Putting numbers to a metaphor: Soil Quality, Health or Fitness?

Andrew Whitmore (andy.whitmore@rothamsted.ac.uk)
Rothamsted Research https://orcid.org/0000-0001-8984-1436

Kirsty Hassall
Rothamsted Research

Alice Milne
Rothamsted Research https://orcid.org/0000-0002-4509-0578

Arthur Dailey
Rothamsted Research

Margaret Glendining
Rothamsted Research

Steve P. McGrath
Rothamsted Research https://orcid.org/0000-0003-0952-8947

Joanna Zawadzka
Cranfield University

Jim Harris
Cranfield University

Ron Corstanje
Cranfield University

Aidan Keith
UK Centre for Ecology & Hydrology https://orcid.org/0000-0001-9619-1320

Lindsay Todman
University of Reading https://orcid.org/0000-0003-1232-294X

Joanna Clark
University of Reading

Paul Alexander
Bulrush Horticulture Ltd

Philippa Arnold
National Farmers Union

Amanda Bennett
AHDB

Anne Bhogal
ADAS

Felicity Crotty

Royal Agricultural University
Claire Horrocks
Rothamsted Research
Nicola Noble
National Sheep Association
Robert Rees
Scottish Rural University College
Matthew Shepherd
Natural England
Elizabeth Stockdale
NIAB
E Tipping
Inst of Freshwater Ecology

Physical Sciences - Article

Keywords: soil quality, soil health, metaphors

Posted Date: May 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-477831/v1

Putting numbers to a metaphor: Soil Quality, Health or Fitness?

Hassall, Kirsty, L.[1], Zawadzka, Joanna[2], Milne, Alice[1], E., Corstanje, Ronald[2], Harris, James, A.[2], Dailey, A. Gordon[1] Keith, Aidan M.[3], Todman, Lindsay C.[1,15] & *Whitmore, Andrew P.[1]

*Corresponding Author: andy.whitmore@rothamsted.ac.uk, 0000-0001-8984-1436

KLH: 0000-0003-0033-1136

SMG: 0000-0003-0952-8947

MJG: 0000-0002-6466-4629

AMK: 0000-0001-9619-1320

JZ: 0000-0002-6172-901X

JAH: 0000-0001-9266-4979

LCT: 0000-0001-8984-1436

with

Alexander, Paul.[4], Arnold, Philippa[5], Bennett, Amanda, J[7], Bhogal, Anne[6], Clark, Joanna, M.[8], Crotty, Felicity,V[9], Glendining, Margaret, J.[1], Horrocks, Claire[10], McGrath, Steve, P[1], Noble, Nicola[11] , Rees, Robert[12], Shepherd, Matthew.[13], Stockdale, Elizabeth A[14], Tipping, Edward, W.[3]

[1]Rothamsted Research, Harpenden, Hertfordshire. AL5 2JQ. UK.

[2]Cranfield University, Bedfordshire. MK45 0AL. UK

[3]UK centre for Ecology & Hydrology, Lancaster, LA1 4AP. UK.

[4]Royal Horticultural Society, Wisley, Woking, Surrey, GU23 6QB. *Current address* Bulrush Horticulture Ltd, Newferry Road, Bellaghy, Magherafelt, Co. Londonderry, N. Ireland, BT45 8ND.

[5]National Farmer's Union, Stoneleigh, Warwickshire, CV8 2LG. UK

[6] ADAS Gleadthorpe, Mansfield, United Kingdom NG20 9PD

[7]Agriculture & Horticulture Development Board, Stoneleigh Park, Kenilworth, Warwickshire, CV8 2TL, UK.

[8]Soil Research Centre, School of Archaeology, Geography and Environmental Science, University of Reading, Reading, UK, RG6 6D orcid: 0000-0002-0412-8824

[9]Royal Agricultural University, Stroud Road, Cirencester, GL7 6JS. Orcid ID 0000-0002-7363-224X

[10]Rothamsted Research, North Wyke, Okehampton, Devon EX20 2SB

[11]National Sheep Association. The Sheep Centre, Malvern, Worcestershire, WR13 6PH

[12] Scotland's Rural College, West Mains Road, Edinburgh EH9 3JG, UK. orcid: 0000-0003-1348-8693

[13]Natural England, Sterling House, Dix's Field, Exeter EX1 1QA orcid: 0000-0001-8616-963X

[14]NIAB, Cambridge, CB3 0LE, UK. orcid 0000-0003-4834-859

[15]University of Reading, Reading, UK, RG6 6D. UK.

