Article (refereed) - postprint

Barton, D.N.; Kelemen, E.; Dick, J.; Martin-Lopez, B.; Gómez-Baggethun, E.; Jacobs, S.; Hendriks, C.M.A.; Termansen, M.; García-Llorente, M.; Primmer, E.; Dunford, R.; Harrison, P.A.; Turkelboom, F.; Saarikoski, H.; van Dijk, J.; Rusch, G.M.; Palomo, I.; Yli-Pelkonen, V.J.; Carvalho, L.; Baró, F.; Langemeyer, J.; van der Wal, J. Tjalling; Mederly, P.; Priess, J.A.; Luque, S.; Berry, P.; Santos, R.; Odee, D.; Pastur, G. Martines; García Blanco, G.; Saarela, S.-R.; Silaghi, D.; Pataki, G.; Masi, F.; Vădineanu, A.; Mukhopadhyay, R.; Lapola, D.M.. 2018. (**Dis) integrated valuation – assessing the information gaps in ecosystem service appraisals for governance support [in special issue: SI: Synthesizing OpenNESS]** *Ecosystem Services*, 29 (C). 529-541. https://doi.org/10.1016/j.ecoser.2017.10.021

(Dis) integrated valuation – assessing the information gaps in ecosystem service appraisals for governance support

Barton D.N.[1] *, *E. Kelemen*[2], *J. Dick*[3], *B. Martin-Lopez*[4], *E. Gómez-Baggethun*[5], *S. Jacobs*[6], *C.M.A. Hendriks*[7], *M. Termansen*[8], *M. García- Llorente*[9], *E. Primmer*[10], *R. Dunford*[11] [12], *P. A. Harrison*[13], *F.Turkelboom*[6], *H.Saarikoski*[14], *J. van Dijk*[15], *G. M. Rusch*[16], *I. Palomo*[16], *V. J. Yli-Pelkonen*[17], *L. Carvalho*[18], *F. Baró*[19], *J. Langemeyer*[20], *J. Tjalling van der Wal*[20], *P. Mederly*[21], *J.A. Priess*[22], *S. Luque*[23], *P. Berry*[24], *R. Santos*[25], *D. Odee*[3] [26], *G. Martines Pastur*[27], *G. García Blanco*[28], *S-R. Saarela*[10], *D. Silaghi* [29], G. Pataki[30], *F. Masi*[31], *A. Vădineanu*[32], *R. Mukhopadhyay*[33], *D. M. Lapola*[34]

[1] *Corresponding author. david.barton@nina.no. Tel. 004792442111 Norwegian Institute for Nature Research (NINA), Gaustadalléen 21, 0349 Oslo,Norway.
[2] Environmental Social Science Research Group (ESSRG Ltd.), Rómer Flóris u.38.,1024 Budapest, Hungary.
Department of Decision Sciences, Corvinus University of Budapest, Fővám tér 8., 1093 Budapest, Hungary.
[3] Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian, EH260QB, UK
[4] Leuphana University of Lüneburg, Faculty of Sustainability, Institute of Ethics and Transdisciplinary Sustainability Research, Scharnhorststraβe 1, 21355 Lüneburg, Germany.
[5] Department of International Environment and Development Studies (Noragric), Norwegian University of Life Sciences (NMBU), P.O. Box 5003, N-1432 Ås, Norway.
[6] Research Group nature and Society, Research Institute for Nature and Forest (INBO), Kliniekstraat 25, 1070 Brussels, Belgium.
[7] Wageningen University and Research, Environmental Research (Alterra), P.O.Box 47, 6700 AA, Wageningen, The Netherlands.
[8] ENVS, Aarhus University, Denmark
[9] Department of Applied Research and Agricultural Extension, Madrid Institute for Rural, Agricultural and Food Research and Development (IMIDRA), Ctra. Madrid-Barcelona (N-II), KM. 38.200, 28802 Alcalá de Henares, Madrid, Spain
[10] Finnish Environment Institute, P.O. Box 140, FI-00251 Helsinki, Finland
[11] Environmental Change Institute, Dyson Perrins Building, South Parks Road, Oxford, OX1 3QY, UK
[12] Centre for Ecology & Hydrology Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK
[13] Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK.
[14] Finnish Environment Institute, P.O. Box 140, FI-00251 Helsinki, Finland
[15] Norwegian Institute for Nature Research (NINA), Postboks 5685 Torgarden, 7485 Trondheim.
[16] Basque Centre for Climate Change (BC3), Sede Building 1, 1st floor, Scientific Campus of the University of the Basque Country, 48940 Leioa, Spain
[17] Department of Environmental Sciences, P.O. Box 65, FI-00014 University of Helsinki, Finland.
[18] Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK.
[19] Institute of Environmental Science and Technology (ICTA), Universitat Autònoma de Barcelona (UAB), Edifici Z (ICTA-ICP), Carrer de les Columnes s/n, Campus de la UAB, 08193 Cerdanyola del Vallès (Barcelona), Spain.
[20] Wageningen University and Research, Marine Research (IMARES), P.O. Box 57, 1780 AB, Den Helder, Wageningen, The Netherlands.
[21] Department of Ecology and Environmental Sciences, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 94901 Nitra, Slovakia.
[22] Helmholtz-Centre for Environmental Research - UFZ, Permoserstraße 15, 04318 Leipzig, Germany.
[23] National Research Institute of Science and Technology for Environment and Agriculture, IRSTEA, UMR TETIS, 500 rue JF BRETON Montpellier 34000 CEDEX 5, France
[24] Environmental Change Institute, Dyson Perrins Building, South Parks Road, Oxford, OX1 3QY, UK.
[25] CENSE - Centre for Environmental and Sustainability Research, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
[26] Kenya Forestry Research Institute (KEFRI), P.O.Box 20412-0200, Nairobi, Kenya.
[27] Centro Austral de Investigaciones Científicas (CADIC CONICET), Houssay 200, Ushuaia (9140) Tierra del Fuego, Argentina.
[28] Fundación Tecnalia Research & Innovation, Energy and Environment Division**,** Parque Tecnológico de Bizkaia, C/Geldo, Edificio 700 , E-48160 Derio - Bizkaia (Spain)
[29] National Institute for Research and Development in Forestry "Marin Dracea", Eroilor Blvd 128, Voluntari, Romania.
[30] Institute of Ecology and Botany, MTA Centre for Ecological Research, Alkotmány u. 2-4., 2163 Vácrátót, Hungary.
[31] IRIDRA Srl, Via La Marmora, 51 . 50121 - Florence – Italy
[32] University of Bucharest - Research Center in Systems Ecology and Sustainability, Splaiul Independentei 91-95, 050095, Bucharest, Romania
[33] IBRAD (Indian Institute of Bio Social Research and Development), VIP Road, Kestopur, Prafulla Kanan, Kolkata 700101, West Bengal, India.
[34] UNESP - Universidade Estadual Paulista, Ecology Department - LabTerra, Av.24-A, 1515 CEP: 13506-900, Rio Claro, São Paulo, Brazil.

Abstract

The operational challenges of integrated ecosystem service (ES) appraisals are determined by study purpose, system complexity and uncertainty, decision-makers requirements for reliability and accuracy of methods, and approaches to stakeholder-science interaction in different decision contexts. To explore these factors we defined an information gap hypothesis, based on a theory of cumulative uncertainty in ES appraisals. When decision context requirements for accuracy and reliability increase, and the expected uncertainty of the ES appraisal methods also increases, the likelihood of methods being used is expected to drop, creating a potential information gap in governance. In order to test this information gap hypothesis, we evaluate 26 case studies and 80 ecosystem services appraisals in a large integrated EU research project. We find some support for a decreasing likelihood of ES appraisal methods coinciding with increasing accuracy and reliability requirements of the decision-support context, and with increasing uncertainty. We do not find that information costs are the explanation for this information gap, but rather that the research project interacted mostly with stakeholders outside the most decision-relevant contexts. The paper discusses how alternative definitions of integrated valuation can lead to different interpretations of decision-support information, and different governance approaches to dealing with uncertainty.

Key words: integrated valuation, ecosystem service appraisal, ecosystem service governance, information costs, uncertainty, valuation, eccosystem services cascade

Highlights:

- a theory of cumulative uncertainty in integrated valuation
- an information gap hypothesis for governance support of ES appraisals
- comprehensive classification of purposes for ecosystem service appraisal
- researcher led assessment of appropriateness of ES appraisals methods for governance support
- stakeholder assessment of participation in ES appraisal

Supplementary Material

S1 Conceptual basis for the value of information for decisive purposes
S2 definition of methods
S3 Selected stakeholder survey questions
S4 Additional descriptive data

1. Introduction

Ecosystem service (ES) appraisal methods include biophysical mapping, modeling, socio-cultural and economic assessments. Recent reviews point to a persistent gap in the promise of ES appraisal methods to provide readily usable information for decision-support (Laurans et al. 2013; Ruckelhaus et al. 2015; Martinez-Harms et al. 2015). The challenge is broader than lacking operationalisation of monetary valuation in decision-making (TEEB, 2010). Valuation in its broad sense of 'assigning importance' is inherently part of decisions on natural resource and land use (Gómez-Baggethun *et al.*, 2014; Jacobs, 2016). The literature reviews testify to a general pattern of 'disintegrated valuation', both in terms of integration of appraisals across the ES cascade (Haines-Young and Potschin, 2010), and operational application of ES appraisals in governance (Primmer *et al.*, 2015; Verburg *et al.*, 2016).

