

Article (refereed) - postprint

Norton, Lisa; Greene, Sheila; Scholefield, Paul; Dunbar, Mike. 2015. **The importance of scale in the development of ecosystem service indicators?**

© 2015 Elsevier Ltd.

This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>



This version available <http://nora.nerc.ac.uk/511919/>

NERC has developed NORA to enable users to access research outputs wholly or partially funded by NERC. Copyright and other rights for material on this site are retained by the rights owners. Users should read the terms and conditions of use of this material at <http://nora.nerc.ac.uk/policies.html#access>

NOTICE: this is the author's version of a work that was accepted for publication in *Ecological Indicators*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published *Ecological Indicators*. [10.1016/j.ecolind.2015.08.051](https://doi.org/10.1016/j.ecolind.2015.08.051)

www.elsevier.com/

Contact CEH NORA team at
noraceh@ceh.ac.uk

The importance of scale in the development of ecosystem service indicators?

Lisa Norton¹, Sheila Greene², Paul Scholefield¹, Mike Dunbar³.

¹Centre for Ecology and Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP (lrn@ceh.ac.uk), ²Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, ³formerly at CEH Wallingford, now at Environment Agency, Kings Meadow House, Reading, RG1 8DQ.

Abstract

Understanding the interactions between ecosystems and their underlying environmental constraints, the services which they provide, and the people benefiting from those services, are essential for the effective management and sustainability of socio-ecosystems (ecosystems which support and are impacted upon by humans). Ecosystem service (ES) indicators attempt to provide a means of measuring service provision, but the scale at which they are developed is likely to impact on how they can be used to influence the effective management of socio-ecosystems. This paper compares science and practice in the development of service measures at contrasting scales in: a) an active research project, focused on local catchment management to improve water quality at Loweswater in the English Lake District, and b) a science-based study developing national scale indicators of water quality using the Countryside Survey dataset.

The paper explores different approaches taken towards the production of ecological measures, which inform on either single or multiple ES delivery across the land/water interface, dependent on scale. It considers how scale impacts on the process of gathering data and on the types of data which can contribute to ES indicators. It further reflects on how service indicators representing different scales of study may be used and by whom. Local scales, in this case the catchment scale, provide a valuable socio-ecological unit for exploring ES delivery, but the extent to which ecosystem service indicators may be used by local actors is uncertain. Larger scale studies may be confined to single services by virtue of data availability but can provide useful policy tools for targeting action. The paper concludes that 'scale' is an important consideration when developing ES indicators. It also concludes that questions around the utility of such indicators should consider the relevance of scale and how it relates to governance.

Introduction

All ecosystems, both managed and unmanaged have an important role in supporting human well-being (Millenium Ecosystem Assessment, 2005) through the provision of multiple ecosystem services (ES) including, for example: food production, clean air and water and cultural services, such as recreation. These services are themselves underpinned by natural capital: soils, waters, habitats and species. In managed ecosystems, an understanding of the actual and potential impacts that land managers and occupiers have on their environment is vital to achieving sustainable use of natural capital and the ES that flow from it (Daily and Matson, 2008; Swinton et al., 2007). Rather than tackling complexity by reducing the scope of our science, the challenge for scientists is to open out our science and address the inherent complexities of socio-ecosystems; ecosystems which support and are impacted upon by humans (Carpenter et al., 2009; de Lange et al., 2010). Science seeking to support sustainable resource management needs to be able to produce indicators that reflect such complexity.

Scientists have employed a range of approaches to try to take on board the inherent complexity of socioecological systems when developing indicators (McVittie et al., 2015). Such approaches attempt to measure both indirect and direct human impact on natural capital through for example, land management, or diffuse pollution impacts on waterbodies. Through measuring such impacts the potential for improving management to enhance the production of ES is increased. A number of key factors are likely to be important for measuring and managing human-ecosystem interactions. This paper considers the relevance of 'scale', and factors associated with it, for the development of ES indicators.

Scale

Identifying scales which are relevant to both social and natural systems is difficult, especially in a globalised world where the goods produced by one part of the world may be consumed by humans from another, e.g. tea produced in India and drunk in the UK, or lamb produced in northern England and consumed in the south-east of England. Inevitably the larger the scale of assessment, the coarser the resolution of data and hence the resultant information/advice on resource management that can be produced. For resource management, an essential consideration in relation to scale is that of environmental governance (Görg 2007; Padmanabhan and Jungcurt 2012). Most land is managed at parcel (field) scales with broader governance structures, where they exist, effective at local to national scales. Governance with an ES agenda may be increasingly widespread following the MA (Millenium Ecosystem Assessment, 2005) but little research has been done to inform land managers (and those institutions with the responsibility for ensuring large scale ES

provision) about optimal landscape scale management for the delivery of ES (Hassan et al. 2005).

The scale at which science is carried out tends to be prescribed by the constraints imposed by funding mechanisms and their required deliverables. Most natural science studies focus either on the micro-scale, and study a reduced set of variables with relatively high control, or focus on the landscape scale using large amounts of data collected over a wide range of sites to identify effects/trends (Bilotta et al., 2010; Boix-Fayos et al., 2009; Collins et al., 2007). The potential value of socio-ecological landscape scale approaches, which take into account human impacts on and use of ES, is increasingly recognised by natural scientists and others (Potschin and Haines-Young 2006, Görg 2007). Despite this, many gaps remain in our understanding of the interactions between ecosystems and with humans because relatively little science which incorporates both natural and social science perspectives (interdisciplinary science) is actually done in this area.

Recognition of the lack of studies which could provide valuable insight into how interactions between humans and ecosystems impact on the delivery of services from ecosystems has led to research initiatives which seek to promote interdisciplinary approaches. These include the Rural Economy and Land Use (RELU) programme (Lowe and Phillipson 2006) under which the Loweswater work described here was conducted. Another key pillar of the RELU programme, recognising the role that publics play in environmental decision making (Lowe and Phillipson 2006), was stakeholder engagement. 'Stakeholder' covers a wide range of potential participants in research, from individual land owners to national policy makers and wider publics affected by environmental decision making. Increasingly stakeholder engagement is recognised as an important facet of studies investigating and attempting to address socio-ecological issues (Krueger et al. 2012), reflecting the role of stakeholders in environmental management at multiple scales. In the development of ES measures, understanding human interactions with the environment and the roles of stakeholders therein are intrinsically connected to 'place' (Berkes 2004; Lowe et al. 2009; Fish et al. 2010; Waterton et al. 2015). For example, the provision and enjoyment of clean water in a catchment needs to take into account those impacting on and using the water, rather than assessing land and water quality in that catchment independently of human interactions.

Problems associated with specific localities or 'places' tend to require approaches which reflect the specific social and natural composition of the locality. Methods currently used to derive ecosystem indicators vary widely with some focusing on modelling at field/experimental scales (Lavelle et al. 2014, Williams and Hedlund 2014) and many others at regional/national scales (Maes et al. 2012, Locatelli et al. 2014). The extent to which

indicators developed at large scales translate to particular localities has not been widely investigated. Similarly comparisons to highlight the constraints and benefits of working at different scales, and their impacts on the value of the ES indicators developed, are rare.

This paper seeks to address this issue by carrying out and then contrasting alternative approaches to providing indicators of ES delivery at different scales, focusing particularly on the interactions between land management and water quality indicators, but also incorporating a wider range of service indicators, where practical. The work on which it is based was developed in response to a political and scientific desire to understand the important factors influencing the value of ES indicators related to the scale at which they are produced. The study focuses on data from two 'sites': 1) Great Britain (GB), as sampled by the Countryside Survey, a national long term monitoring programme (Norton et al. 2012) and 2) the Loweswater catchment in the English Lake District, as sampled by the Loweswater Rural Economy and Land Use (RELU) study. The former study takes an integrated ecological sampling approach at a representative sample of GB sites on an 8-10 year cycle (5 surveys since 1978). The Loweswater study was an interdisciplinary, stakeholder engaged active research project to investigate approaches to community catchment management in a catchment suffering from poor water quality. Modelling approaches described here build on previous work exploring the provision of clean water at the two 'sites' using water quality indicators (Dunbar et al. 2010 – Countryside Survey and Norton et al. 2011b - Loweswater). Previous approaches include process based modelling and mixed effects linear models, whilst new approaches described here include the use of Bayesian Network (BN) analysis and Boosted Regression Tree (BRT) modelling. Bayesian networks are now widely used to model environmental systems due to their capacity for integrating multiple issues and investigating trade-offs (Barton et al. 2008, Chen and Pollino 2012), such as those between water quality and food production. Understanding how the production of one ES may trade-off against another is important when making decisions for sustainable ES delivery. In contrast to BN's, Boosted Regression Tree models are a relatively new approach to modelling large and complex ecological data sets. They include features which are highly suitable for typical characteristics of ecological data (complex, non-linear relationships, non-normality, missing data, variable data formats and inter-correlated predictor data).