Abstract

Soil Quality or Soil Health are terms adopted by the scientific community as a metaphor for the effects of differing land management practices on the properties and functions of soil. Many other terms and metaphors are in use that defy neat quantification: human health, for example. Our challenge is to understand the importance of using such metaphors, but without compromising the underlying scientific understanding upon which they are based. We present here an approach based on expert elicitation in the field of soil quality and management, which offers a universal way of putting numbers to the metaphor. Like humans, soils differ and so do the ways in which they become unhealthy.

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 deduce the value of additional data to the precision of estimates of soil quality and health and infer the likely state of soil at locations in England & Wales. We conclude that the value of soil is best scored as its fitness for purpose or its utility. Our methodology has general applicability and could be deployed elsewhere or in other disciplines.

Main

Soil Quality and Soil Health (SQH) are subjective concepts[1,2] 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. For this reason, SQH is often expressed as an index and used to suggest that this soil or that practice is better than another [3,4]. Such one dimensionality engenders spurious certainty, whereas multi-dimensionality seems specific but can be vague. Given the need to answer the questions 'How much better?' or 'What should be done to improve this soil?', one or more intrinsic properties of a soil may be used to articulate SQH in specific units or in a multi-dimensional index. An example is the well-known radar plot [5,6] to compare several properties together. However, such figures fail to capture interactions between the components that determine the dynamic nature of SQH[7]. Too often, the components are chosen in an *ad hoc* fashion without due consideration as to their functional importance or relevance. Visual assessment overcomes the limitations of single or multiple unconnected indices[8,9] but is not fully objective.

SQH is hard to quantify in a consistent, objective fashion[10]. This has its origin in the complexity and multiplicity of the functions that soil is expected to fulfil and the fact that it harbours a great diversity of organisms[11]. It would undoubtedly be useful if the properties of soil that enable it to deliver substantial crop yields, for example, are the same as those that enable it to buffer water flow, resist erosion[12], regulate greenhouse gases and provide habitat for biodiversity[13,14]. 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 ways. Soils, contexts and functions differ. Soil formation many factors [15], and so soils vary spatially. Humans have adapted agriculture to soil; semi natural systems in the UK are often located in cool climates with acid or nutrient-poor geology. SQH has become a catch-all for the different aspirations of different stakeholders who may interpret a score or value of an index in different ways[6]. Given this background of expectation, context and variation, it is timely to see how far it is possible to operationalise the concept of SQH with methodology that takes account of both qualitative and quantitative components such that each stakeholder gains a metric that is (i) consistent with purpose, (ii) is sufficiently self-consistent with other purposes to be close to universal and (iii) which explicitly captures interactions. We propose that fitness in the Darwinian sense may be a better metaphor and that given the closeness of the concept to the idea of utility in economics, fitness is a more quantitative, contextual measure of the goodness of functioning of soil than quality or health.

Of the attempts to develop 'expert-system' approaches, Bayesian methods are among the most promising means to structure knowledge [16]. Bayesian Belief Networks (or Bayes Nets, BN) are graph-based, directional networks that can incorporate probability distributions of the component variables. They have had diverse application in biological and social sciences[17,18,19]. 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[20] 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[21]. A BN, therefore, gives us not so much an index for SQH as a quantified aggregate of all that experts think is vital.

Notwithstanding the need for universality, a single network for all soils, relevant to all stakeholders would be difficult to achieve. Context, such as agriculture or nature is key. Therefore, we addressed three broad land-uses: (i) arable, (ii) livestock agriculture or (iii) semi-natural land-use. Almost all land in GB 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 tried to retain the ability of each broad land use to express the multi-faceted nature of soil function – production as well as environmental quality, for example.

The three networks (Fig 1a-c) reflect experts' preferences as well as the nature of the land-uses. Consequently, they differ in composition. In all cases, however, the networks were structured such that inference is anchored in measurable data nodes (Fig. 1). Inferences proceed from sink (or child) nodes (ultimately SQH) that are conditionally independent from all other nodes in the network given the connection to the direct source or parent nodes. At most, three nodes parent any one child node, although one parent may influence multiple children. An inference chain builds between SQH and the data nodes. The power of this approach derives from quantitative representation of the influence of objective, measurable information on the value of SQH (area of SQH node and size of data node) and the strength of interactions between components (thickness of arrows). The precise nature of each data node and their sources are given in Supplementary Information. In what follows, we present the findings from the BNs, which in turn are inferred from the views of the experts.