There are recent examples of integrated valuation across multiple types of value which are exceptions to the broad patterns revealed by literature reviews cited above, e.g. (Fontaine *et al.*, 2014; Martin-Lopez *et al.*, 2014; IPBES, 2015; van Oudenhoven *et al.*, 2015; Yee *et al.*, 2015; Barton *et al.*, 2016; Grêt-Regamey *et al.*, 2016). However, gaps in the operationalisation of ES appraisal for governance support remain. Martinez-Harms *et al.* (2015) review 144 studies, and find that ecosystem service appraisals do not capture the core steps of the decision making process, with much of the literature focused on quantifying and mapping the supply of ecosystem services. A review by Laurans *et al.* (2013) shows that a majority of studies have been conducted for informative uses, and only 2% of the 313 studies reviewed have targeted decision-making. Among the reasons explaining the lack of valuation applications in policy, Laurens et al. (2013) suggest data inaccuracy, information costs, lack of expertise among decision-makers, and lack of compatibility between valuation methods and regulatory frameworks.

Based on empirical case study experiences in the Natural Capital Project, Ruckelshaus *et al.* (2015) define four impact pathways for ES appraisal: (i) conducting research that is disseminated, (ii) making stakeholder differences transparent, and mediating through changing their perspectives, (iii) generating action in new policy and finance mechanisms, and (iv) producing outcomes for biodiversity, ecosystem services and well-being. Their impact pathways highlight that ES appraisals can have multiple and iterative purposes ranging from explorative research for systems understanding, to supporting design of policy instruments and documenting their outcomes. While most studies in the Natural Capital project addressed the first three purposes to some degree, only one of 22 case studies had documented outcomes during the six years of the project. Ruckelshaus et al. (2015) argue that this is to be expected, given the time taken to conduct studies and the significant time lags between multi-sector planning processes involved in ecosystem management and measurable impact on the ground.

A broad research community is working to conceptualise the multiple values of nature and nature's contributions to people, beyond monetary valuation (IPBES, 2015; Pascual, 2017; Jacobs, 2018). Similarly to the efforts to mainstream the economics of nature (TEEB, 2010), efforts are also under way to operationalize plural values in decision-making. For example in a recent special issue[35], Jacobs et al. (2016) call for a 'new school of integrated valuation', which would address multiple values, through self-critical reflexive research, learning from real world applications and aiming at societal, rather than only academic impact. They call for research that understands the socio-political setting of decision-making mechanisms and provides instrumental criteria of credibility and legitimacy that can help determine the appropriate level of integration.

The aim of this paper is to evaluate the practical challenges of integrated ES appraisals, in particular how study purpose, uncertainty, stakeholder requirements for reliability and accuracy, and stakeholder-

[35] See special issue on Integrated valuation of ecosystem services: challenges and solutions Ecosystem Services, Volume 22, Part B, Pages 213-402 (December 2016) http://www.sciencedirect.com/science/journal/22120416/22/part/PB

science interaction co-determine the uptake and influence of integrated ES appraisal methods in different decision contexts. We analyse a diverse set of real-world case studies of the EU FP7 OpenNESS project (Dick, 2018), which have operationalised the ES concept and applied a range of biophysical, socio-cultural and monetary ES appraisal methods, as well as 'hybrid' and 'integrating' methods(Gómez-Baggethun *et al.*, 2014). Through this analysis, we report on the extent to which the project as a whole succeeded in narrowing the gap between ES appraisal and governance support – or how far we came in putting the parentheses in '(dis)integrated valuation'.

2. A theory of cumulative uncertainty in integrated ES appraisal

The accuracy and reliability required of ES appraisal methods must be seen in the context of the socio-ecological system in which it is applied. Uncertainty in ES appraisals results from the number of units and values to be considered, and the extent and complexity of the social system and ecosystems that are being appraised. Integrated ES appraisal for decision-support faces a challenge of addressing uncertainty from biophysical, socio-cultural and value heterogeneity (Gómez-Baggethun *et al.*, 2014). Integrated environmental modeling shows us that methodological and measurement error across conditionally dependent models is cumulative (Barton et al. 2016). Conditionally dependent appraisal methods can expected to accumulate uncertainty across sequentially linked models, especially if model parameters are not estimated jointly across models. Clearly, this will further challenge application of integrated valuation in governance. Those governance mechanisms that address the entire ES cascade, hierarchical governance in particular, would require comprehensive appraisal (Primmer et al. 2015). Paradoxically, comprehensive appraisals that document uncertainty might not seem appealing for operational use in governance support if they uncover large uncertainty about system response to policy (Barton *et al.*, 2008). Here we argue that cumulative uncertainty is one reason we would expect to find few valuation studies integrated in decision-support appraisals, as documented by Laurans *et al.* (2013) and Martinez-Harms *et al.* (2015).

Figure 1 outlines a theory of cumulative uncertainty in an integrated chain of ES appraisal methods. It builds on the ES cascade framework (Haines-Young and Potschin, 2010) which suggests a sequential cascade of linked ecosystem appraisals, leading to final valuation outcomes. In the original ES cascade framework valuation outcomes feed back to policy actions that control pressures. The ES cascade framework has been criticized and further developed in ways that provide support for a cumulative theory of uncertainty in integrated ES appraisal (Spangenberg *et al.*, 2014; Nassl and Loffler, 2015; Primmer *et al.*, 2015; Potschin and Haines-Young, 2016; Hausknost *et al.*, 2017). Spangenberg et al. (2014) propose that different kinds of agency are required to transform biophysical structures and processes into ecosystem service values, including use value attribution, mobilisation of ecosystem service potentials into ecosystem services, appropriation of ecosystem services as benefits of use, and commercialisation of use values into exchange values. The transformation and maintenance of nature for societally desired states as depicted by Spangenberg et al. (2015) requires a constant input of energy (Hausknost et al. 2017). Hausknost et al. (2017) suggest that the ES cascade be depicted as a 'stairway' metaphor where each step in appraisal or appropriation requires the investment of human agency, involving work, time and other resources. ES appraisal involves the investment of human agency in the production of information to reduce uncertainty (entropy) about states of nature resulting from actions to transform and maintain it for societal purposes.