The new models were used to explore the limits of the available data in order to gain an understanding about the importance of scale and approach, from both modelling and data collection perspectives, to what we can learn about ES indicators. The approaches are compared systematically in terms of how scale impacts on: a) the data required, b) the extent to which the approaches conceptualise the socio-ecosystems under study, c) the

ecosystem indicators produced and on d) the potential relevance of the approaches to land managers, institutional frameworks and policy makers.

Methods

The methods used in this paper consist of two parts. The first (modelling approaches) describes the construction of the two modelling approaches used to provide ES indicators at two different scales. The second (comparative approaches) describes the comparison between the two approaches in terms of scale and associated factors.

1. Modelling approaches

Pre-existing data and previously published models (Dunbar et al 2010; Norton et al. 2011b) were supplemented by additional data and new modelling approaches for two 'sites': Loweswater and Great Britain (GB). The GB model (using Countryside Survey (CS) data) produced an indicator of water quality whilst that for Loweswater produced indicators of landscape quality (as a cultural service, i.e. a pleasant landscape to visit and look at) and farmer income (provisioning service) as well as water quality. Detailed modelling approaches for each of the sites and for comparison between them are described below.

Site - Loweswater

Loweswater is a small lake (0.6 km²), couched within a mixed upland and lowland farmland catchment, also containing woodland, totalling 7.6 km² in area. Over 85% of the catchment is farmed by eight beef and sheep farmers with small scale low-income enterprises. The catchment's sparse population is supplemented with modest numbers of visitors to the area with residential, visitor accommodation and farm buildings occupying approximately 1% of the catchment. Loweswater experiences regular, potentially toxic blooms of blue-green algae resulting from high concentrations of Phosphorus (P) in the water.

The catchment has been the focus of a number of studies post 2004, the first of which was a natural science modelling exercise to identify the potential impacts of farming practices on Loweswater (Maberly et al. 2006). The Loweswater catchment was subsequently the focus of a RELU project (2007-2011) which aimed to: 1) understand the social and environmental perspectives surrounding the provision of the ES of clean water at Loweswater (Norton et al. 2011b) and 2) help to create a mechanism that would enable community and institutional involvement in decision making (collaborative governance) as a basis for long-term ecological, economic and social sustainability within the Loweswater catchment, the Loweswater Care Project (Tsouvalis and Waterton 2011). Within the RELU project the catchment provided an opportunity to try out modelling approaches which incorporated a wide range of expertise from land management to scientific measurement. Thus the

modelling was embedded in the community mechanism aimed at improving water quality and incorporated inputs from both scientists and non-scientists.

Loweswater data:

Table 1 provides a summary of data from Loweswater used in previous (RELU) models and the additional data incorporated into new models described here.

Table 1: Summary of data used in Loweswater models. BBN refers to the Bayesian Belief Network model.

Data used in models (Norton et al. 2011b)	Additional data used for BBN model
Catchment boundary	Presence , extent and condition of landscape features (i.e. walls, hedges, lines of trees and individual trees)*
Catchment land cover and land use (collected in the field)	Condition of inbye land (as measured in standard vegetation plots size 200m ² , on each farm*
Detailed land management information from farmers (stocking, feeding, housing, agri-environment scheme payments).	Farming income aggregated into two crude classes of 'moderate' and 'low moderate' income [†] .
Soil Phosphorus (P) sampling data from farmers' fields	Relief (difference between highest and lowest altitude, by farm)
Information on septic tank use and P loss	
Flow data (for inflows)	
Weather data (rainfall, air temperature, wind speed and humidity)	
Lake water quality data (water temperature, chlorophyll concentration, soluble reactive phosphate, nitrate and silicate and phytoplankton composition and abundance)	

*Data collection approaches followed those for Countryside Survey, see Norton et al. (2012).

[†] 'Moderate' income was considered as income in line with income for similar farms in the North West (from the Farm Business Survey for 2007) at £7,000 per annum. 'Low Moderate' was more than 25% lower than 'moderate'.

Site – GB (Countryside Survey)

Countryside Survey is a large scale long-term monitoring programme which investigates stock and change in habitats, landscape features, vegetation, soils and freshwaters across Great Britain (Norton et al.2012). Data collection takes place across a randomly stratified sample of GB 1km squares which represent 45 ITE Land classes (landscape types) (Norton et al. 2012; Bunce et al 1996). In 2007, 591 squares were surveyed. Following the 2007 survey, the recognised need for national data on ecosystems and their role in the production of ES led to the production of the CS Integrated Assessment (Smart et al. 2010). This piece of work involved scientists working closely with a policy ‘topic’ group (consisting of representatives of national UK government bodies with responsibility for the environment) to recast CS data as indicators relevant to ES provision. This paper focuses on CS freshwater data.

CS data:

CS freshwater data relevant to water quality includes a) actual counts of invertebrate species in sampled headwater streams and b) a suite of physical variables for each stream (plus alkalinity) which are used to predict an expected ‘reference’ macroinvertebrate community at a stream or river site using the River Prediction and Invertebrate Classification System (RIVPACS) software (see Dunbar et al. 2010). This predicted community (b) is then compared to measured stream community (a) for each site, to provide an observed/expected ratio which in an un-impacted site will be close to one. As degradation, associated with human impacts increases, the observed index value fails to meet expectations and the value of the ratio falls below one.

Whilst this work builds on a modelling approach used by Dunbar et al. (2010), which investigated factors impacting on water quality within the CS 1km sample squares themselves, extrapolation to the national scale has required the use of broadly comparable data outside of those squares. Table 2 provides a summary of data used in previous modelling approaches and in the new model described here.

Table 2: Summary of data used in CS models. Explanatory variables for the Boosted Regression Tree (BRT) models are numbered 1) to 10).

Data used in CS 1km ² models (Dunbar et al. 2010)	Data used in national 1km ² BRT models and catchment model (C)
Biological Monitoring Working Party (BMWP) ⁺ invertebrate taxa score – observed/expected	Biological Monitoring Working Party (BMWP) ⁺ invertebrate taxa score – observed/expected

% Improved grassland in the 1km survey squares from CS	1)% Arable, 2)% Improved Grassland, 3) % Urban in 1km square from LCM* and (C)
% woody cover in plots associated with the headwater stream	4)% woody cover along the headwater stream
Cover weighted canopy height in plots associated with the headwater stream	
	5) Slope - over a 1km length centred on the sampling site i.e. from a point 500 m upstream to a point 500m downstream Average slope (C)
	6) Altitude of sampling site Average altitude (C)
	7) Easting and 8) Northing co-ordinates
	9) Strahler ^{&} stream order (1,2 or 3)
	10) Survey year

⁺ Biological Monitoring Workers Party (BMWP) score is an index for measuring the biological quality of rivers using selected families of macroinvertebrates as biological indicators (Armitage et al., 1983).

^{*}LCM is the UK Land Cover Map, a map of UK habitats with data derived from satellite imagery (Morton et al. 2011).

[&]Strahler stream order is used to define stream size. When two **first-order** streams come together, they form a **second-order** stream. When two second-order streams come together, they form a **third-order** stream. CS samples only headwater streams, i.e. streams of Strahler order 1, 2 or 3.

Modelling

Detailed descriptions of historic modelling approaches discussed in brief below can be found in the original publications (Norton et al. 2011b – Loweswater; Dunbar et al. 2010 – CS).

Descriptions of the Bayesian Belief Network and Boosted Regression Tree models relevant to the aims of this paper are provided here. Detailed information on both approaches is included in Supplementary Information (S1).