In contrast to the other soils, SQH in arable soil has a direct parent data node – *Ecological Diversity.* Consequently, this node assumes great importance in inferring SQH relative to the other measurable data (Fig 1a, 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, *Productivity Consistency* and *Regulation* are also highly important in signifying arable SQH *(Productivity Consistency* is half as important again as the other two factors which is represented by the arrow thickness in Fig1a. *Productivity Consistency* : *Regulation* : *Ecological Diversity* = 1 : 0.65 : 0.63). Where data, are available for the intermediate nodes (grey discs in Fig. 1), this may be preferred because all upstream parent nodes become unnecessary. Finding above average *Productivity Consistency* is key if the best probability of Good SQH is to be inferred in arable soils (arrow thickness, Fig 1a). If *Productivity Consistency* is Uncertain then 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 S4.1a).

SQH in livestock agriculture can only be inferred reliably from good knowledge of several measurable factors (data nodes), chief among these are the extent of *Compaction* or poaching, the *pH* of the soil and *SOM* – Soil Organic Matter content - (Fig 1b). The complexity of livestock production (soil generates grass, grass supports livestock, but animals impact the soil) 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 1b compared with Figs 1a or 1c). Good *Productivity* is essential for best SQH (Fig 1b & S4.1b), followed by a positive score for *Environment*; *Forage Yield* must be close to Potential.

Soil Moisture explains more variation in SQH in semi-natural soils compared to other data nodes, with *Soil Nitrate* and whether or not the soil is *Bare* almost equally as important as one another (SQH node Fig 1c). Good environmental *Water Regulation* and *SOM* are strongly associated with good SQH in these soils. Note that the Δ*pH* node here represents the difference of the observed pH from the expected pH of a pristine soil (see SI). Although the impact of *Appropriate Chemistry* on *Water Quality* is very large (arrow size Fig 1c) the eventual value of *Water Regulation* on 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).

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 2a). Upland soil is generally in semi-natural land-use whilst many of the livestock soils of intermediate quality are found in the West (Fig 2b). Arable soils appear likely to be worse in the North than the South of the country which is suggested largely by the presence of relatively more non-agricultural species found in the Countryside Survey in the South[22] and so potentially greater levels of *Ecological Diversity*. The risk of *Excess Nutrients* is deduced from context. Even if farmers take action to control nitrate, there is an enhanced risk of poor fitness at locations within a Nitrate Vulnerable Zone 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 1c). Both factors account for much of the geographical variation in values of SQH (Fig 2a)

Where arable land-use predominates in England and Wales, it is likely that more soils should 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 3a). Farmers may, indeed should, already be doing so. Although knowledge about *Ecological Diversity* is the most important directly measurable contributor to good SQH (Fig 1a), *Productivity Consistency* and *Regulation* (into which *Excess Nutrients* feed, Fig 1a) are together vital for the very best quality (ED Fig 1a). 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[23] and the Countryside Survey[22]. Managing nutrients more carefully, by whatever means is thus likely to lead to soils which are fitter in the sense that they are less eutrophic or transmit less to the wider environment.

Many opportunities exist to ameliorate livestock soils that suffer from *Compaction* (1132 instances out of 2063) or where the *pH* is too low (Fig 2). Alkaline soils are difficult to change. Direct interventions to improve *Productivity* or *Environment* (Fig 1b, ED Fig 1b), would help too, in so far as this is possible. Other states of nodes that are associated with poor SQH are difficult (*SOM,* because it increases slowly) or impossible (*Texture, Slope*) to change. If a poor soil is impossible 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, *pH* has improved in recent years[24]. Nonetheless, the discrepancy between actual and expected pH (ΔpH) 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 1c, ED Fig 1c), 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 soil that are above ambient background levels, affects *Appropriate Chemistry* (Fig 1c) and is an issue in about 40% of semi-natural soils. (Fig 3, 286 instances in High exceedance and a further 87 in Moderate exceedance for at least one metal out of 1253 samples where metals data is available.) High metal content in soil may imply mine spoils, deposition from the air or some other historical misuse of land that reduces its function but can also be natural due to 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, however.