1

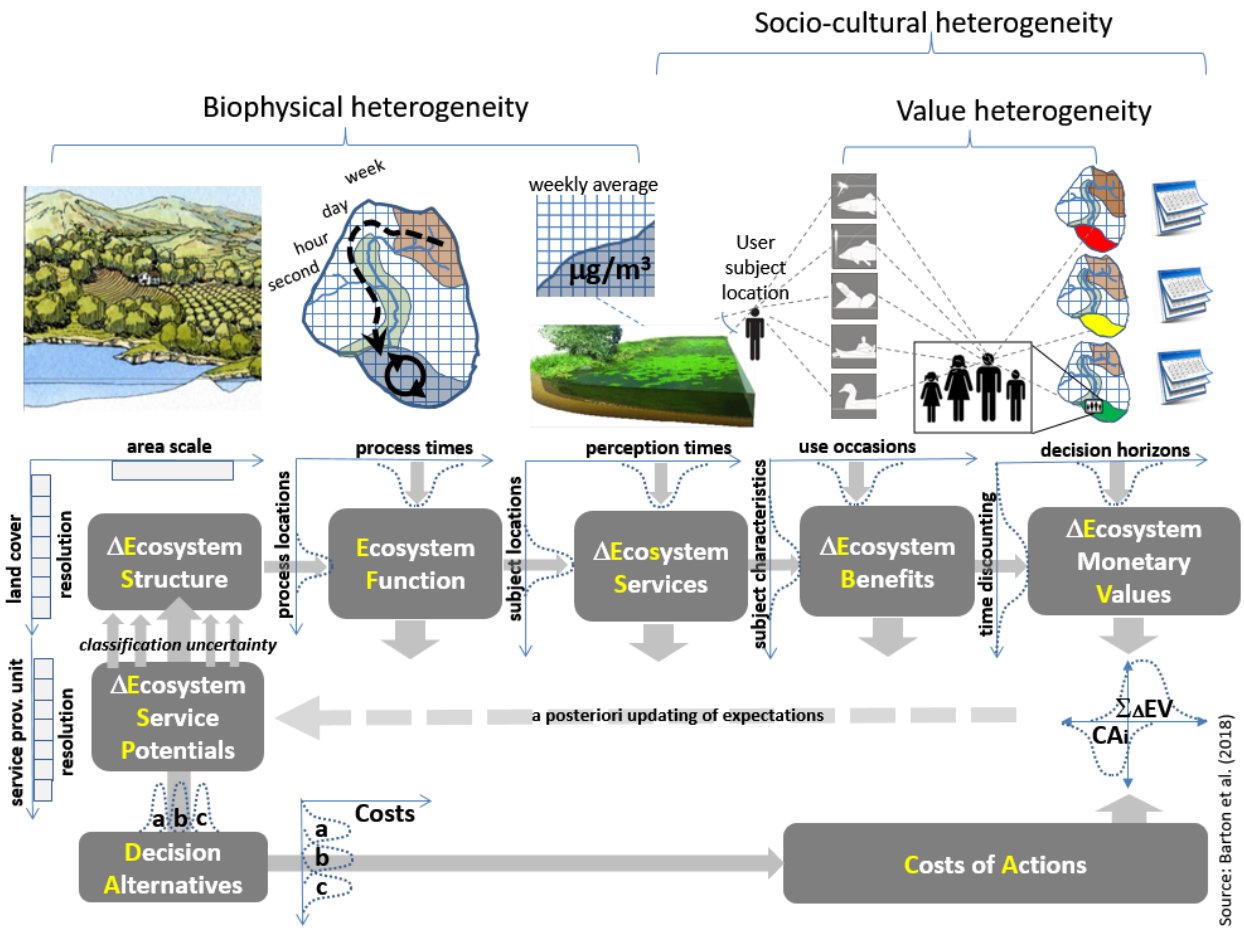

Figure 1 A theory of expected cumulative uncertainty

Integrated ES appraisal links actions based on expected changes in ecosystem service potentials to changes in ecosystem structure, changes in ecosystem condition, changes in ecosystem services, to changes in benefits and values of those changes. Decision support is provided by comparing uncertain costs of alternative actions (CA) with uncertain ecosystem benefits, such as exchange values (EV), in a e.g. Benefit-Cost analysis. Uncertainty is cumulative in this framework, challenging accuracy and reliability as appraisal integrates down a causal chain. Accuracy of ecosystem service assessments decreases as the appraisal integrates across biophysical, socio-cultural and value heterogeneity. Reliability in consistently ranking of alternatives increases along the chain as more type of impact information is generated at each step. With the focus on appraisal information in a theory of expected cumulative uncertainty, we have modified the ES casade version by Spangenberg et al. (2015) and the ES stairway by Hausknost et al. (2017). The epistemological step of attributing of ecosystem service potential to the landscape codetermines which ecosystem structures are targeted for landuse decisions. The attribution of ES potential is subject to classification uncertainty regarding which structures have what kinds of ES potential. ES appraisals update these expectations for every appraisal iteration. Figure credits: Landscape artwork http://www.dpz.com/Initiatives/Transect. Lake artwork and use icons: Hime, S. and I.J. Bateman, A transferable water quality ladder for conveying use and ecological information within public surveys, CSERGE, University of East Anglia. 2008.

Indeed, Nassl and Loffler (2015) discuss the ecosystem service cascade as an integrated environmental assessment framework, drawing parallels to the driver-pressure-state-impact-response (DPSIR) framework. ES appraisal to support decisions requires some form of integration across such a DPSIR chain, including (i) identification of changes in values, subject to (ii) change in ecosystem services, subject to (iii) change in ecosystem condition, subject to (iv) identification of ecosystem service potential associated with (v) the ecosystem structures in question.

A theory of cumulative uncertainty proposes that integrated ES appraisal as defined in Figure 1 is subject to cumulative uncertainty in its outcomes. Cumulative uncertainty can be actual, as when a series of models are linked to simulate impacts of management decisions (Barton *et al.*, 2016). Cumulative uncertainty can also be expected, as when decision-makers are already familiar with the complexity of the system they are managing. Rational decision-makers will assess how accurately and reliably study outcomes are linked to the decision alternatives at hand. While an ES appraisal may not have conducted sensitivity analysis or probability simulation, rational decision-makers will formulate expectations about how uncertain outcomes are, relative to the purpose of interest to them. Their expectations will be based jointly on the complexity of the context, and the complexity of assumptions of the ES appraisal method. Primmer et al. (2015) discuss how different governance mechanisms have arisen to address uncertainty and different information inputs along the ES cascade. Hierarchical, scientific-technical, adaptive-collaborative and strategic governance rest on different information assumptions, and use information in differing ways to reduce uncertainty for different governance purposes (Primmer et al. 2015).

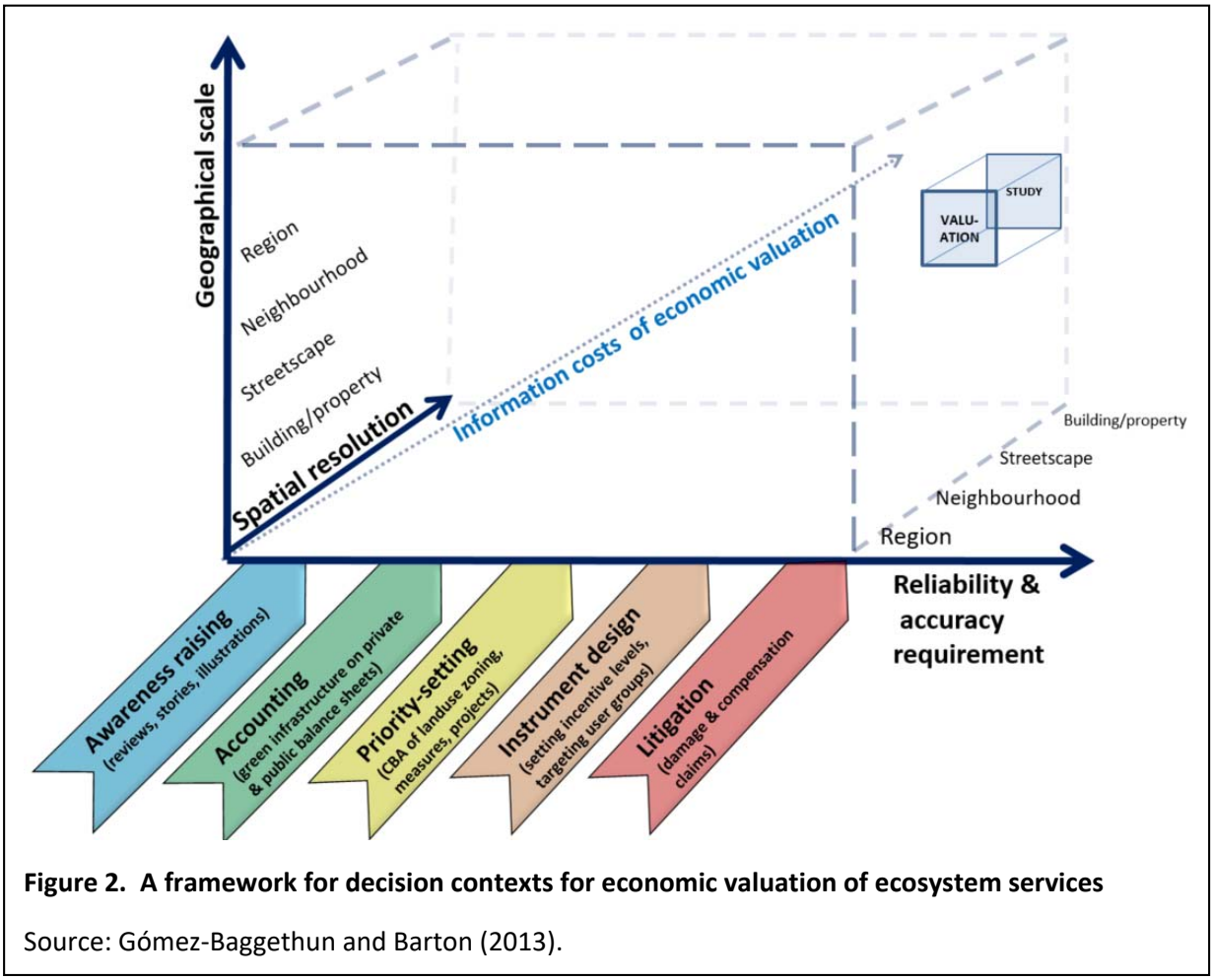

Figure 2. A framework for decision contexts for economic valuation of ecosystem services

Source: Gómez-Baggethun and Barton (2013).

Gómez-Baggethun and Barton (2013) identified a range of governance purposes of urban ES valuation (Figure 2), including (i) awareness-raising, (ii) accounting, (iii) priority-setting, (iv) instrument design, and (v) litigation. Schröter *et al.* (2014) proposed a similar ordering of study purposes in the context of ecosystem accounting at regional and national scales. They argued that different purposes can be organised along a gradient of increasing decision-maker expectations for accuracy and reliability. The more accuracy and reliability required of the purpose of the analysis, the more tailored and information intensive ES appraisal methods need to be.

Accuracy and reliability of different ES appraisal methods are key concerns for practitioners (Dick, 2018; Dunford, 2018; Harrison, 2018). Literature reviews highlight information costs as one of several challenges to operationalisation of the ecosystem services concept (Bagstad *et al.*, 2013; Laurans *et al.*, 2013; Ruckelshaus *et al.*, 2015). Bagstad *et al.* (2013) argue that the time and resources needed to obtain data, analyse and communicate results constitute one of the most important challenges. The fact that ecosystem service appraisals are costly is trivial. The point we wish to make is that information costs increase incrementally with both system complexity and decision-support purpose. Further, decision-makers' requirements for accuracy and reliability will be correlated with increasing information costs if, and only if, ES appraisers make efforts to meet those requirements (Gómez-Baggethun and Barton, 2013). In other words, both decision-makers and ES appraisers need to have a common terminology for ES appraisal uncertainty and decision-purpose to avoid a gap in expectations. To this end we formulate an information gap hypothesis and show how it can be used to compare across ES appraisal methods and their purposes in practice.

3. The information gap hypothesis

The information gap hypothesis is based on two propositions, the first regarding the expected uncertainty of appraisal methods and the second regarding decision-makers' requirements (Figure 3).

Expected uncertainty in outcomes. The theory of cumulative uncertainty in ES appraisal (Figure 1) proposes a classification of ES appraisal methods along the ES stairway in order of increasing expected uncertainty in outcomes. Expected uncertainty increases with the position of the method in the appraisal chain. The further along the appraisal chain the lower its likelihood of application assuming decision-makers avoid appraisal uncertainty. The information gap hypothesis therefore predicts an ordering of methods' likelihood of application as follows. (1) ES mapping methods aimed at attributing ecosystem service potentials. (2) ES modelling methods aimed at describing the response function between changes in structure and services. (3) Socio-cultural assessments aimed at identifying qualitative ES benefits from use values. (4) Monetary valuation aimed at identifying economic values of ES benefits. (5) Synthesizing methods combining two or more types of ES appraisal output (see Table 2, next section for examples). While synthesizing methods accumulate uncertainty from multiple methods, their purpose is to compare and triangulate different indicators of importance, thereby reducing overall uncertainty about decision alternatives (Jacobs *et al.*, 2016; Jacobs, 2018). This would lead us to expect some intermediate likelihood of application, as the effects of entropy in complexity and information through triangulation are balanced against oneanother.