Bayesian Network (BN) - Loweswater

Bayesian networks are graphical models that encode probabilistic relationships among variables of interest in a network from primary cause to final effect (Chen and Pollino 2012), effectively providing an influence diagram. Model development begins with a conceptual outline of the 'issue', which in this case was to develop an understanding of the joint production of ES at Loweswater. Previous successful process based modelling of water quality in Loweswater which used a series of linked models (Norton et al. 2011b) to relate catchment variables (including land management) to the likelihood of blue-green algal

blooms (indicative of poor water quality) provided a good basis for the BN. The next steps were to incorporate additional information from the catchment, which would enable inclusion of indicators for farm income and cultural service provision for Loweswater to be modelled alongside water quality. In order to avoid over-complicating the network, a pragmatic approach to its development was taken focusing on key areas of human impact.

The Bayesian Network (Figure 1) was created in Netica (Norsys, 2003). Probabilistic relationships were altered to deterministic relationships where there were known dependencies between model nodes (through data collection and modelling). This included known endpoints in relation to farmer income and a good understanding of processes leading to the presence of blue green algae. This meant that rather than being used to provide new information on all three services, the model was primarily being used to provide new measures of cultural services (landscape quality scores), provided by the farmers within the catchment, and information on how the variables within the model jointly contributed to multiple ES delivery. Landscape quality scores were based on values assigned to key catchment characteristics. These values were essentially 'expert' judgement based on previous collaborative work looking at landscape cultural services (Norton et al. 2011a). As required by BN's, values for nodes were aggregated into categories (discretised) rather than entered as continuous variables. This process both simplified the model and diluted statistical accuracy.

Model evaluation and testing was carried out to validate the model structure and the model detail. Previous modelling and data collection in the catchment (Norton et al. 2011b) facilitated this process. The decision node in the network, the 'farmer' node, enabled an assessment of a) how individual farmer practice/ownership influenced the three ES measures and, b) how collectively farmers influenced ES measures across the catchment as a whole. Although the modelling process built on previous stakeholder engagement and interdisciplinary approaches at Loweswater, engagement with catchment stakeholders regarding BN development was only possible at the end of the process.

Boosted Regression Tree (BRT) – Countryside Survey

Boosted Regression Tree models are a machine learning modelling technique which uses a progressive ensemble approach that combines the strength of two algorithms: i) regression trees (models that relate a response to their predictors by recursive binary splits) and ii) boosting (an adaptive method for combining many simple models to give improved predictive performance) (Elith *et al.*, 2008). The models comprised the observed Biological Monitoring Working Party (BMWP) score (Box-Cox transformed, lambda 0.628) as the response variable and the 10 explanatory variables (Table 2) as the potential predictors. The models were used to identify those explanatory variables that may account for, or help explain

trends in the observed BMWP scores. The best-fit models were determined by adjusting values of two model parameters (tree complexity and the learning rate) until model predictive deviance was minimized without data overfitting. The models were initially trained on a subset of the CS 1km squares and tested on the remainder before being extended to the national scale at 1km² and catchment scales. Model performance was evaluated based on the proportion of the deviance explained (pseudo R²), the Pearson correlation coefficient (c) and the root mean square error (RMSE) between fitted and observed data. Residuals were examined using histograms and Shapiro-Wilk tests to test whether predictions follow normal distributions and to confirm model assumptions were met.

Final models were used for prediction to the national scale for each 1km² grid and catchment containing a stream/river of Strahler order 1, 2 or 3 across GB, comprising a total of 152,639 grids. The national model incorporated the catchment of Loweswater, hence results for the catchment boundary, sub catchment boundaries and 1km² grids for Loweswater were generated, each containing a random point on a river for which predictions would be made. The 10 explanatory variables in both models were generated for all prediction areas. In order to produce predicted o/e BMWP values for the unmonitored sites, expected values for BMWP (predicted) were required and these were generated using the 45 ITE land classes as a base. The expected BMWP scores from the CS data (data derived from RIVPACS using real, sampled environmental attributes at each site) were averaged for each land class. This value was used as the predicted expected BMWP values for the randomly generated river sampling site in each unmonitored grid square. Predicted o/e values were calculated by dividing the predicted observed (from BRTs) by the predicted expected (average scores for ITE land classes).

As with the Loweswater work model described here, the BRT models built on previous stakeholder engaged approaches to understanding the requirements of (in this case) policy makers for national measures of ES. A future workshop with policy stakeholders will discuss the value of both models in that context. Model findings, for both models were communicated to the steering group of the new community based project (the Loweswater Care Programme) at Loweswater in June 2014.

2. Comparative approaches

Following the two modelling exercises the approaches were compared systematically. Comparisons focused on how the scale of the different approaches affected: a) the data required, b) the extent to which the approaches conceptualised the socio-ecosystems under study, c) the resultant ecosystem service indicators produced and d) the relevance of the results to sustainable land management. The latter was assessed through engagement with

stakeholders at local, and to a lesser extent, national policy scales. This included presentation of both models to the steering group of the new community based project (the Loweswater Care Programme) at Loweswater. Consideration of the relevance of ES indicators from a national policy perspective refer largely to work on indicator development post CS and to the commissioning of this work by policy makers.

Results

1. Modelling approaches

Detailed model results are provided in S1. Summary results are given below.

Bayesian Network (BN) models – Loweswater

The Bayesian network analysis produced results for all possible combinations of variables across the model. Results for just one of the scenarios representative of conditions in the catchment at the time of data collection (2008) are given in Table 3. These results represent the scores for each of the farmers for each of the three service measures at high levels of hydrologically effective rainfall (HER), the part of precipitation that reaches stream channels as runoff. Probabilities for blue-green algae out of 100 are relatively low for each farmer but around 25% for the catchment as a whole. Variance in probability between farmers (1.22 - 5.22%) reflects management practices in terms of farm inputs and the area of improved land managed in the catchment as well as the known contribution of P from septic tanks associated with each farm.

The highest % for a credible outcome in terms of farmer income (i.e. one which is not in the 'not present' category), as given in table 3, is the known income level for each farmer. Lack of variability in income in the catchment makes it difficult to test for any relationship between impact on blue-green algae and farming income, but given that the 'low moderate' income farmer has the second highest impact on blue-green algae any expected relationship between low income and blue-green algae probability can be discounted.

Landscape quality scores varied quite considerably between farmers in the catchment (from 21.5 out of 100 to 68 out of 100). Smaller farms (or parts of farms within the catchment) tended to have lower landscape quality scores (ANOVA, $F_{1,7}=4.25$, $P<0.08$), despite account having been taken for area in the measure of landscape features. Not all farms followed this pattern with two farms covering similar areas within the catchment (62 and 66 Ha) having landscape quality scores of 65 and 34 respectively. These farms differed primarily in a) the presence of farm buildings within the catchment and b) the extent of hedges and walls.

Linear regression showed that there was no relationship between the probability of blue-green algae and the landscape quality score ($F_{1,7}=2.11$, $P<0.2$). Overall the results indicate

that at Loweswater the production of desirable landscapes is not linked to high income levels or to particular impacts on blue-green algae.

Table 3: Ecosystem service scores by farmer and for the whole catchment (TOTAL) under high levels of hydrologically effective rainfall, (highest probabilities only for blue-green algae and farming income and mean and Standard Deviation for landscape quality score).

Farmer	Blue-green algae (probability %)	Farming income (probability %)	Landscape quality (score)
A	(High) 1.82	Moderate (71.0)	56.1 (±19)
B	(High) 5.22	Moderate (97.0)	55.6 (±19)
C	(High) 4.37	Low moderate (78)	59 (±14)
D	(High) 3.17	Moderate (44)	64.6 (±23)
E	(High) 2.78	Moderate (58)	21.5 (±10)
F	(High) 1.22	Moderate (98)	24.4 (±15)
G	(High) 3.36	Moderate (97)	68 (±14)
H	(High) 2.69	Moderate (56)	34.1 (±10)
MEAN	(High) 3.07	Moderate	47.9
TOTAL	(High) 24.63	N/A	383.3

Boosted Regression Tree (BRT) models – Countryside Survey

The models that showed the best explanatory power indicated that the 10 variables (Table 2 column 2) were significant predictors of observed BMWP at both grid and catchment scales. Predictor variables were inter-correlated, but attempts to simplify the models by removing some correlated predictor variables indicated that it was beneficial to retain all variables, as the model accounted for the interactions. The 1km² grid model explained 86% of total BMWP variance (pseudo R²). The RMSE for the grid based model was 37.42 and c was 0.5 while RMSE for the catchment based model was lower at 35.48 and a higher c of 0.57. Shapiro–Wilk tests provided evidence to show that the residuals were from a normally distributed population, meeting model assumptions.