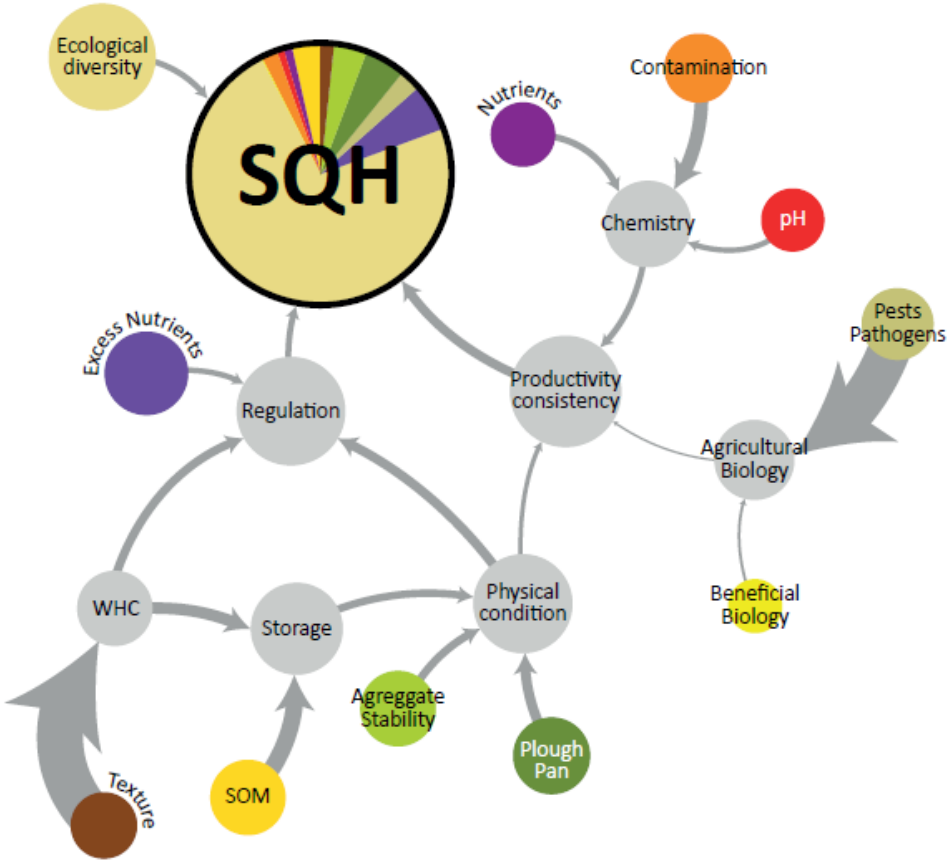
Few soils could not be improved by one measure or another (Fig 3) or by a change of land-use. This is not to say that almost all UK soils are in a parlous state but that most could be improved at least slightly and some, in agricultural use, substantially (Fig 2).

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 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, here arable, livestock farming or semi-natural land-use. A natural ecosystem will be *fitter* if the pH does not differ from what is expected; there is more *utility* in increasing SOM in arable soils above 1% than there is above 3%. Soil *Fitness* would seem a better metaphor than *Quality* or *Health*.

These arguments apply to most soils under similar land use in the temperate regions. Beyond SQH, the method is general and applicable to many other metaphorical terms such as sustainability, resilience, air and water quality, human health or well-being.

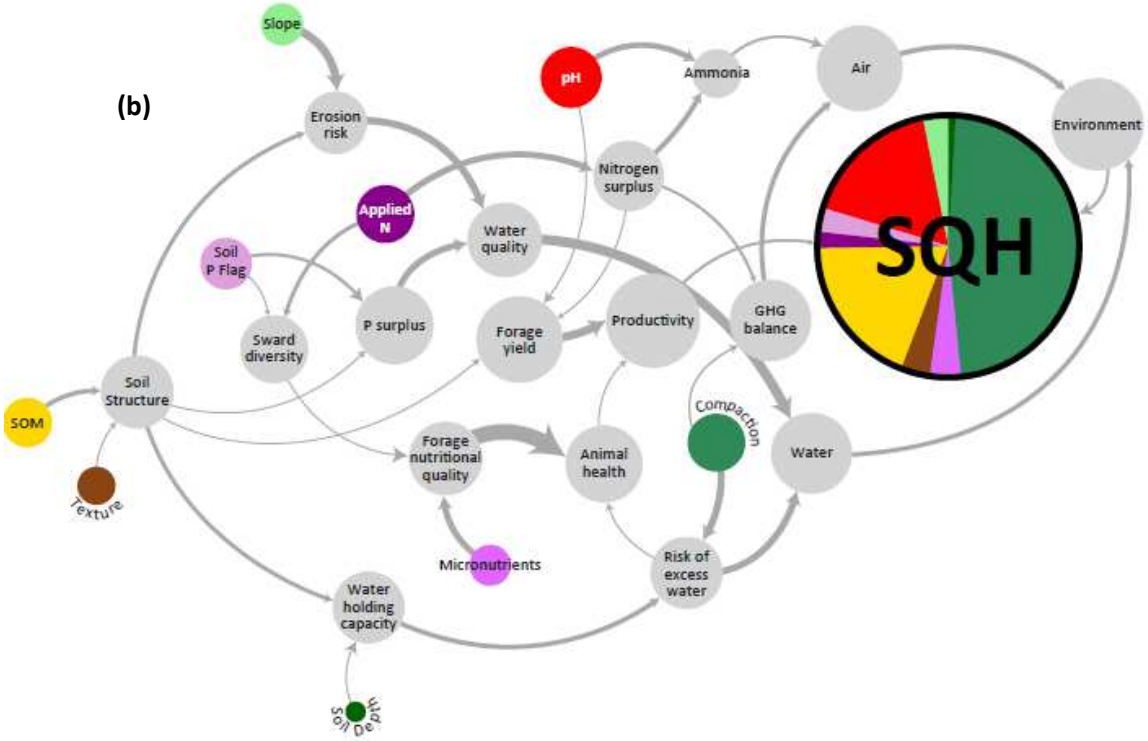
(a)