Decision-makers' requirements. The typology of decision purposes and their requirements is based on previous reviews (Gómez-Baggethun and Barton, 2013; Laurans *et al.*, 2013; Martinez-Harms *et al.*, 2015; Ruckelshaus *et al.*, 2015). Four broad typologies of ES appraisal purpose are defined, including *explorative, informative, decisive* and *design* purposes (see Table 1 for definitions). The requirements for accuracy and reliability increase systematically across these four broad study purposes, following the rationale in Figure 2. *Explorative purposes* rely on research agenda driven analyses, with low

expectations for direct application to decision-making, generating few costs for governance support. *Informative purposes* include awareness raising about importances of ecosystem services, with no specification of decision context. *Decisive purposes* require ES appraisal to distinguish between the desirability of alternative decisions. *Design* purposes have the highest requirements for accuracy and reliability, as values are used to calibrate the scope and targeting of policy instruments and management actions.

Combining these two propositions, the information gap hypothesis states that *the more demanding the decision-support purpose of the ES appraisal, the lower the likelihood that the ES appraisal will be considered appropriate for decision-support, conditional on an expected uncertainty defined by system and method complexity.* Decreasing likelihood of the method being appropriate indicates an information gap in ES appraisal for governance support (Figure 3). The information gap may result whether researchers identify the uncertainty of their methods or not, it relies crucially on decision-makers expectations.

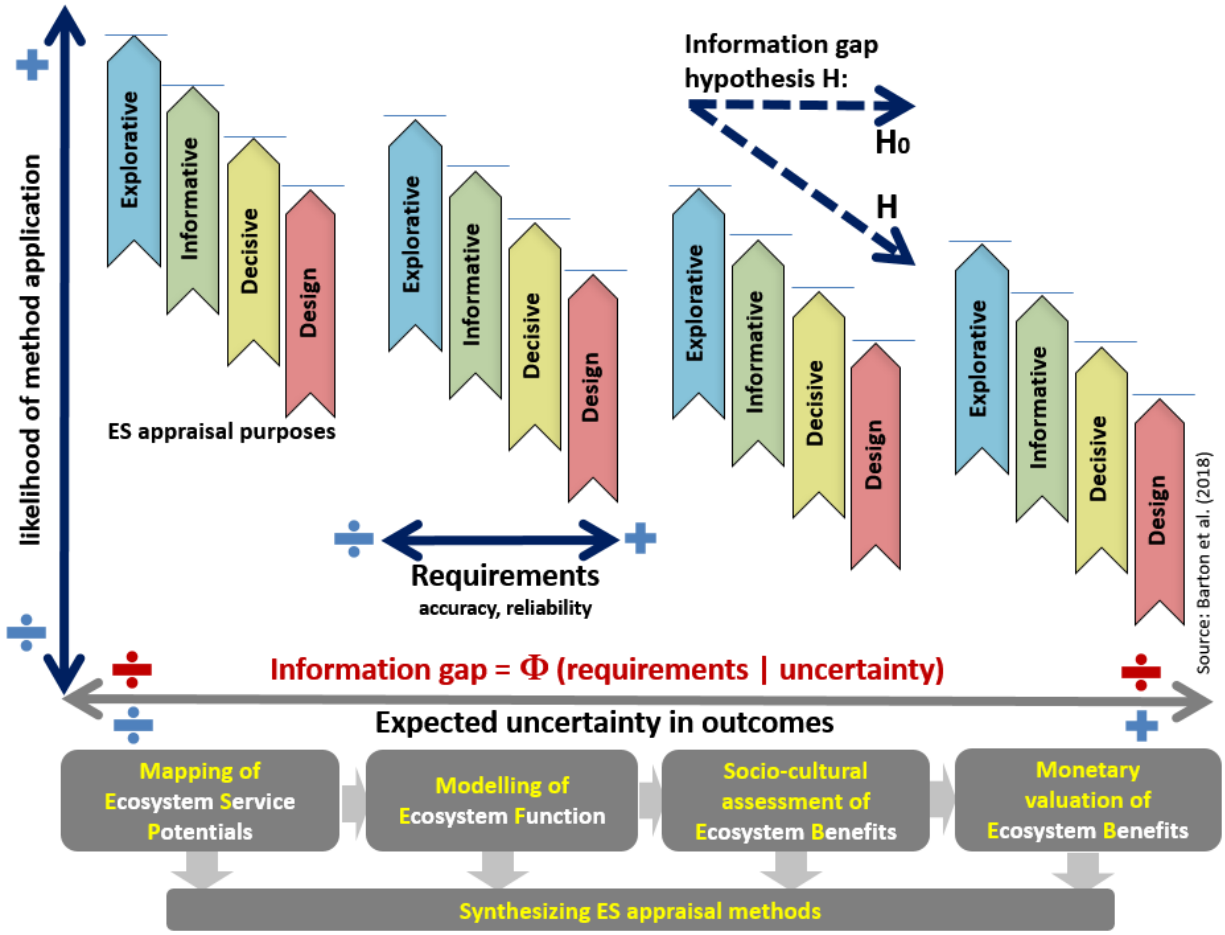

Figure 3. The information gap hypothesis in ecosystem service appraisals for governance support

The information gap hypothesis is based on a theory of cumulative uncertainty. Successive steps in a 'stairway' of ES appraisals - from mapping, modeling to benefit assessment and valuation - lead stakeholders to expect higher cumulative uncertainty in outcomes. Given an expectation about uncertainty, the more demanding the decision-support purpose of the ES appraisal, the lower the likelihood that the ES appraisal method will be considered appropriate for decision-support. The causal mechanism behind the hypothesis is a function of decision-makers' requirements for accuracy and reliability in a particular study purpose, conditional on the expected uncertainty of the ES appraisal method. This is denoted in the figure as: Information gap = Φ (requirement | uncertainty). For any given type of method we would expect to see decreasing likelihood(L) of method application, ordered by purposes as follows: L(explorative)>>L(informative)>>L(decisive)>>L(design). The relative likelihood of successive method types being appropriate – mapping-modelling-socio-cultural-valuation - would also be expected to decrease with the more cumulative uncertainty in outcomes that is expected by decision-makers. Figure 3 shows that we have no expectation of the ordering of synthesizing methods. The coloured chevrons in the figure depict the combination of these two ordering effects.

The information gap hypothesis is based on a theory of cumulative uncertainty and a classification and ordering of method and their purposes, either of which may be disproved. The information gap hypothesis could be rejected (H_0) by limitations in the classifications of methods, purposes and the requirements assumed to order them. For example, explorative research purposes may have high

accuracy requirements if testing and developing theories that make precise predictions. Large data sets with high information costs may be needed to obtain sufficient statistical power - the likelihood of observing methods applied for explorative purposes might in such cases be low, rather than high as predicted by the information gap hypothesis. At the other end of the scale it could be argued that technical design may be understood as repetitive and requiring little accuracy and reliability. In that case, low information costs would lead us to expect a high likelihood of methods being applied for design purposes. Furthermore, individual methods may be misclassified in terms of the method type they are assumed to belong to (Table 2, Supplementary Material S2). If such misclassification is significant,. the information gap hypothesis (H) in Figure 3 would also be rejected. We would see no ordering as in H_0, or even increasing likelihood of methods with increasing uncertainty and requirements.

As explained above, we expect constraints on data, budget and time available for appraisals supporting governance to be important reasons for the likelihood of method selection decreasing according to the information gap hypothesis. But there may also be alternative explanations for observing 'much mapping and little valuation'. These may include the lack of participation of decision-mandated stakeholders, lack of compatibility between research and operational needs, as well as how risk is managed by different types of governance. These alternative explanations are addressed in the discussion section.

4. Materials and methods

The OpenNESS project was designed to operationalise the ecosystem services concept in case studies across Europe, Brazil, Kenya and India, through close collaboration between researchers and stakeholders(Dick, 2018)In the last year of the project, 26 case study leaders and 246 stakeholders responded to targeted surveys evaluating the different methods applied during the project period from 2013 to 2016. The survey of case study leaders provided information on the researcher perceived purpose of studies and operational constraints. The stakeholder survey provided information on stakeholders perceived interaction with the appraisal methods.

4.1 Case study leader survey

A survey was circulated to case study research teams in 2016, in which the case study leaders were asked to identify all the appraisal methods used within their case study. For each appraisal method, the respondents were asked to score considerations that influenced their decision to use that method within their case study. For the complete survey protocol of case study coordinators, see Dick et al. (2018) Supplementary Material – "Annex 4 Case study context reporting forms part 2". For an analysis of a comprehensive evaluation of case study considerations see Dunford (2018) for an analysis of a comprehensive evaluation of case study considerations. In this paper we focus on the subset of considerations concerning study purposes and information constraints. Table 1 describes 19 different study purposes in ES assessment, based on a synthesis of decision contexts and impact pathways in the literature (Gómez-Baggethun and Barton, 2013; Laurans *et al.*, 2013; Martinez-Harms *et al.*, 2015; Ruckelshaus *et al.*, 2015) and adapted to the OpenNESS 26 case studies included in this study.