Percentage of woody cover and degree of topographical slope were the most influential drivers of observed BMWP values at the grid scale. Mean catchment altitude and slope were the most prominent for the catchment scale. The data display significant spatial correlation, as denoted by the easting and northing coordinates. Partial dependency plots for the 10 variables show the non-linear effect of each variable on observed BMWP values (fitted function) after accounting for the average effects of all other variables in the model.

The predicted o/e BMWP values at the national scale are presented in Figure 2. They show a strong southeast/northwest pattern, with higher water quality in western and northern areas (with the exception of coastal areas particularly in west Wales and northwest Scotland), and lower water quality in the east and south across England. Predictions for Loweswater gridded points and sub-catchments are shown in Figures 3 and 4 respectively. The predicted o/e BMWP value for a point just after the lake outlet at Loweswater, taking the data within the whole catchment as explanatory variables, was 1.36. The predicted o/e ratio values range from 0.24 to 2.10. Low values (under 1.0) indicated that the predicted observed BMWP scores were lower than the predicted expected scores, i.e. less families were predicted to have been observed than should have been according to inherent environmental variables that were used to generate the expected BMWP scores (also predicted based on CS expected). Loweswater stream water quality predictions are varied at the 1km scale from 0.69 to 1.14. For sub-catchment areas water quality predictions range from 0.46 to 1.25.

2. Comparative approaches

a) Data requirements

Tables 1 and 2 summarise the data used in the two different scale modelling approaches. Data from Loweswater comprised detailed field and farm level data, collected by professionals and non-scientists including land managers, to ensure the best possible representation of the catchment (from social and natural perspectives). Use of the Bayesian approach resulted in some constraints to the resolution at which data could be used. For example, raw data values used in the process based modelling for water quality (under RELU), were converted to coarser categories of data for the BN.

Data used in the national BRT models built on relationships established through detailed field sampling in 1km squares (Table 2), extrapolated using variables from national mapping products. The data used were natural science only. The large scale of the modelling approach limited the resolution (1km scale) and volume of data that could be used in the model.

b) Conceptualisation of socio-ecosystems

The national BRT models incorporated information on socio-ecosystems by proxy, essentially using land cover type to provide information on land management impacts on water quality. In contrast the BBN approach at Loweswater attempted to conceptualise the socio-ecosystem there, focusing on stakeholder engagement in order to understand their roles in the provision of ecosystem services. Table 4 summarises the key contrasts between the two approaches for providing ES indicators at each of the study sites, in terms of; scale, interdisciplinarity of approaches and stakeholder engagement.

Table 4: Summary of site contrasts in terms of scale, stakeholder engagement and interdisciplinarity of project and/or modelling approaches.

Site	Scale	Stakeholder engagements	Interdisciplinarity
Loweswater	Small catchment (7.6km ²)	Project: Stakeholders included local residents, farmers, representatives of local environmental institutions and scientists. Project results communicated to representatives of national stakeholder institutions. Model: Stakeholder data included.	Project: Social and natural scientists from academia and among local stakeholders working together in the project. Model: Interdisciplinary approach using social and natural data.
Great Britain (CS)	Large (country) (209,331km ²)	Project: Policy stakeholders were involved in the development of indicators from the start of the project. Model: No local stakeholder input	Project: Natural scientists only. Model: Natural science approach.

c) Ecosystem service indicators

Detailed data collected across the Loweswater catchment made it possible to produce indicators of three different services. The catchment scale had direct relevance for water quality and less relevance for farmer income or landscape quality but was generally adequate. Farmer income is reliant on whole farms (rather than the area within a particular catchment) and landscape or farm type. However, it was possible to relate income to land in the catchment, as the majority of land for most farmers was in catchment, with land outside of catchment not adjudged to significantly impact on income. Landscape quality measures were to some extent hampered by the scale, as inclusion of whole farms in the catchment would mean, at least, that farm buildings (an important contributor to landscape quality in the model) would be present on all farms. However it was possible to construct measures of

landscape quality for land belonging to each of the farmers in the catchment, and using the BN, link that to farming income and impacts on water quality.

This work demonstrates that CS data collected at a 1km square scale can be used to model a national water quality indicator, using the relationships identified between landscape variables and water quality in those squares. Catchment is also clearly an important scale for consideration of water quality, hence the BRT was used to produce a catchment model for water quality at Loweswater.

d) Relevance of results to sustainable land management.

When presented to stakeholders at Loweswater the models elicited several responses. One response led to a direct change in the indicator developed from the BRT model from a BWMP score for water quality to a relative observed/expected BWMP score. This provides more useful information on stream condition relative to what might be expected for the area under study. The models also raised a number of questions discussed (in 'Relevance to Users', below). These included: 'which modelling approach is appropriate for us? And, how do we use it?'

National policy stakeholders helped to both fund and shape the work carried out under CS to provide a national assessment of water quality (alongside other ES), based on indicators developed from CS data. Policy stakeholders also supported the work reported here in order to gain a better understanding of a) the issues and constraints associated with using national level data to provide information at local scales and b) the importance of scale for ES indicators. Feedback from policy users in relation to the models presented here has been affected by changes in government and the individuals responsible for commissioning this work, leading to limited interactions between researchers and policy makers. It has also been affected by deep government budget cuts and resultant impacts on staffing and workloads of staff within government departments.

Discussion

Model results

Bayesian Network (BN) models – Loweswater

For water quality, the BN results reflect what is already known from the catchment from previous models and data collection (Norton et al. 2011). Farming negatively impacts on lake water quality by contributing to nutrient loss from the land. Farming impacts on water quality in the catchment are captured in the BN through land cover type and inputs related to stocking, and to a lesser extent through delivery from septic tanks. The BN captures impacts in the form of a probability of high, medium or low likelihood of blue-green algae. Indicators

of farm income from the BN simply reflect known information. Farmer income in the catchment is highly constrained by the nature of the catchment and its extent (and as a direct result, the size of farms in the catchment) (Norton et al. 2014). As a result of this, only substantial increases in either government subsidy (through agri-environment schemes and CAP pillar one payments) or stock prices, or potential amalgamation of all farms in the catchment (at a minimum), could improve farming income.

Landscape quality scores varied substantially between farmers with buildings and the presence of landscape features driving differences. These scores were based on information about cultural variables (which are important to people in their experience of landscapes) from previous work (Norton et al. 2011a). It is acknowledged that scoring could potentially have been improved through consultation with users of the Loweswater landscape. For example, species richness, (although only accorded low impact in this study), may in fact have no impact, whereas the presence of sheep or cattle may be important. From a landscape quality perspective, the whole catchment and its context is more relevant to visitors than individual farms, but like water quality, understanding contribution at a farm level is important in any policy aimed at enhancing services. Another key landscape quality issue would be water quality impacts on landscape quality. No attempt has been made to include that in this model, in part because lake water quality is a catchment rather than a farm issue.

The BN was essentially used as a quantitative influence diagram providing a quantification of how different (largely farmer related) variables in the catchment influenced service indicators. Even though highly rationalised, the BN indicates that multiple service delivery at farm and catchment scale is complex. In its current form there are no clear optimal farm management strategies which maximise delivery of all three services, which is probably a good reflection of reality. At the catchment scale it would be important to consider non-farmed land (woodland and water) and their impacts on services alongside farmed land. Because of its basis on process based models (Evans 2012), the BN lends itself to the consideration of alternative scenarios and their impacts on multiple service delivery. Exposure of the BN to the Loweswater Care Programme steering group highlighted the issue of who might use the BN. This is discussed further below (Relevance to users).

Boosted Regression Tree (BRT) models – GB

Both of the BRT models provided a high explanation of variance of the response variable. Relative to more traditional means of modelling (e.g. multiple linear regression, generalised additive models), studies have shown that BRT models tend to show superior performance (Leathwick et al. 2006; Brown et al., 2012), however fits are subject to user controls during

the modelling procedure (Elith et al., 2008). For example, if it is not appropriately controlled, the risk of overfitting (i.e. allowing the learning algorithm to fit obscure aspects of the training data that are unlikely to improve prediction accuracy) is much higher in machine learning techniques, such as BRTs, than other modelling methods. Hence regularization methods were used in model development to balance model fit and predictive performance. The accuracy of the predicted values (gauged as RMSE and c values) indicated that reasonable and sensible estimates were generated, however although the grid model showed a higher R^2 , its RMSE and c values were higher and lower, respectively, when compared with the catchment model, hinting at a possible slight overfitting of the data in comparison.