(a)


Ecological diversity
SQH
Nutrients
Contamination
Chemistry
pH
Excess Nutrients
Regulation
Productivity consistency
Pests Pathogens
Agricultural Biology
Beneficial Biology
WHC
Storage
Physical condition
Agreggate Stability
Plough Pan
SOM
Texture

225

(b)


Slope
Erosion risk
pH
Ammonia
Air
Environment
Nitrogen surplus
Applied N
Water quality
SQH
Soil P Flag
Sward diversity
P surplus
Forage yield
Productivity
GHG balance
SOM
Soil Structure
Compaction
Water
Forage nutritional quality
Animal health
Texture
Micronutrients
Risk of excess water
Water holding capacity
Soil Depth

226

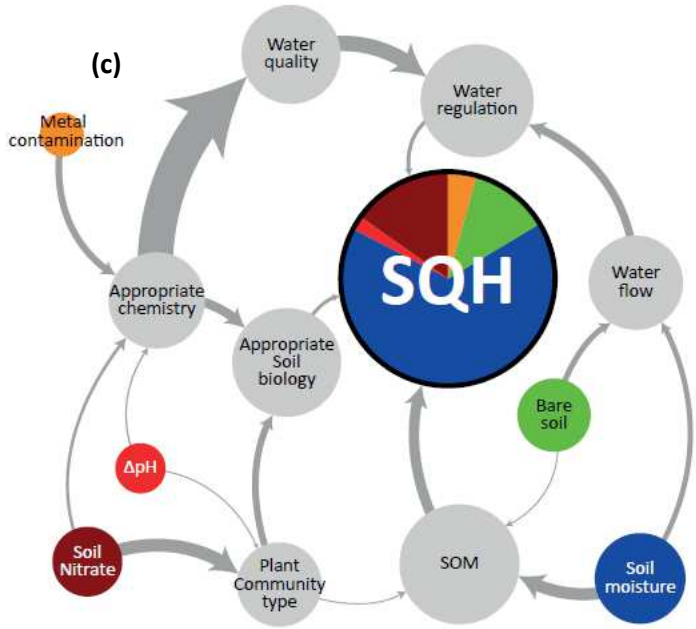


Figure 1: Bayesian Networks that define SQH *of soils under (a) arable, (b) livestock, and (c) semi-natural land-uses. Discs are nodes in all BNs, black-ringed nodes are the SQH end-point. Distinct colours of the sectors in the SQH node correspond to those of the data nodes* i.e. *those which consist 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 nodes represent. 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. 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*