Table 1 Range of study purposes of each ES appraisal method scored by case study representatives

Explorative	**Conduct research** aimed at developing science and changing understanding of research peers
E1	Theory and concept development
E2	Hypothesis formulation and testing
E3	Method development and testing
Informative	**Change perspectives** of public & stakeholders
I1	Assessment of current state
I2	Assessment of long-term historic trends
I3	Assessment of potential future conditions
I4	Evaluation of existing projects and policies
I5	Raising awareness of the importance of ES
I6	Raising awareness of trade-offs and conflicts between ES
Decisive	**Generate action** in specific decision problems by stakeholders
D1	Decision problem formulation and structuring
D2	Criteria for screening alternatives
D3	Criteria for ranking alternatives
D4	Criteria for spatial targeting (zoning & planning of alternatives)
D5	Arguments for negotiation, shared norms & conflict resolution
Design	**Produce outcomes** through design and implementation of policy instruments with stakeholders
T1	Standards & policy target-setting
T2	Land and natural resource management rules & regulations
T3	Licencing / permitting / certification
T4	Pricing, setting incentive levels
T5	Establishing levels of damage compensation

Study purposes were classified into broad categories including explorative, informative, decisive and technical design in order to test the information gap hypothesis in Figure 2. A total of 80 ES appraisals in the OpenNESS project were ordered into 5 method types (Table 2) following the framework in Figure 1. Method descriptions are provided in Supplementary material S2. Most case studies applied more than one method. The survey evaluated case studies' reasons for selecting methods across a large number of criteria (Dunford, 2018). Integrated mapping-modelling methods and decision-support methods were grouped together in a single class of "synthesising" methods.

In a research project such as OpenNESS we do not observe the actual frequency of method applications to different purposes, given the limited timespan of the project(Ruckelshaus *et al.*, 2015). Instead we use a proxy indicator of the likelihood of methods application - for each method 26 case study leaders were asked 'To what extent is the way that you use the method in your case study described by the purposes listed (in Table 1)'? (scores: 0 = "not relevance"; 1 = "relevant" 2 = "primary purpose"). This question provided the indicator of how well a method addressed the explorative, informative, decisive or technical design purposes listed in Table 1. All methods were scored for all 18 purposes in Table 1. A single method could have multiple purposes. Regarding considerations for method selection, case study leaders were asked, 'To what extent are the following practical/research-related considerations factors that influenced your choice of this method?' (scores: 0 = "not at all"; 1 = "to some extent" 2 = "very

much"). From the list of considerations we used the answers regarding data, budget and time constraints as proxy indicators for information cost[36].

Table 2. Classification of ES appraisal methods used by case studies

Mapping (n=24)	**Biophysical Modelling** (n=10)	**Socio-cultural valuation** (n=25)	**Monetary valuation** (n=12)	**Synthesising methods** (n=9)
• ESTIMAP • Matrix approach • Quickscan • Smarthphone Apps • PPGIS	• SITE Landuse model • Bayesian belief network • Climate envelope modelling • Hydrological model • Meta-analysis	• Questionnaire & narrative analysis • Photoseries analysis • Preference assessment • Time use • Q-method • Deliberative valuation • Stakeholder analysis • Focus group	• Time use value • Value transfer • Cost-based • Revealed preference • Stated preference	• Integrated mapping-modelling (INVEST, EcoServ) • Scenario development • Multi-criteria decision analysis • Benefit-cost analysis

4.2 Case study stakeholder survey

After three years of OpenNESS case study work in close consultation with case study advisory boards (CAB), a standard questionnaire was administered to 246 case study stakeholders in 2016. For the complete survey protocol of stakeholders, see Dick et al. (this issue) with Supplementary Material – "Annex 1 Practitioner's perspective questionnaire". Three methods were used for selecting respondents: (i) restricting the respondents to CAB members (8 case studies), (ii) complementing the CAB respondent group with stakeholders outside the CAB (8 case studies), and (iii) including all stakeholders with relevant involvement in the process, as evaluated by the CAB and case study leader (11 case studies). Given the flexibility and variation across case studies in stakeholder participation during the three years of case study, the importance of CAB membership was assessed in relative terms. For the OpenNESS project as a whole stakeholders self-reported their 'membership of the CAB' as follows: *very applicable* (39%), *applicable* (14%), *somewhat applicable* (3%), *little bit applicable* (6%) and *not applicable* (38%). Partial involvement in the CAB reflects the dynamic nature of CAB membership, with individuals leaving and new members joining during the lifetime of the project in some case studies (Dick, 2018).

The stakeholder questionnaire was structured into four main topics (i) self-characterisation of stakeholders, (ii) perception of the participatory process followed in the case study, (iii) perceived impact, and (iv) practical usefulness of appraisal methods, allowing the stakeholders to feed back their experiences anonymously. We use a selection of the stakeholder survey data for our analysis - in particular, the stakeholders' degree of participation in the case study advisory board; the extent to which they participated in study design; method selection; knowledge co-production and how informed they

[36] We also applied a survey to case study leaders to quantify person months and expenses allocated per appraisal method. Application of multiple methods per case study meant that several case studies were unable to assign costs exactly. Due to missing responses we reverted to use of the qualitative question described above which was responded to by all but one case study.

were about results. Knowledge co-production was defined as 'attending workshops/meetings/ and stakeholder engagement activities' (question wording provided in Supplementary Material S3).

4.3 Testing the information gap hypothesis

We defined a t-test of differences in the mean scores for ordinal scoring of method relevance (Rasch *et al.*, 2007; Fagerland *et al.*, 2011), comparing pairwise samples of different study purposes. The T-test is robust to testing independent samples even under conditions of non-normality. Study purposes were ordered by expected information gap as explained in Figure 3. Using the scores assigned by case study leaders for each method tested in their case studies, we calculated the mean scores (s) of each method class (c) over each class of study purpose (p) (as defined in Table 2 and Table 1, respectively).

Given that we were testing whether there was a reduction (as opposed to increase) in the appropriateness of methods compared over pairwise purposes, we use the reported confidence levels of one side of the t-test,

$Pr(T < t)$, Ho: mean(diff) = 0 , Ha: mean(diff) < 0 ,

Mean(diff) was the difference in mean scores between pairwise consecutive study purpose categories organised in order of increasing information gap, as defined in Table 1 (Explorative – Informative(E-I), Informative – Decisive(I-D), Decisive – Technical design(D-T), Informative – Technical Design(I-T)). In other words, the expectation is that

mean(E-I)<0, mean(I-D)<0, mean(D-T)<0, and mean(I-T)<0,

If we inferred that for example mean(I-D)<0 is true, we would be confident that the mean score of *informative* method applications is higher than the mean score of *decisive* method applications. Table 3 reports different confidence levels of $Pr(T < t)$. Three of the method classes had relatively few observations (modelling, monetary valuation, and synthesising methods). For these method classes the power of the t-test is lower, and the test has less power to observe the hypothesised difference in means.

We next contrasted the results of the hypothesis tests with case study leaders' self-reported evaluation of importance of data, time and budget constraints in selecting the methods actually used. These results are reported in Figure 4. Recalling Laurans *et al.* (2013), we also evaluated complementary explanations to the information gap hypothesis using stakeholders' self-reported roles in the project, their degree of participation in decision-making, their perceptions of the general impact of the appraisal methods on decision-making, and prior researcher-stakeholder collaboration experience. We use a two-way fractional polynomial plot (Stata, 2013) to illustrate (i) whether stakeholders' participation in case study advisory board was correlated with co-design of ES appraisal methods, and (ii) whether the length of time researchers had worked with stakeholders in the case study prior to OpenNESS was correlated with the importance scores for different study purposes (Supplementarr Material Figure S2).

5 Results

5.1 Uses of methods for different purposes

Figure 4 shows the mean relevance scores for each of the 5 method classes (defined in Table 2) and for each of the four groups of study purposes (defined in Table 1). Our hypothesis that the applicability of appraisal methods is negatively correlated with an ordering of study purposes by increasing information requirements was moderately supported by the case study leader survey responses. Although we did not test it statistically, visual inspection of Figure 4 also shows an ordering of method types as expected by ta theory of cumulative uncertainty. The detailed distributions of relevance scores are given in Figure S1, Supplementary material. Comparing methods for the same purpose, we see that 'mapping' methods scored almost as highly as 'synthesising' methods for *decisive* purposes. Synthesing methods scored the highest for decisive context and in the median range for other purposes. Mapping scored more highly than other individual methods across all purposes, and monetary methods scored the lowest across all purposes (Figure S1).

A closer inspection of the distributions using the t-test shows that most of the adjacent purposes are different by order of increasing study costs (Table 3). There are some exceptions. For example, in the case of modellling methods we are only 83% confident that *decisive* purpose is more prevalent than *technical design*; for socio-cultural methods we are only 79% confident that *informative* purposes are more prevalent than *decisive* purposes. However, Table 3 shows that there is a significant ordering of methods' appropriateness across purposes with increasing requirements. The mean scores by purpose class in Figure 5 hide a lot of heterogeneity within each class. Figure S1 in Supplementary materials provides further detail across the 19 different detailed study purposes. From this more detailed picture we see that among *explorative* study purposes *methods development & design* was the most important study purpose for case studies. Among *informative* purposes creating awareness of the current state and importance of ecosystem services were the dominant purposes across the different case studies. *Decisive* purposes had no dominant detailed purpose. Decision-support tools (MCDA, BCA) and integrated mapping-modelling methods were on the whole more relevant for *decisive* purposes than were mapping, modelling and valuation methods on their own. For *technical design* purposes, input to design of natural resource *management rules and regulations* was the most relevant technical design purpose across OpenNESS case studies, scoring as high as decisive purposes for both mapping methods and synthesis methods.