The outputs from BRT analysis indicate that there is a high level of association in the relationship between the BMWP score of a river and metrics of land use and biophysical characteristics at both 1km square and catchment scales, and that these relationships show complex, non-linear relationships. Although previous studies have shown that such relationships exist (Dunbar et al. 2010), the analyses here has been able to isolate the relative impacts and interactions of individual explanatory variables, together with looking at the influence of spatial scale studied. Notably, the three metrics of land use collectively influence BMWP scores less in the catchment-scale analysis (14%) when compared with the grid-scale study (24%). In both cases the biophysical aspects of both the grid and catchment boundaries (such as slope, altitude and position north and east) have greater influence on water quality. Differences between results at catchment and 1km scales reflect the extent to which water moves through a landscape and is subject to influences along its passage rather than at any particular point in the landscape.

For the Loweswater catchment, BRT results at both 1km² and catchment scale indicate variability across the catchment in terms of water quality and indicate moderate levels of pollution. Model predictions at the catchment scale indicate particularly low water quality in one sub-catchment which is almost certainly inaccurate. However, for the most part results are broadly in line with measures of lake and stream water quality in the catchment. Variability in the macroinvertebrate community dataset and the structure of the BMWP scoring system may have contributed to uncertainty in the model outputs. Even with the absence of nutrient pollution, the community of macro-invertebrates sampled varies between different sites and rivers as a result of inherent environmental variation, including attributes such as, sediment type, flow, underlying geology, river depth, river width, climate etc. Therefore, the BMWP scores are subject to natural variation that may not be accounted for in the model.

Scale and the development of indicators

a) Data requirements

The contrasting scales of approach shown here presented different data challenges. Modelling at finer scales demands more detailed data in order to predict outcomes effectively; at larger scales statistical patterns become more regular and the use of coarser proxies more rational (Levin 1992). Data provided by the community relevant to income, land and septic tank management data and weather data as well as more general information, such as drainage patterns affecting the lake watershed, are all key parts of the BN model. These data reflect input by social and natural scientists as well as catchment stakeholders, all aided by the small scale of Loweswater (Norton et al. 2011b; Norton et al. 2014). The contrasting (national) scale of the BRT models led to a much more restricted set of (natural science only) data being used in the models. Modelling parameters used were constrained by the relationships identified between stream water quality and measured 1km square level variables (Dunbar et al. 2010), as well as by data available at national scales (to enable extrapolation). A further minor constraint was computing power, as the modelling involved sizeable computing tasks.

b) Conceptualisation of socio-ecosystems

If we recognise ecosystems as socio-ecosystems and acknowledge the vital role that humans play in ensuring their long term sustainability (Berkes and Folke 1998), then we must also recognise that science which attempts to measure ES needs to take approaches which engage with both social and natural parts of those systems (Klein 2004). Despite the recognised need for these interdisciplinary approaches to ecosystem management (Spash 2012, Young et al. 2014), and initiatives which seek to promote these (Lowe et al. 2009), key challenges remain (Schoolman et al. 2012, Cullen and White 2013, Botey et al. 2014). Scale constraints on the extent to which the production of ecological indicators can incorporate both social and natural components of ecosystems relate, in part, to data issues (as discussed above).

There are, however, other constraints including the availability of funding for such approaches and beyond that the 'vision' of potential funders for interdisciplinary approaches to tackle socio-ecological issues. Both of the approaches described here were fundamentally reliant on funders prioritising 'ecosystem service' approaches. A specific call for funded interdisciplinary research provided the opportunity for joint working at Loweswater (Lowe and Phillipson 2006), but such research programmes remain in the minority compared to single discipline research programmes. The development of a water quality indicator using CS data

resulted from working alongside a policy stakeholder group to identify how CS could be used to develop national ES indicators (Smart et al. 2010). The same policy stakeholders provided funding for this comparative work. The scale of the national work limited the extent to which the modelling tool took into account both social and natural perspectives. However to an extent, the 'social' perspective was the policy stakeholder desire for the model and for how its results (in particular land cover) may help to influence policy.

There are a range of benefits of working at a scale that is relevant to a specific community of people and stakeholders, such as Loweswater. Fish et al. (2010) point out that engagement with stakeholders, combined with interdisciplinary research and collaborative governance emphasises reciprocity, relationships learning and creativity. The collaborative governance approach at Loweswater, involving institutional stakeholders and locals, created the capacity for deeper understanding and appreciation of human impacts on Loweswater. In this respect the work echoed other work on water catchments, e.g. The SLIM (Social Learning for the Integrated Management) Project (Collins et al. 2007).

Prior to the LCP a more restricted (purely natural science) project on water quality modelling for the catchment used earth observed data on the catchment land cover and made assumptions on management based on land cover type (Maberly et al, 2006). This exercise failed to engage effectively with farmers or the wider Loweswater community, because it was carried out independently of any local involvement and farmers were unable to link their actions to the data used in the models. The difference made by spending time with farmers and community members to learn about their knowledge of place, and trying to embed that knowledge in the science from the beginning of the project, was that the science then became more relevant and more accessible to them (Waterton et al., 2015). Direct links between farmer and wider community data and catchment use encouraged them to address issues where possible (including changed fertiliser applications, septic tanks management and co-operative vegetation management at the lake outlet). Engagement of farmers and community members also resulted in local action to further understandings of the catchment (through several small further research projects which arose out of the LCP, see Norton et al. 2014). Ultimately the project has led to the successful application for further funding by community representatives. This is enabling a range of practical and research initiatives (relating to water quality) to be undertaken in the catchment. From a scientific perspective engagement changed the scientific process, improved data quality and allowed the science to move further away from measures of natural capital towards ES measures.

c) *Ecosystem service indicators*

The discussion above focuses on the water quality indicator which was incorporated into the BN and was the primary focus of the RELU work. The 'Loveswater' scale was ideal for water quality indicators as it forms a natural water catchment within which stakeholder impacts can be evaluated (Hooper 2005; Mitchell 1990). It is important to note that the exploration of a wider set of ES indicators in the BN modelling, which was carried out following the completion of the RELU project, was not co-developed with the Loveswater Care Programme (although it relied on data collected within the project). Rather, the BN modelling was developed in response to a political and scientific desire for comparing the production of ES indicators at different scales (discussed further below). The Loveswater catchment (for reasons discussed under a) and b) above was a useful scale for producing ES indicators using a BN approach, despite issues surrounding scale mismatches (see results) which may have impacted on their validity. The degree to which common factors influenced all three indicators made them appropriate for the BN model, although necessarily somewhat coarse.

The CS sampling strategy is designed to enable the generation of national level data relevant to ecosystems, e.g. measures of natural capital relevant to ES delivery (Dunbar et al. 2010; Norton et al. 2011a; Maskell et al. 2013). ES indicators at national scales are only ever likely to be indicative rather than absolute, because of the resolution of the data used. Potentially, it may be more accurate to refer to the indicators produced here as indicators of natural capital rather than of ES, given that data used is a proxy for human impacts rather than a direct measure (such as stocking levels or fertiliser application). There is currently no scope for accessing and using the kinds of detailed information collected at the Loveswater catchment scale at national scales. Improvements in the availability of data which is already collected, e.g. farm management data, would enhance the quality of such indicators and the potential for understanding and predicting changes in them. Weaving spatially resolved social data on land management, income or visitor usage together with data on natural capital within national sampling sites, and more broadly across national scales, can only enhance scientific capacity for providing ES indicators and understanding key drivers of change.

d) Relevance to users

This work was developed in response to a political and scientific desire to understand the factors influencing the value of ES indicators related to the scale at which they are produced. The work shows that the scale at which indicators are developed can have a profound effect on the indicators developed. It is possible to produce a national indicator of water quality using data on land cover and land type which can be used to provide relative assessments of where water quality is poor or good on a national scale. This may be useful for targeting funding for action at such scales. It is also possible to scale such data to local catchments

or to a standard unit (in this case 1km²), but the relevance and accuracy of such data at local levels is constrained by a lack of information on human's roles within the ecosystem (Eigenbrod et al., 2010).