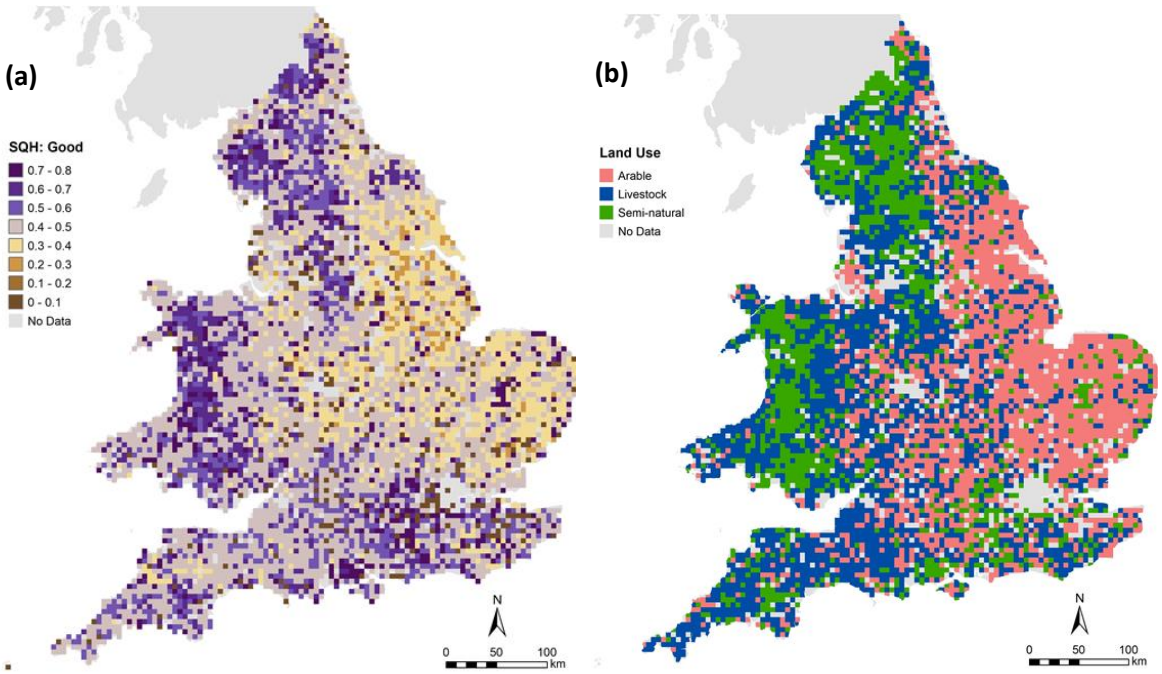
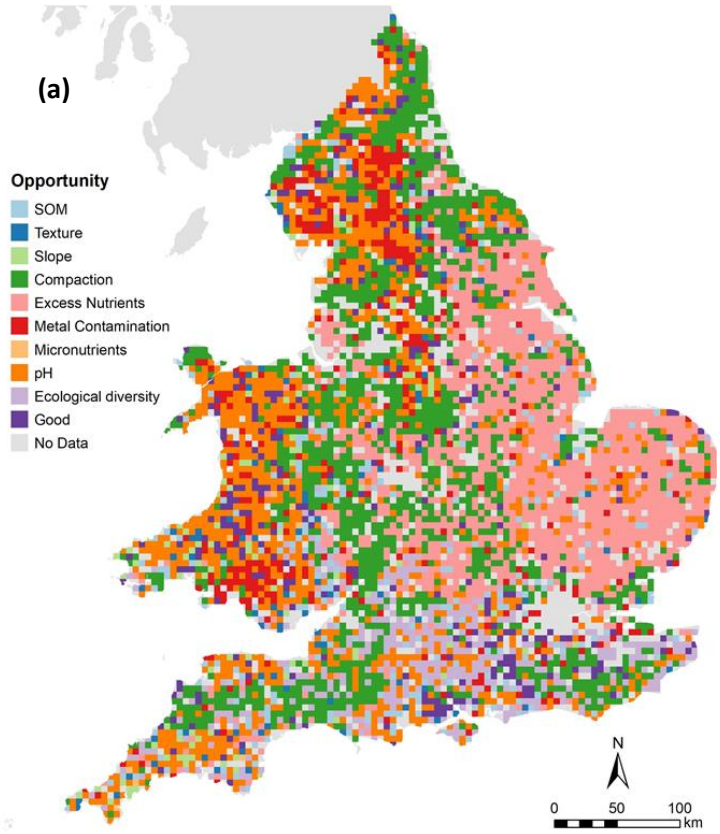


Fig 2: State of SQH in England & Wales: *(a) map of the probability of finding* Good *SQH in England and Wales based on data on land-use, texture and nutrient status in the National Soils Inventory on a 5km grid and on inferred biological parameters derived from the Countryside Survey. All land-uses. The three different land-use nets were applied at locations in (b) and have been used to compile the map in (a).*



(b)

Arable Nodes	Frequency	Livestock Nodes	Frequency	Semi-Natural Nodes	Frequency
Excess Nutrients	1252	Compaction	1132	pH	533
Ecological diversity	375	pH	630	Metal Contamination	373
Compaction	214	SOM	233	Excess Nutrients	54
Texture	13	Slope	117		
SOM	12	Texture	82		
Aggregate stability	11	Excess Nutrients	23		
pH	10	Micronutrients	13		
Metal Contamination	4	Soil Depth	2		
Good	1	Good	23	Good	293

Fig 3: **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 2b. 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 is difficult to change e.g. texture, the opportunity is to change the land-use, i.e. convert arable to grassland under livestock.*

References

1 Janzen, H.H, Janzen, D.W. & Gregorich, E.G (2021) The 'soil health' metaphor: illuminating or illusory? Soil Biology and Biochemistry. https://doi.org/10.1016/j.soilbio.2021.108167.