1

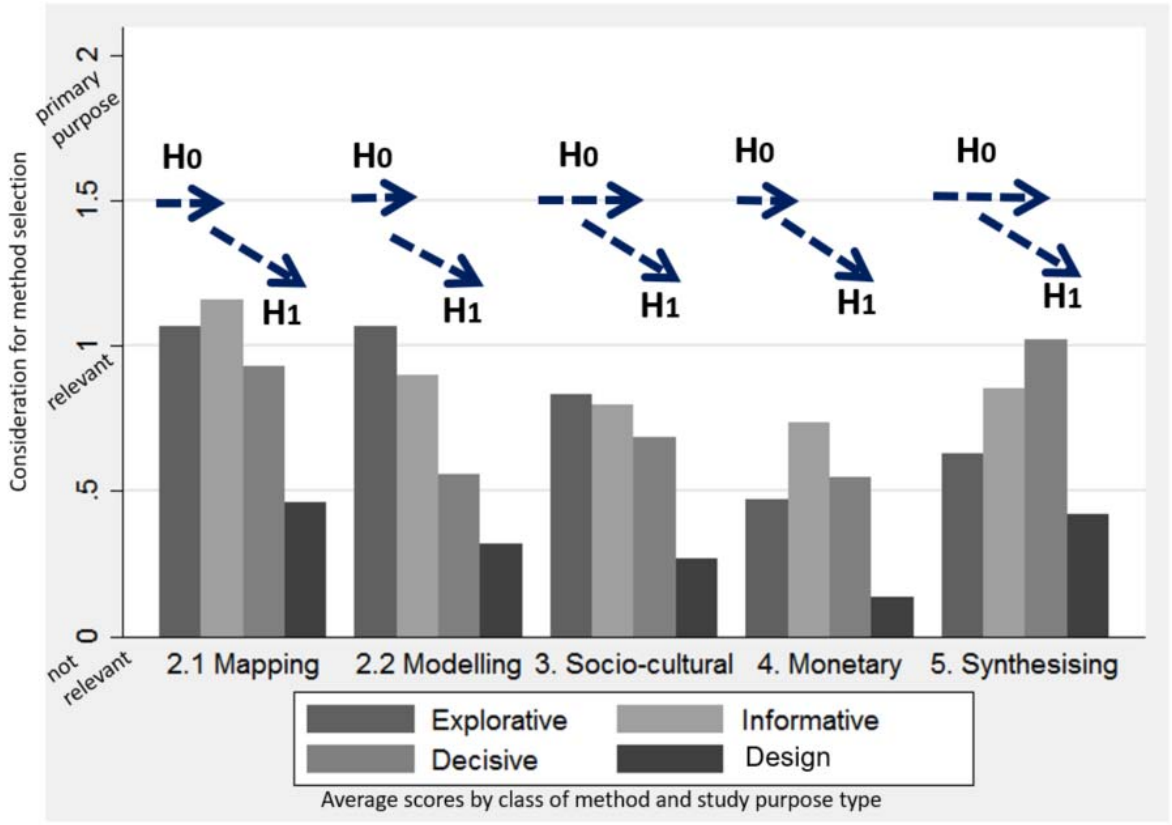

Figure 4 Mean relevance scores of the purposes of 80 methods from 26 case study leader respondents

Table 3 T-test of the difference in the mean of method relevance scores between consecutive study purposes

Pr(T < t) for Ho: mean(diff) = 0 , Ha: mean(diff) < 0				
Purposes: **Method groups:**	**Explorative-Informatve** mean(E-I)	**Informative-Decisive** mean(I-D)	**Decisive-Technical Design** mean(D-T)	**Informative-Technical design** mean (I-T)
Mapping	t = -0.9725 Pr= 0.1678	t = 2.3522 Pr= 0.9886	t = 4.0675 Pr= 0.9999	t = 7.6059 Pr= 1.0000
Modelling	t = 1.1971 Pr= 0.8766	t = 1.4362 Pr= 0.9159	t = 0.9909 Pr= 0.8326	t = 3.8562 Pr= 0.9994
Socio-cultural	t = 0.2905 Pr= 0.6136	t = 0.8008 Pr= 0.7863	t = 2.5861 Pr= 0.9935	t = 4.5755 Pr= 1.0000
Monetary	t = -1.6848 Pr= 0.0531	t = 1.3250 Pr= 0.9006	t = 3.3544 Pr= 0.9986	t = 6.1671 Pr= 1.0000
Synthesis	t = -0.9104 Pr= 0.1881	t = -0.6888 Pr= 0.2504	t = 2.6572 Pr= 0.9914	t = 2.0364 Pr= 0.9707

Confidence level: 99% 95% 90% <90%

Figure 5 illustrates the relative importance – as assessed by case study leaders - of data, time and budget constraints in the selection of methods used in the case studies, compared to alternative methods. Data constraints within the project are more important in relative terms than time and budget constraints for 'modelling', and to some extent for 'monetary valuation'. But more striking is the fact that the mean importance of data, time and budget constraints in method selection is low, varying from 'to some extent' to 'definitely not'. In other words, none of these constraints – as judged by case study leaders – were 'definitely' important in selecting the methods they used in case studies. We also looked at whether stakeholder-researcher familiarity increased the likelihood of decisive and technical design purposes of studies, using as a proxy the number of years they had worked with stakeholders before initiating OpenNESS. There was no visible effect for any of the study purposes (the results are shown in Supplementary Material, Figure S2). In section 6.4 we discuss alternative explanations for the information gap hypothesis.

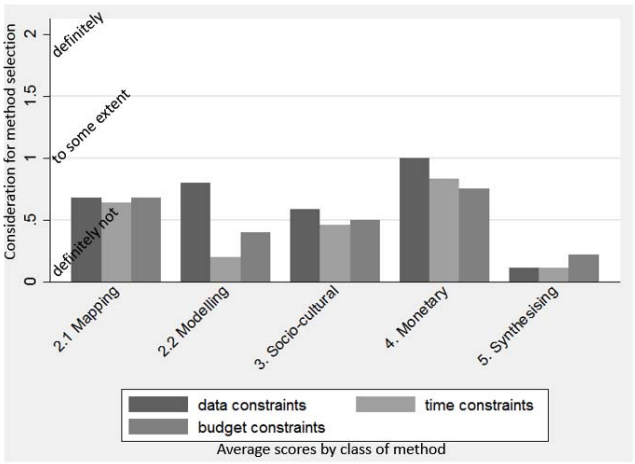

Figure 5 The importance of constraints in choice of methods relative to their alternatives as reported by case study leaders.

5.2 Participation in appraisal method selection

In this section, we further question information costs as an underlying mechanism for the information gap between ES appraisal and governance application, using responses to the stakeholder survey. Slightly fewer than half of the stakeholders interviewed "make decisions" about the ecosystem services investigated, while over half "contribute to decision-making"; the majority of stakeholders were "affected" by or "interested" in the ecosystem services issues assessed in the case studies (Figure S3). A majority of stakeholders found that appraisal methods lead to a "change in future vision in the area", "change in the way information and tools are used to support decisions", "change in decision-making" and "change in actions" (Figure S4). The stakeholder survey shows that the type and degree of governance support of ES appraisal methods depends on the level of interaction with researchers in the case study advisory board. Figure 6 shows that stakeholders' advisory role is weakly correlated with some, but not all, types of science-policy interaction. For example, stakeholder participation in *study design, knowledge co-production and recipient of information* increased increased signficantly for stakeholders were members of the advisory board(very applicable), versus those that were not(not applicable). This was not the case for method selection.