It is clear from the discussion above that development of indicators at local scales through a process of engagement may have a range of benefits, from improved scientific accuracy to eliciting local action to address issues negatively impacting on indicators. However, the extent to which such a process can promote action depends on the nature of the specific indicators, how they have been developed and the extent to which locals have the 'power' or 'resource' to affect them. At Loweswater targeting practices towards improving water quality has happened (and is continuing) following on from the models linking catchment management to water quality (Norton et al. 2011). Although the BN presented a more rounded approach to ecosystem delivery in the catchment, the potential for locals to engage with the model and its outputs was constrained by a lack of 'power' or 'resource' to affect outcomes, particularly in terms of largely market driven farmer income. Hence the question, 'how do we use it?' in response to the model being presented to locals. To which, the answer was perhaps, 'you don't, but perhaps it at least allows you to see the bigger picture?'

In the CS project, interactions with policy stakeholders were, as at the local scale, beneficial on both sides. Sharing the scientific process with stakeholders enhanced both the accessibility and understanding of the science and the scientific process for stakeholders (Waterton et al., 2015). Of particular importance was how engagement led to a better understanding by policy stakeholders and scientists of the constraints of the natural science data, in terms of direct links with ES. From a science perspective because the project was focused on data already collected, there was perhaps little change to the scientific process. However, through a series of meetings with policymakers which ran throughout the project, the way the science was communicated and presented changed to become more relevant to a non-scientific audience and to policy targets. Specifically this involved which words were used and how they were understood in the description of measures and uncertainties. It also involved the production of figures, in particular maps, which can provide a very accessible picture of national ES delivery.

Fundamental questions concerning the relevance of scale and stakeholder engaged approaches to the production of ES indicators and the decision contexts surrounding them, include: who are the indicators for? Who is going to make decisions based on the indicators? For national scale policy makers, implementing the ecosystem approach and trying to do so on the basis of available evidence, it is important to understand the importance of scale and approach and the constraints therein. The relevant scale for indicator production depends on

the scale at which they will be used. Case studies and national predictions are useful for policy insights but in order to affect change, translation into relevant indicators at local scales is essential (Cash et al. 2003, Norton et al. 2011b).

Despite its use of local input the BN was effectively developed as a science and policy tool to explore multiple ecosystem service delivery at a catchment scale. A particular challenge from the work is to discover whether some of the advantages of the place-based science can be effectively brought into larger scale investigations (e.g. improved data on land use contributing to the evaluation of multiple services), or whether it is enough to recognise that work at different scales has different uses and advantages. In the longer term, greater awareness amongst the scientists of stakeholder needs will, it is to be hoped, encourage changes in the scientific process which will better meet the needs of society.

Concluding remarks

This work described the development of two models to produce indicators of ES at a national scale for GB (water quality), and at a local scale for a site within GB (water quality, income, landscape quality). A comparison of the two models and the approaches taken in their development provided insights into how scale can influence the production of ES indicators.

The results indicate that the production of ecosystem indicators at national scales is supported by a nationally representative sampling approach which can be extrapolated using measured relationships with variables available at national scales. They also indicate that such indicators may be locally inaccurate due to lack of knowledge about local level variables, but may provide a means of targeting more in-depth studies. Indicators produced from in-depth local level studies benefit from engagement with local stakeholders early on in the process and may be used to direct changes in local level management. Local studies which engage stakeholders provide the potential for development of multiple ecosystem service indicators which may be more relevant for policy decision making at larger scales. Governance structures determine the extent to which ecosystem service indicators can be used for landscape management at different scales.

Acknowledgements

This work was carried out under the 'Ecosystem Interactions' project funded by the UK Department for the Environment Food and Rural Affairs. It uses data collected under the Joint Research Councils RELU funded project 'Testing a community approach to catchment management' and the UK Countryside Survey project, funded by a large number of UK organisations, see website: <http://www.countrysidesurvey.org.uk/about/partners-and-people>

The authors are grateful to Ron Smith at CEH Edinburgh and Inge Aalders at the James Hutton Institute for advice on the use of the Bayesian Belief Network software.

References

Barton, D. N., T. Saloranta, S. J. Moe, H. O. Eggestad, and S. Kuikka. 2008. Bayesian belief networks as a meta-modelling tool in integrated river basin management - Pros and cons in evaluating nutrient abatement decisions under uncertainty in a Norwegian river basin. *Ecological Economics* 66:91-104.

Berkes, F. 2004. Rethinking community-based conservation. *Conservation Biology* 18:621-630.

Berkes F. and Folke C., 1998. *Linking social and ecological systems: Management practices and social mechanisms for building resilience*. New York, NY: Cambridge University Press.

Bilotta, G.S., Krueger, T., Brazier, R.E., Butler, P., Freer, J., Hawkins, J.M.B., Haygarth, P.M., Macleod, C.J.A. and Quinton, J.N., 2010. Assessing catchment-scale erosion and yields of suspended solids from improved temperate grassland. *Journal of Environmental Monitoring* 12: 731-739.

Boix-Fayos, C., de Vente, J., Albaladejo, J. and Martinez-Mena, M., 2009. Soil carbon erosion and stock as affected by land use changes at the catchment scale in Mediterranean ecosystems. *Agriculture Ecosystems & Environment* 133: 75-85.

Botey, A. P., T. Garvin, and R. Szostak. 2014. Interdisciplinary Research for Ecosystem Management. *Ecosystems* 17:512-521.

Bunce, R.G.H., Barr, C.J., Gillespie, M.K., Howard, D.C., 1996. The ITE Land Classification: providing an environmental stratification of Great Britain. *Environmental Monitoring and Assessment* 39, 39-46.

Carpenter, S.R., Mooney, H.A., Agard, J., Capistrano, D., DeFries, R.S., Diaz, S., Dietz, T., Duraiappah, A.K., Oteng-Yeboah, A., Pereira, H.M., Perrings, C., Reid, W.V., Sarukhan, J., Scholes, R.J. and Whyte, A., 2009. Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment. *Proceedings of the National Academy of Sciences of the United States of America* 106: 1305-1312.

Chen, S. H., and C. A. Pollino. 2012. Good practice in Bayesian network modelling. *Environmental Modelling & Software* 37:134-145.

Collins K., Blackmore C., Morris D. and Watson D., 2007. A systemic approach to managing multiple perspectives and stakeholding in water catchments: some findings from three UK case studies. *Environmental Science & Policy* 10: 564-574.

Cullen, R., and P. C. L. White. 2013. Interdisciplinarity in biodiversity project evaluation: a work in progress. *Wildlife Research* 40:163-168.

Daily, G.C. and Matson, P.A., 2008. Ecosystem services: From theory to implementation. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9455-9456.

de Lange, W.J., Wise R.M., Forsyth, G.G. and Nahman, A., 2010. Integrating socio-economic and biophysical data to support water allocations within river basins: An example from the Inkomati Water Management Area in South Africa. *Environmental Modelling and Software* 25: 43-50.

Dunbar, M., Murphy, J., Clarke, R., Baker, R., Davies, C., Scarlett, P. (2010) Countryside Survey: Headwater Streams Report from 2007. Technical Report No. 8/07 NERC/Centre for Ecology & Hydrology 67pp. (CEH Project Number: C03259).

Eigenbrod, F., Armsworth, P.R., Anderson, B.J., Heinemeyer, A., Gillings, S., Roy, D.B., Thomas, C.D., Gaston, K.J. (2010). The impact of proxy-based methods on mapping the distribution of ecosystem services. *Journal of Applied Ecology*, 47, (2), 377-385.

Elith, J., Leathwick, J.R., and Hastie, T. (2008). Boosted regression trees – a new technique for modelling ecological data. *Journal of Animal Ecology* 77: 802-813.

Fish, R. D., A. A. R. Ioris, and N. M. Watson. 2010. Integrating water and agricultural management: Collaborative governance for a complex policy problem. *Science of The Total Environment* 408: 5623-5630.

Görg C., 2007. Landscape governance. The 'politics of scale' and the 'natural' conditions of places. *Geoforum* 38: 954-966.

Hassan R., Scholes R. and Ash N., 2005. Ecosystems and human well-being: Current state and trends. Washington, D.C.: Island Press.

Hooper, B. (ed.), 2005. Integrated River Basin Governance. IWA Publishing, London.

Krueger, T., T. Page, K. Hubacek, L. Smith, and K. Hiscock. 2012. The role of expert opinion in environmental modelling. *Environmental Modelling & Software* 36:4-18.