2 Powlson, D. S. (2020) Soil health—useful terminology for communication or meaningless concept? Or both? *Frontiers in Agricultural Science and Engineering.* https://doi.org/10.15302/J-FASE-2020326

3 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 Science Society of America Journal 61, 4-10.

4 Karlen, DL and Stott,D E (1994) A framework for evaluating physical and chemical indicators of soil quality Eds by: Doran, JW; Coleman, DC; Bezdicek, DF; et al.Conference: Symposium on Defining Soil Quality for a Sustainable Environment, SSSA special publications 35, 53-72.

5 Kleijn, D., Bommarco, R., Fijen, T.P.M., Garibaldi, L.A., Potts, S.G., and 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

6 Rutgers,M., van Wijnen, H,J., Schouten, A.J., Mulder, C., Kuiten, A.M.P., Brussaard, L. and Breure, A.M. (2012) A method to assess ecosystem services developed from soil attributes with stakeholders and data of four arable farms. Science of the Total Environment 415: 39-48. [doi:10.1016/j.scitotenv.2011.04.041]

7 Wagenet, R.J & Hutson, J.L. (1997) Soil Quality and its Dependence on Dynamic Physical Processes, Journal of Environmental Quality 26, 41-48.

8 Schipper LA and Sparling GP (2000) Performance of soil condition indicators across taxonomic groups and land uses Soil Science Society of America, Journal 64, 300-311

9 Guimarães, R., Ball, B & Tormena, C. (2011). Improvements in visual evaluation of soil structure. Soil Use and Management. 27. 395 - 403. 10.1111/j.1475-2743.2011.00354.x.

10 Bünemann, E.K., et al., (2018) Soil quality – A critical review. Soil Biology and Biochemistry 120 105-125 doi.org/10.1016/j.soilbio.2018.01.03

11 Bardgett, R.D. and van der Putten, W.H. (2014) Belowground biodiversity and ecosystem functioning Nature 515,505-511. doi:10.1038/nature13855

12 Rickson 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)

13 Doran, JW and Zeiss MR (2000) Soil health and sustainability: managing the biotic component of soil quality Applied Soil Ecology 15, 3–11

14 Black, H.I.J. Ritz, K. Harris, J.A. Cameron, C.M. Campbell, C.D. Chamberlain, P.M. Creamer, R. Pawlett, M. Wood, C. and Singh, B.K. (2011) Scoping biological indicators of soil quality Phase II. Defra Final Contract Report SP0534. Defra, 237pp. (Defra Project no: SP0534)

15 Jenny, H. (1941). Factors of soil formation—a system of quantitative pedology. Dover Earth Science. 1994 ed. pp281.

16 Bui, E.N., Lougldead, A. & Corner, R. 1999. Extracting soil-landscape rules from previous soil surveys. Australian Journal of Soil Research, 37, 495–508.

17 Aalders 2008. Modeling Land-use Decision Behaviour with Bayesian Belief Networks. Ecology and Society 13.

18 Corstanje, R., Deeks, L., Whitmore, A., Gregory, A.S., Ritz, K., Probing the basis of Soil Resilience. Soil Use and Management, 31, 72-81 (2015).

19 Levontin, P., Kulmala, S., Haapasaari, P. and 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 Journal of Marine Science. 68, 632-638.

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

21 Taalab, K., Corstanje, R., Whelan, M.K., Creamer, R., On the use of expert knowledge in Soil Mapping. European Journal of Soil Science, 66, 930–941 (2015)

22 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).

23 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.

24 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. Ecological Indicators 68, 52-62 https://doi.org/10.1016/j.ecolind.2016.01.005

Methods

Notwithstanding our conclusion that fitness is a suitable metric, we persist with SQH because that is what we asked experts to consider.

Building the Bayesian Networks (BN)

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
(v) interfacing BNs with measured data wherever possible.

Steps i-iv required experts and a setting such as a workshop in which to elicit their opinions. The protocol below evolved from a series of practice workshops that we held to hone our procedure. 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 enthusiasm. The experts ranged from soil scientists and soil surveyors to policy makers, land-owners and managers, farmers and growers. Five to eight experts is considered the optimal number [25] 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).