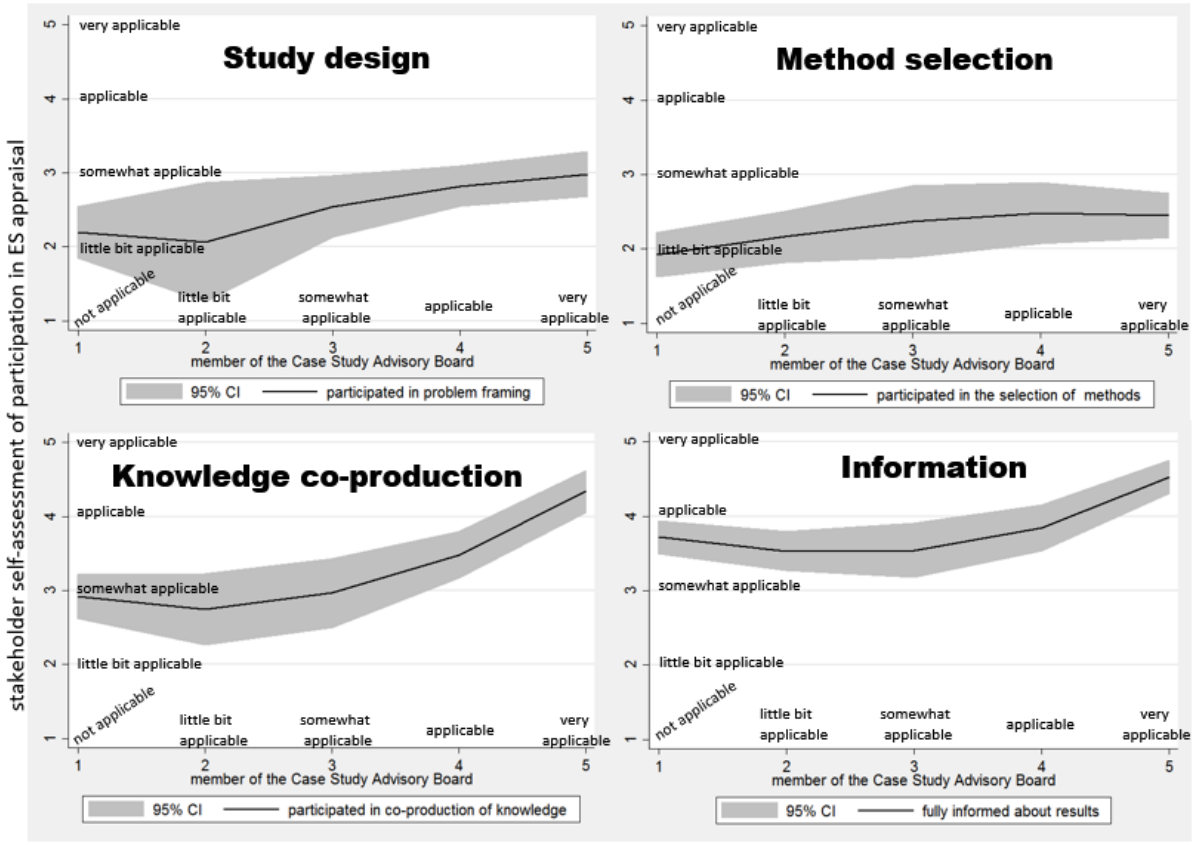

Figure 6 Stakeholder-researcher interactions in ES appraisals in OpenNESS reported by stakeholder respondents (n=246) STATA Two-way fractional polynomial plot with confidence interval (95%).

6 Discussion

6.1 Uses of methods for different purposes

To investigate the gap between ES appraisal research and its operational use, we analysed a large set of case study experiences. With the aim to explore the reasons behind the choice of different appraisal methods and their purposes, we formulated an information gap hypothesis predicting an ordering of methods that would be expected in the presence of increasing cumulative uncertainty and method requirements (Figure 2). We observe some significant ordering effects of methods across study purposes - in particular the difference between *decisive* and *technical design* purposes, and to a lesser extent the difference between *informative* and *decisive* purposes are apparent (Figure 3). Individual methods for *informative* purposes from different parts of the ES appraisal cascade are predominant. Synthesising methods that focus on valuation end-points are less frequent. This could also be explained by the novelty of ES appraisals for decision support (Ruckelshaus *et al.*, 2015).

Contrary to the expectation that limited use of appraisals for *decisive* and *technical design* purposes is due to constraints on information, we find that on average data, time and budget constraints were not perceived as strong constraints by case study coordinators (Figure 5). The lack of importance of these factors could be due to *explorative* and *informative* purposes being more important than *decisive* and *technical design* purposes in a research oriented project like OpenNESS. An additional explanation may be that as a research project, methods were relatively well-resourced, compared to what they would be

in e.g. a consultancy. Also, in several case studies that had longer term engagements with their case study sites there were opportunities to combine several financing sources, as well as having access to established data bases. Regarding whether stakeholder-researcher familiarity increased the likelihood of more challenging and more costly purposes of ES appraisal, we find no clear correlation between the importance of *decisive* or *technical design* study purposes, and the number of years researchers have worked with stakeholders before the OpenNESS case study started (Figure S2).

6.2 Co-design of purpose and method selection by stakeholders

The stakeholder survey results complemented that of the case study leader survey above. The OpenNESS project was operational in the sense of having active participation of stakeholders who are also decision-makers. We looked closer at the type of engagement and the way in which stakeholders interacted with the project. About half the stakeholders did not participate in the problem framing of the project, and slightly more than half did not participate in selecting the assessment methods (summing 'not applicable' and 'slightly applicable' responses). Stakeholder participation in scoping and selection of methods was generally low. This could be explained by the novelty of the project in a large proportion of the cases. In cases with established relationships before OpenNESS, scoping and method selection happened more often. On average the OpenNESS case studies scored higher in terms of co-production of knowledge, and in providing information inputs to case study advisory boards - what Primmer et al. 2015 called 'governance support', rather than explicit decision support.

Despite case study leaders perception of the decision-support role of their "synthesis" methods, there are indications that the decision-making role of these methods was potential, rather than actual (Figure S3, S4). As also reported in Dick et al. (this issue), only a bit more than 20% of stakeholders thought that OpenNESS assessment methods had already resulted in "a change in decision making" in the case study site. A further 40% thought that such a change was likely to take place, while the remainder did not assume such influence. The divergence between researchers' and stakeholders' perceptions of the decision-making role of appraisal methods can partly be traced back to the engagement processes that took place in the case studies. It highlights that involving the 'right' stakeholders (in terms of having decision-making power) is crucial to realize real decision-support. The question of which stakeholders to interact with to increase impact is complicated by hierarchical governance (Primmer et al. 2015). Still, our assessment is somewhat more optimistic than findings from the Natural Capital project (Ruckelshaus et al., 2015) where a 3-4 year research project was deemed insufficient to observe operational changes in decision-making.

The question remains whether the expectation of change in decision-making procedure can be interpreted as an actual sign of operationalisation. Based on stakeholder survey results the OpenNESS project was relatively successful in achieving decision-maker participation in the CABs (Figure S3). However, case studies were on the whole less effective in engaging them in study design and methods selection than in knowledge co-production and keeping them informed (Figure 5). Laurans and Mermet (2014) point out that studies for informative purposes can have a decisive effect over time as they can help reframe the policy debate. Ruckelshaus *et al.* (2015) argue that case study purposes evolve over the longer term as dynamic 'impact pathways'.

6.3 ES appraisal methods as value articulating institutions

Valuation methods range from biophysical, monetary, social to integrative methods. ES appraisal methods all generate results on 'importances', or what Gómez-Baggethun et al. (2014) call plural values of ES. Environmental appraisal methods in this broad sense are value articulating institutions (Vatn, 2005, 2009). Jacobs et al. (2017) evaluate application of 21 different valuation methods in various real-life case studies. They argue that the methods *generate* values rather than measuring them as an external objects. Every method articulates different values. Methods should be carefully selected and

combined to obtain a valid, just and credible decisions (Jacobs et al 2017). In this paper, we detail this argument further by identifying the variation in institutional (decision) contexts that a given appraisal method may be designed for.

The OpenNESS experience reveals that monetary valuation was applied in a small minority of case studies. Furthermore, where monetary valuation methods were applied, they were mainly conducted for informative purposes, in particular 'awareness-raising on the *importance of ecosystem services'* (Figure S1). This finding could be taken to confirm the decision-support gap in monetary valuation found in the review by Laurans et al. (2013), but we have admittedly a very small sample. It may also be a result of the classification of individual monetary valuation methods and benefit-cost analysis into separate categories. The majority of monetary valuation methods found in the academic literature by Laurans et al. (2013) may indeed be monetary valuation methods observed out of their potential decision-support role in applied benefit-cost analysis. Indeed, the formal documentation of the information value of monetary valuation for decision-making is demanding (Barton, 2007).

Looking at detailed purposes it is also interesting to note that mapping is perceived to be conducted, as we would expect, for informative purposes ('current state', 'awareness of ES importance'), but also for decisive purposes at a similar level of importance as 'synthesising' methods (lower left hand panel Figure S1). Within the group of decisive purposes, mapping methods were predominantly conducted for '*spatial targeting'*, but notably also as '*arguments for negotiation, shared norms & conflict reduction'*. This provides support for the assertion that mapping can also be a tool for framing and articulating values (Hauck *et al.*, 2013; Martín-López *et al.*, 2014). This was particularly true in OpenNESS where a number of participatory mapping methods were tested.

6.4 Information requirements of appraisal methods

While we found a significant ordering of methods across purposes with increasing requirements we could not explain the ordering using information costs. The constraints of data availability, information and resources are expected to be the strongest for novel decision-support problems. OpenNESS case studies often applied novel methods for their study sites (Dunford, 2018), with *explorative* purposes being among the most important for mapping, modelling, and socio-cultural methods. Here, there may be a selection bias from the point of view of decision-makers in the case study advisory boards. Novel studies need to be *explorative* until their reliability is tried and tested. While methods are being tested they may also be *informative*, to the extent that engagement with stakeholders is actively pursued. While OpenNESS was designed to promote co-production of knowledge, it seems reasonable to assume that cautious decision-makers will be less likely to use novel methods for *decisive* and *technical design* purposes, despite researchers' perception of their methods' potential. Figure 1 suggests that ES appraisal methods - and integrated valuation more widely (Jacobs et al. 2016) - must be perceived as reliable before being used by decision-makers to prioritise between alternative actions. Ecosystem services appraisal requires tailor-made forms and modes of societal involvement. These also require new qualities of public policies and new roles of governments (Verburg et al. 2016). These new roles include deliberative processes at various governance levels (Primmer et al. 2015) and new types of contracts, ample resources, sufficient knowledge on ES and long term commitment (Verburg et al. 2016). Our contention is that OpenNESS' success as a research project in applying novel knowledge co-production methods with stakeholders, to some extent 'self-selected' away from *decisive* and *technical design* purposes. The exception may be the handful of 'synthesis' methods, including for example multi-criteria decision analysis (MCDA). Some individual cases also actively pursued *decisive* and *technical design* purposes, e.g. (Barton *et al.*, 2015; Peri *et al.*, 2017; Dick, 2018).