Lavelle, P., N. Rodriguez, O. Arguello, J. Bernal, C. Botero, P. Chaparro, Y. Gomez, A. Gutierrez, M. del Pilar Hurtado, S. Loaiza, S. Xiomara Pullido, E. Rodriguez, C. Sanabria, E. Velasquez, and S. J. Fonte. 2014. Soil ecosystem services and land use in the rapidly changing Orinoco River Basin of Colombia. *Agriculture Ecosystems & Environment* 185:106-117.

Levin, S.A. 1992. The problem of pattern and scale in ecology. The Robert H. MacArthur Award Lecture. *Ecology* 73, no 6: 1943-1967.

Locatelli, B., P. Imbach, and S. Wunder. 2014. Synergies and trade-offs between ecosystem services in Costa Rica. *Environmental Conservation* 41:27-36.

Lowe, P., and J. Phillipson. 2006. Reflexive interdisciplinary research: The making of a research programme on the Rural Economy and Land Use. *Journal of Agricultural Economics* 57:165-184.

Lowe, P., G. Whitman, and J. Phillipson. 2009. Ecology and the social sciences. *Journal of Applied Ecology* 46:297-305.

Maes, J., M. L. Paracchini, G. Zulian, M. B. Dunbar, and R. Alkemade. 2012. Synergies and trade-offs between ecosystem service supply, biodiversity, and habitat conservation status in Europe. *Biological Conservation* 155:1-12.

Malawska, A., C. J. Topping, and H. O. Nielsen. 2014. Why do we need to integrate farmer decision making and wildlife models for policy evaluation? *Land Use Policy* 38:732-740.

Maskell, Lindsay C., Crowe, Andrew, Dunbar, Michael J., Emmett, Bridget, Henrys, Peter, Keith, Aidan M., Norton, Lisa R., Scholefield, Paul, Clark, Douglas B., Simpson, Ian C.,

Smart, Simon M. 2013. Exploring the ecological constraints to multiple ecosystem service delivery and biodiversity. *Journal of Applied Ecology* 50 (3): 561-571. 10.1111/1365-2664.12085

McVittie, A., Norton, L., Matin-Ortega, J., Siameti, I., Glenk, K., Aalders, I. (2015) Operationalising an ecosystem services-based approach using Bayesian Belief Networks: An application to riparian buffer strips. *Ecological Economics* 100, 15-27.

MEA, 2005. Millennium Ecosystem Assessment, Ecosystems and Human Well-being: Synthesis, www.MAweb.org.

Morton, D., Rowland, C., Wood, C. Meek, L., Marston, C., Smith, G., Wadsworth, R., Simpson, I.C. (2011). Final Report for LCM2007 - the new UK land cover map. Countryside Survey Technical Report No 11/07 NERC/Centre for Ecology & Hydrology 112pp. (CEH Project Number: C03259).

Norton, L., Maberly, S., Waterton, C., Watson, N., Tsouvalis, J. (2014). The relevance of scale to water governance: An example from Loweswater, UK. In *Scale Sensitive Governance of the Environment* (Eds Frans Padt, Paul Opdam, Nico Polman and Catrien Termeer). John Wiley and Sons, Ltd.

Norton, L.R., Maskell, L.C., Smart, S.S., Dunbar, M.J., Emmett, B.A., Carey, P.D., Williams, P., Crowe, A., Chandler, K., Scott, W.A., Wood, C.M.(2012) Measuring stock and change in the GB countryside for policy – key findings and developments from the Countryside Survey 2007 field survey. *Journal of Environmental Management* 113:117-127.

Norton, L.R., Inwood, H., Crowe, A, Baker, A. (2011a). Trialling a method to quantify the 'cultural services' of the English landscape using Countryside Survey data. *Land Use Policy* 29: 449-455.

Norton, L. Elliott, J.A., Maberly, S, C., May, L. (2011b) Using models to bridge the gap between landuse and algal blooms: an example from the Loweswater catchment, UK. *Environmental Software and Modelling* 36: 64-75.

Padmanabhan M. and Jungcurt S., 2012. Biocomplexity - conceptual challenges for institutional analysis in biodiversity governance. *Ecological Economics* 81: 70-79.

Potschin M. and Haines-Young R., 2006. Rio+10, sustainability science and landscape ecology. *Landscape and Urban Planning* 75(3-4):167-174.

Schoolman, E. D., J. S. Guest, K. F. Bush, and A. R. Bell. 2012. How interdisciplinary is sustainability research? Analyzing the structure of an emerging scientific field. *Sustainability Science* 7:67-80.

Smart, S, Maskell, L.C., Dunbar, M.J., Emmett, B.A., Marks, S., Norton, L.R., Rose, P., Scholefield, P.A., Henrys, P., Simpson, I.C. (2010) Integrated Assessment report: Countryside Survey Technical Report.

<http://www.countryside-survey.org.uk/outputs/integrated-assessment-report-2007>

Spash, C. L. 2012. New foundations for ecological economics. *Ecological Economics* 77:36-47.

Swinton, S.M., Lupi, F., Robertson, G.P. and Hamilton, S.K., 2007. Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. *Ecological Economics* 64: 245-252.

UK National Ecosystem Assessment (2014). The UK National Ecosystem Assessment: Synthesis of the Key Findings. UNEP-WCMC, LWEC, UK.

Waterton, C., Maberly, S.C., Tsouvalis, J., Watson, N., Winfield, Ian. J., Norton, Lisa. R. (2015) Committing to Place: The potential of open collaborations for Trusted Environmental Governance. Public Engagement in Science series, *PLOS Biology* 13(3); e1002081. doi:10.1371/journal.pbio.1002081

Williams, A., and K. Hedlund. 2014. Indicators and trade-offs of ecosystem services in agricultural soils along a landscape heterogeneity gradient. *Applied Soil Ecology* 77:1-8.

Young, J. C., K. A. Waylen, S. Sarkki, S. Albon, I. Bainbridge, E. Balian, J. Davidson, D. Edwards, R. Fairley, C. Margerison, D. McCracken, R. Owen, C. P. Quine, C. Stewart-Roper, D. Thompson, R. Tinch, S. Van den Hove, and A. Watt. 2014. Improving the science-policy dialogue to meet the challenges of biodiversity conservation: having conversations rather than talking at one-another. *Biodiversity and Conservation* 23:387-404.