Roughly one month prior to the workshop we held 15–30 minute one-to-one interviews via video conference with each of the invitees. These video conference 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 SQH?* 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 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. 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 categories. 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 [26] is designed to make completing these conditional probability tables as straightforward as possible. Next, experts were encouraged to think how these functions and processes link together, to infer SQH. In order to maintain parsimony, the experts developed generic nodes whose values were context dependent: *appropriate* biodiversity, *adequate* nutrient levels and so on (See Supplementary Information). 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) semi-natural, (ii) arable, (iii) livestock.

On day two (having practiced the procedure on day one) each group of experts derived the structure of the BN for their allocated land-use type. These BNs became the networks displayed in Fig. 1. With the structure in place we set about eliciting the Conditional Probability Tables (CPT) using our bespoke ACE software [26]. A laptop, networked to the app, was provided to each expert, so that their individual choices could be captured throughout the elicitation process.

In the literature there are two approaches to combining data 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 [27, and references therein] whilst Talaab et al [21] argue that *… any model should go through several iterations until all 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. For specific analytics the individual CPTs were combined (described below).

Following the workshop, the nets and CPTs were incorporated into BNs using the Netica package [28,29]. The CPTs were sense checked and because voice recorders were used to capture discussion it was possible check the intentions of experts for certain issues. We queried odd-seeming views with experts by video conference using open questions so as to not to bias their responses.

The methodology is updateable: views from other experts can be incorporated once elicited.

Mapping SQH

Inference in each net (Fig. 1) begins from measurable quantities (data nodes) to an inference of SQH. 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. Sources of data and the ways they have been used are given in the Extended Data section. Whilst the maps derive from national datasets (Figs 2 & 3), values of SQH can be made at specific locations if a land-manager has local information or a regulatory body mandates measurement. Indeed, local measurements should almost always trump default values. Site specific data could be subjective, such as the relative productivities of the target field compared with any nearby.

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, if amelioration is to succeed

Network analytics

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. 1 through arrow width, node size and SQH pie sector, respectively, with the fourth depicted in the associated heatmaps (Fig. 2). These procedures are described below.

Note, all nodes had an additional parent to that depicted in Fig. 1, specifically, an elicitee (expert) node capturing each individual's CPT.

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

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

i) <u>Arrow Width in Figure 1</u>

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

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

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), the marginal probabilities are obtained by summing over the parental states, i.e.

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

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

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

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

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

and is a measure of the amount 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. 1 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) <u>Node Sizes in Figure 1</u>

To investigate how different nodes affect the probability of SQH, we ran a large simulation study 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 extended data for a full definition of all node states). Given any status, the posterior probability of SQH can be calculated (Equation 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.

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:

Belief(SQH) = Soil_Nitrate + pH + Plant_Community_type + Soil_moisture + Bare_soil + C_Storage
+ Metals_contamination + Appropriate_chemistry + Appropriate_Soil_biology
+ Water_quality + Water_flow + Water_regulation (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 in the fitted regression model and represents the amount of variation in SQH explained by the term. The diameter of each node in Fig. 1 is determined by $\ln(F_{node})$.

iii) <u>SQH sectors in Figure 1</u>

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 pale blue in Fig. 1. 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 1480352 node combinations and 9 data nodes in the livestock network resulting in 110592 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

Belief(SQH) = Soil_Nitrate + pH + Soil_moisture + Bare_soil + Metals_contamination (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 1 are determined by the F-statistics from model (4) expressed as a proportion out of the sum of all F-statistics from model (4).

References

25 Clemen, R.T. & Winkler, R.L. 1985. Limits for the precision and value of information from dependent sources. *Operations Research*, 33, 427–442.

26 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 *Environmental Modelling.* 122, 104539 doi.org/10.1016/j.envsoft.2019.104539

27 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. *International Journal of Forecasting*, *33*, 267279. https://doi.org/10.1016/j.ijforecast.2016.02.008

28 Norsys Software Corp, 2016

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

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. 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.

Author contributions:

APW and RC conceived, wrote the proposal and led the research.

KLH developed the BNs and modelled the interactions. Together with AGD and JZ she populated the nodes with the data.

APW, JAH and LCT conceived the idea of fitness and utility in soil.

All senior authors apart from AMK contributed to the design of the initial survey and workshop protocols.

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.

All senior authors apart from AMK led elicitations during the workshop.

All junior authors and AMK were invited experts who expressed their beliefs from which the BNs were constructed.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryv9.pdf
- Extendeddata.pdf