.org/10.1111/j.1475-2743.2011.00354.x.

Hanea, A.M., McBride, M.F., Burgman, M.A., Wintle, B.C., Fidler, F., Flander, L., Twardy, C.R., Manning, B., Mascaro, S., 2017. InvestigateDiscussEstimateAggregate for structured expert judgement. Int. J. Forecast. 33, 267279. https://doi.org/10.1016/j.ijforecast.2016.02.008.

Harris, J.A., Evans, D.L., Mooney, S.J., 2022. A new theory for soil health. Eur. J. Soil Sci., e13292 https://doi.org/10.1111/ejss.13292.

Hassall, K.L., Dailey, A.G., Zawadska, J., Milne, A.E., Harris, J.A., Corstanje, R., Whitmore, A.P., 2019. Facilitating the elicitation of beliefs for use in Bayesian Belief modelling. Environ. Model. Softw. 122, 104539. https://doi.org/10.1016/j.envsoft.2019.104539].

Janzen, H.H., Janzen, D.W., Gregorich, E.G., 2021. The 'soil health' metaphor: illuminating or illusory? Soil Biol. Biochem. https://doi.org/10.1016/j.soilbio.2021.108167.

Johnston, A.E., Poulton, P.R., 2018. The importance of long-term experiments in agriculture: their management to ensure continued crop production and soil fertility; the Rothamsted experience. Eur. J. Soil Sci. 69, 113–125. https://doi.org/10.1111/ejss.12521.

Karlen, D.L., Stott, D.E., 1994. A framework for evaluating physical and chemical indicators of soil quality. In: Doran, J.W., Coleman, D.C., Bezdicek, D.F., et al. (Eds.), Conference: Symposium on Defining Soil Quality for a Sustainable Environment, 35. SSSA special publications, pp. 53–72.

Karlen, D.L., Mausbach, M.J., Doran, J.W., Cline, R.G., Harris, R.F., Schuman, G.E., 1997. Soil quality: a concept, definition, and framework for evaluation. Soil Sci. Soc. Am. J. 61, 4–10.

Kleijn, D., Bommarco, R., Fijen, T.P.M., Garibaldi, L.A., Potts, S.G., van der Putten, W.H., 2019. Ecological Intensification: Bridging the Gap between Science and Practice Trends in Ecology & Evolution, February 2019, Vol. 34, 154-166. https://doi.org/10.1016/j.tree.2018.11.002.

Levontin, P., Kulmala, S., Haapasaari, P., Kuikka, S., 2011. Integration of biological, economic, and sociological knowledge by Bayesian belief networks: the interdisciplinary evaluation of potential management plans for Baltic salmon. ICES J. Mar. Sci. 68, 632–638.

Marcot, Bruce, G., 2017. Common quandaries and their practical solutions in Bayesian network Modelling. Ecol. Model. 358, 1–9. https://doi.org/10.1016/j.ecolmodel.2017.05.011.

Maskell, L.C., Norton, L.R., Smart, S.M., Scott, R., Carey, P.D., Murphy, J., Chamberlain, P.M., Wood, C.M., Bunce, R.G.H. and Barr, C.J., 2008. CS Technical Report No. 2/07: Vegetation Plots Handbook. NERC/Centre for Ecology & Hydrology, n pp. (CEH Project Number: C03259).

Norsys Software Corp, 2016.

NSRI, 2001. The National Soil Map of England and Wales 1:250,000 scale. National Soil Resources Institute, Cranfield University, UK. http://www.landis.org.uk/data/natmap.cfm.

Powlson, D.S., 2020. Soil health—useful terminology for communication or meaningless concept? Or both? Front. Agric. Sci. Eng. https://doi.org/10.15302/J-FASE-2020326.

Rickson, J. et al. Indicators of the quality of the physical property of soil - SP1611. Defra, 2012 (http://randd.defra.gov.uk/Default.aspx?Module=More&Location=None&ProjectID=17595).

Ritz, K., Black, H.I.J., Campbell, C.D., Harris, J.A., Wood, C., 2009. Selecting biological indicators for monitoring soils: a framework for balancing scientific and technical opinion to assist policy development. Ecol. Indic. 9, 1212–1221. https://doi.org/10.1016/j.ecolind.2009.02.009.

Ros, G.H., Verweij, S.E., Janssen, S.J.C., De Haan, J., Fujita, Y., 2022. Environmental Science & Technology 56 (23), 17375-17384 DOI: 10.1021/acs.est.2c04516.

Rose, R., et al., 2016. Evidence for increases in vegetation species richness across UK Environmental Change Network sites linked to changes in air pollution and weather patterns. Ecol. Indic. 68, 52–62. https://doi.org/10.1016/j.ecolind.2016.01.005.

Rowe, E.C., Emmett, B.A., Frogbrook, Z.L., Robinson, D.A., Hughes, S., 2012. Nitrogen deposition and climate effects on soil nitrogen availability: influences of habitat type and soil characteristics. Sci. Total Environ. 434, 62–70. https://doi.org/10.1016/j.scitotenv.2011.12.027.

Rutgers, M., van Wijnen, H.,J., Schouten, A.J., Mulder, C., Kuiten, A.M.P., Brussaard, L., Breure, A.M., 2012. A method to assess ecosystem services developed from soil attributes with stakeholders and data of four arable farms. Sci. Total Environ. 415, 39–48. https://doi.org/10.1016/j.scitotenv.2011.04.041.

Schipper, L.A., Sparling, G.P., 2000. Performance of soil condition indicators across taxonomic groups and land uses. Soil Sci. Soc. Am. J. 64, 300–311.

Seaton, F.M., Barrett, G., Burden, A., Creer, S., Fitos, E., Garbutt, A., Griffiths, R.I., Henrys, P., Jones, D.L., Keenan, P., Keith, A., Lebron, I., Maskell, L., Pereira, M.G., Reinsch, S., Smart, S.M., Williams, B., Emmett, B.A., Robinson, D.A., 2021. Soil health cluster analysis based on national monitoring of soil indicators. Eur. J. Soil Sci. 72, 2414–2429. https://doi.org/10.1111/ejss.12958.

Taalab, K., Corstanje, R., Whelan, M.K., Creamer, R., 2015. On the use of expert knowledge in Soil Mapping. Eur. J. Soil Sci. 66, 930–941.

Wade, J. Culman, S.W., Gasch, C.K., Lazcano, C., Maltais-Landry, G., Margenot, A.J., Martin, T.K., Potter, T.S., Roper, W.R., Ruark, M.D., Sprunger, C.D., Wallenstein, M. D., 2022. Rigorous, empirical and quantitative: a proposed pipeline for soil health assessments.

Wagenet, R.J., Hutson, J.L., 1997. Soil quality and its dependence on dynamic physical processes. J. Environ. Qual. 26, 41–48.

Watts, C.W., Clark, L.J., Poulton, P.R., Powlson, D.S. and Whitmore, A.P., 2006. The role of clay, organic carbon and cropping on plough draught measured on the Broadbalk Wheat Experiment.

Wilhelm, R.C., van Es, H.M., Buckley, D.H., 2022. Predicting measures of soil health using the microbiome and supervised machine learning. Soil Biol. Biochem. 164, 108472.

Zwetsloot, M.J., Bongiorno, G., Barel, J.M., Lonardo, D.P., Creamer, R.E., 2022. A flexible selection tool for the inclusion of soil biology methods in the assessment of soil multifunctionality. Soil Biol. Biochem. 166, 108514. https://doi.org/10.1016/j.soilbio.2021.108514.