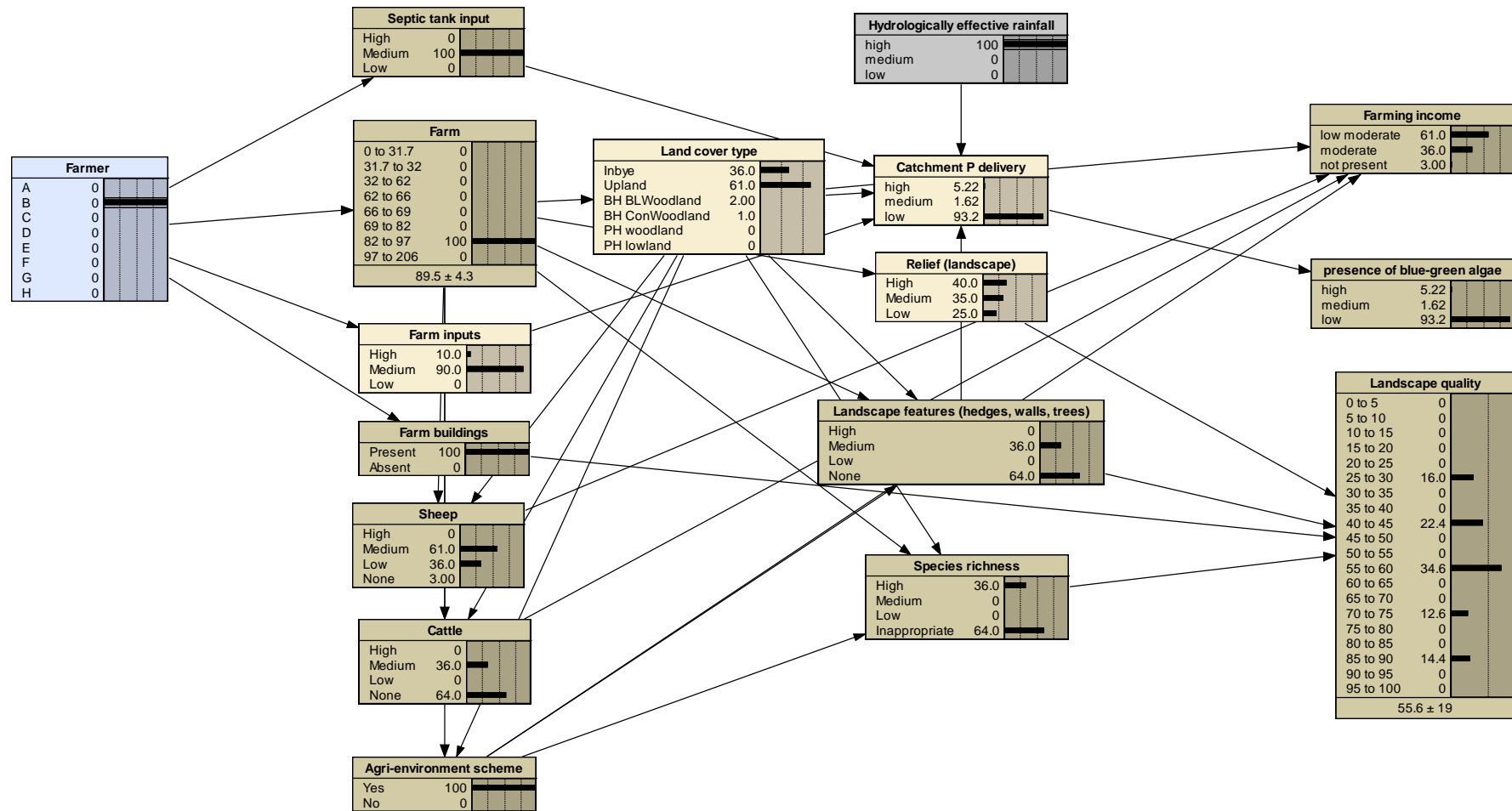


Figure 1: The causal Bayesian network for the provision of ecosystem service indicators: farming income, presence of blue-green algae and landscape quality

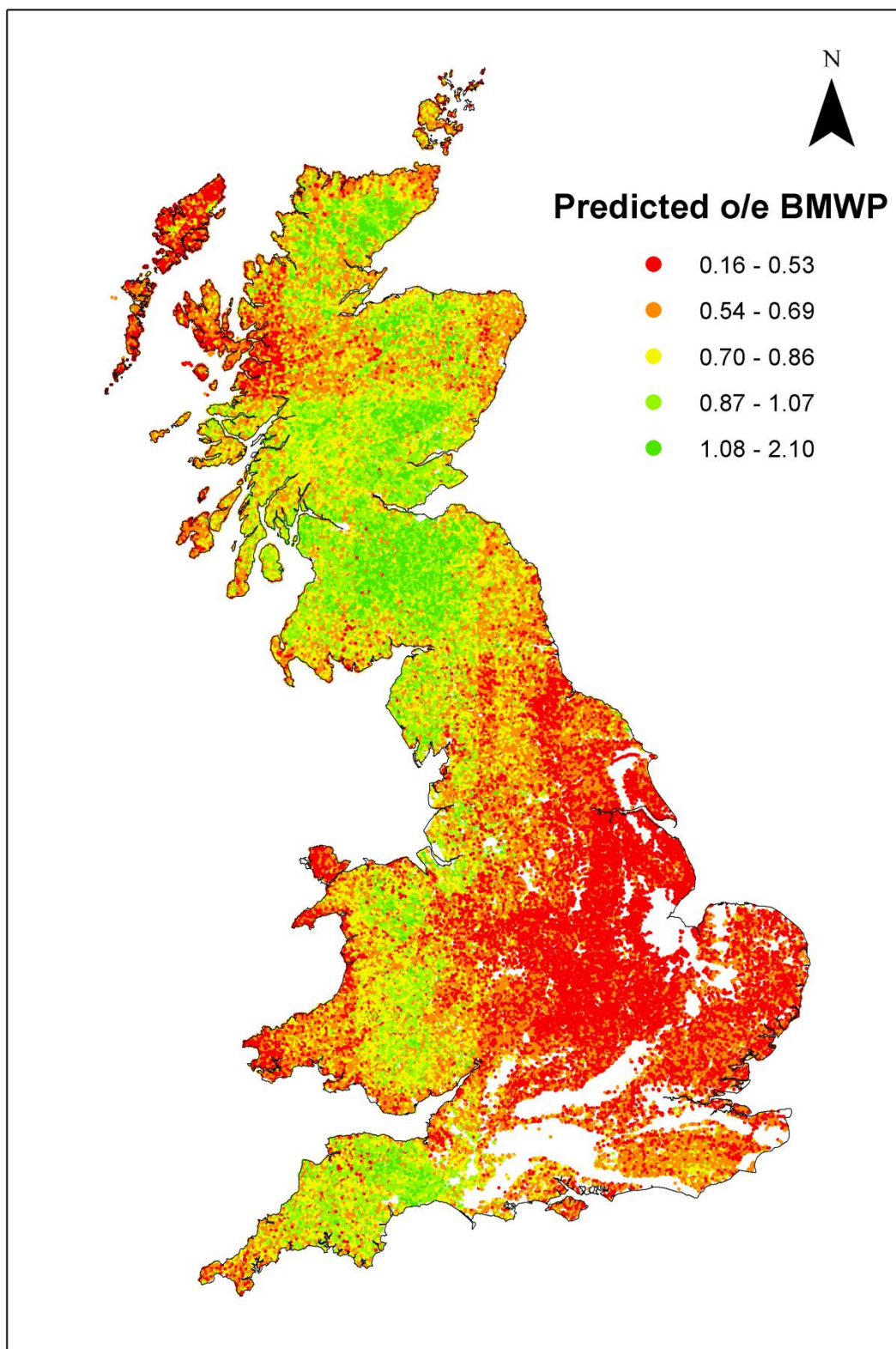


Figure 2: Map showing predicted o/e BMWP values for a headwater stream site in 152,639 1km² grids across GB using the grid model generated in the study.

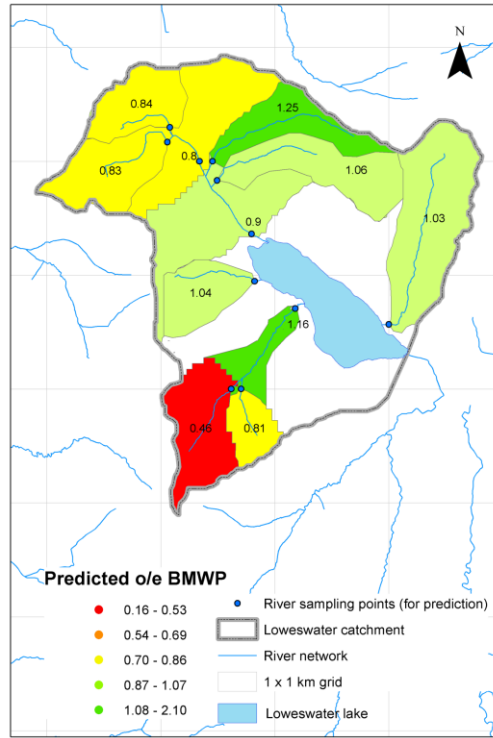
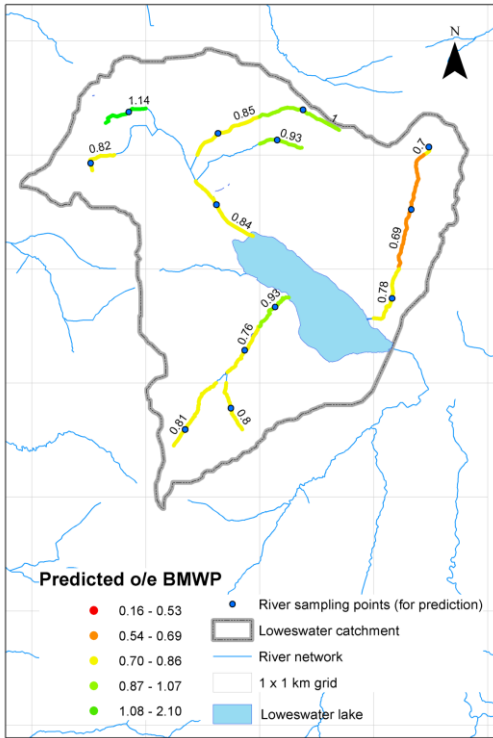


Figure 3: (left) Map showing predicted o/e BMWP values for a randomly selected river/stream site in each 1km² grid within the Loweswater catchment.

Figure 4: (right) Map showing predicted o/e BMWP values for sub catchments within the Loweswater catchment.