6.5 Uncertainty and (dis)integrated valuation

While it is true that integrated environmental impact modeling can lead to very large cumulative uncertainty (Barton *et al.*, 2016), a DPSIR approach or cost-benefit analysis are particular interpretations of integrated valuation which emphasise conditionally dependent models (Figure 7). In other words, they are types of integrated valuation particularly prone to cumulative uncertainty. Multi-criteria decision analysis tries to mitigate this by considering biophysical, social and economic criteria separately, but may not escape uncertainty accumulation if criteria are causally related (as shown in Figure 7), or if they are scaled to a common unit of comparison (Saarikoski *et al.*, 2016). Method triangulation takes a different approach, treating mapping, biophysical modelling, socio-cultural asseessment and monetary valuation as independent plural value inputs to decision-making (Jacobs et al. 2017). Method triangulation cannot avoid correlated errors altogether, but it is suggested that consideration of independent impact criteria with no formal method of comparison can increase reliability of decisions (Jacobs et al. 2017).

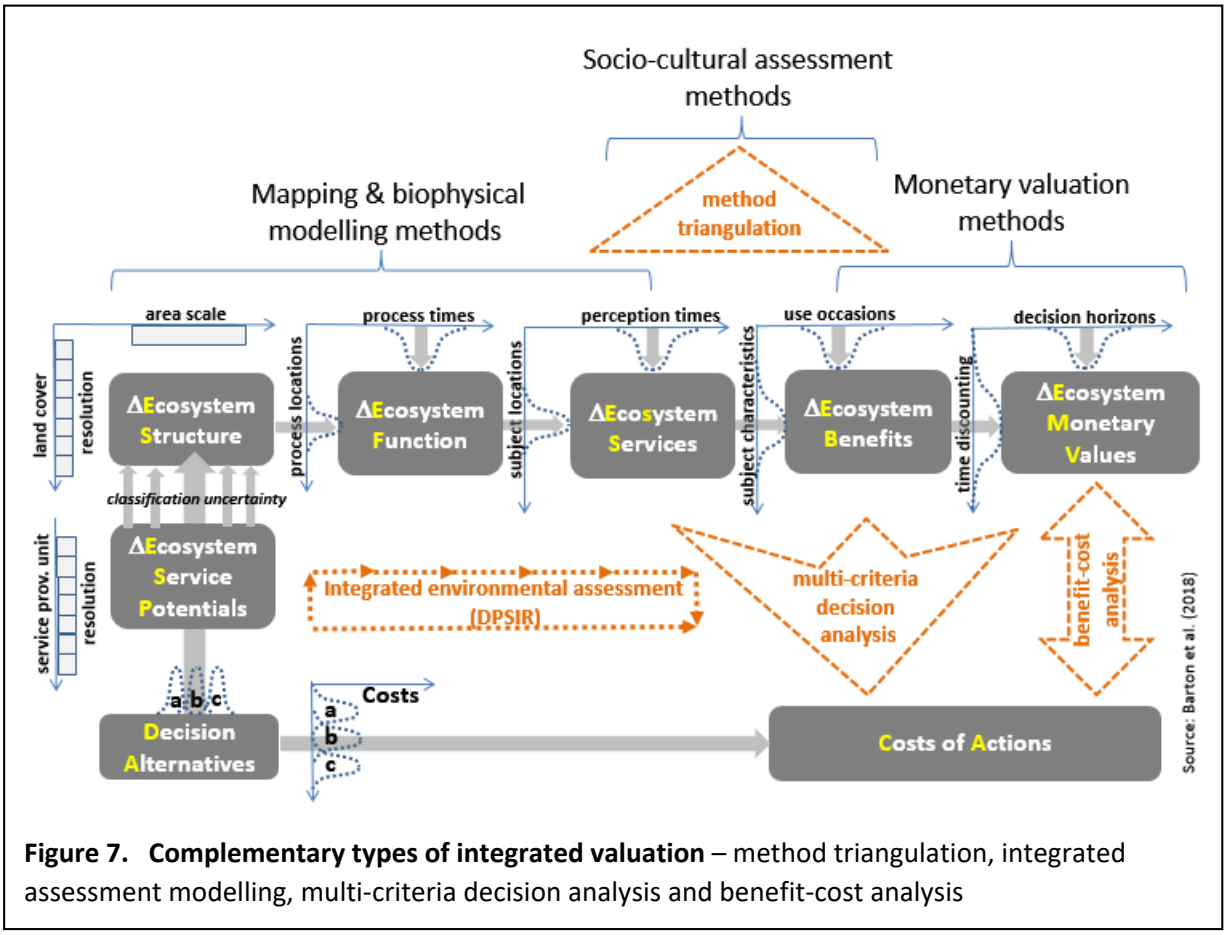

Figure 7. Complementary types of integrated valuation – method triangulation, integrated assessment modelling, multi-criteria decision analysis and benefit-cost analysis

A comparison of costs and benefits of actions– whether quantitative or qualitative - requires some kind of mapping-modelling-valuation synthesis procedure (IPBES, 2015). As we increasingly integrate parallel ES appraisal methods with the aim of decision support, we should expect stochastic events to combine and generate – explicitly or implicitly - a joint probability distribution of predicted policy outcomes. Decision-makers potentially make decision errors because of this probability distribution of outcomes. Knowing this, rational decision-makers will balance the information costs of more ES appraisals against the information value of avoiding "costly actions" (Type I) or "missed opportunities" (Type II) (see Supplementary Material S1). If decision-makers avoid designing one-shot decisions - and consultants avoid offering single-answer ES appraisals - then adaptive management, combined with method triangulation, iteration and updating of appraisals, offers a way to close the information gap.

7 Conclusions

We developed a theory of cumulative uncertainty in integrated ES apprisal based on a 'stairway' framework of ES appraisal. Successive ES appraisals - from mapping, modeling to benefit assessment and valuation - lead stakeholders to expect higher cumulative uncertainty in outcomes. Given an expectation about uncertainty, the more demanding the decision-support purpose of the ES appraisal, the lower the likelihood that the ES appraisal will be considered appropriate for decision-support. We defined an information gap where a particular study purpose's increasing requirements for accuracy and reliability, combine with increasing uncertainty of the ES appraisal methods' outcomes. In order to test the information gap hypothesis, we evaluated responses to survey data of case study coordinators and stakeholders from 26 case studies and 80 ecosystem services appraisals in a large integrated EU research project. We classified different ES appraisal methods in terms of their purpose - explorative, informative, decisive and design. We grouped ES appraisal methods into types – mapping, modeling, socio-cultural assessment, monetary valuation and synthesizing methods. We tested the likelihood of ES method types being appropriate for different purposes, as evaluated by the researchers who used them. We found some support for decreasing likelihood of ES appraisal methods' appropriateness with increasing requirements of study purpose. We discussed whether the likelihood of different method-purpose combinations were conditional on our theoretical expectations about uncertainty of outcomes. We do not find support for resource or time availability as an explanation for information gaps. Instead, stakeholder surveys show that decreasing appropriateness of ES appraisal methods for decisive and design purposes could be due to the research project interacting mostly with stakeholders outside the most decision-relevant contexts. Finally, we discussed how our theory of cumulative uncertainty in ES appraisal is based on a particular understanding of integrated valuation. Cumulative uncertainty and information gaps we uncovered may be particularly relevant for integrated impact assessment and benefit-cost analysis, while being less so for multi-criteria analysis and method triangulation. This leads us to suggest combining to strategies for addressing information gaps. Despite the empirical limitations of our study, we conclude that more systematic consideration of uncertainty and information costs in ES appraisal, together with the active pursuit of knowledge co-production with stakeholders, will contribute to narrowing the operational gap in ES appraisal, moving in the direction of integrated valuation.

Acknowledgements

This research was funded by the European Union EU FP7 project OpenNESS (Grant agreement no. 308428). MGL was fund by a grant from the Spanish National Institute for Agriculture and Food Research and Technology (INIA), which is co-funded by the Social European Fund.

References

Bagstad, K.J., Semmens, D.J., Waage, S., Winthrop, R., 2013. A comparative assessment of decision-support tools for ecosystem services quantification and valuation. Ecosystem Services 5, E27-E39.

Barton, D.N., 2007. How much is enough? The value information from benefit transfers in a policy context. In: Navrud, S., Ready, R. (Eds.), Environmental Value Transfer: Issues and Methods. Springer, Dordrecht, pp. 261-282.

Barton, D.N., Andersen, T., Bergland, O., Engebretsen, A., Moe, S.J., Orderud, G.I., Tominaga, K., Romstad, E., Vogt, R.D., 2016. Eutropia: Integrated Valuation of Lake Eutrophication Abatement Decisions Using a Bayesian Belief Network. pp. 297-320. Chap.14 in Niel, Z. P. (Ed.) Handbook of Applied Systems Science. Routledge, New York and London. DOI: 10.4324/9781315748771.ch14.

Barton, D.N., Saloranta, T., Moe, S.J., Eggestad, H.O., Kuikka, S., 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. Ecol Econ 66, 91-104.
Barton, D.N., Stange, E., Blumentrath, S., Traaholt, N.V., 2015. Economic valuation of ecosystem services for policy. A pilot study on green infrastructure in Oslo [NINA Report 1114, 77p]. http://www.openness-project.eu/node/78
Bateman, I., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D.W., Swandon, J., 2002. Economic Valuation with Stated Preferences Techniques. A Manual.
BBOP, 2009. Business, Biodiversity Offsets and BBOP: An Overview. Business and Biodiversity Offsets Programme (BBOP)., Washington, D.C.
Dick, J.e.a.c.-a., 2018. Stakeholders' perspectives on the operationalisation of the ecosystem service concept: results from 27 case studies. Ecosystem Services (in press).
Dunford, R.e.a.c.-a., 2018. Integrating methods for ecosystem service assessment: experiences from real world situations. Ecosystem Services (in press).
Fagerland, M.W., Sandvik, L., Mowinckel, P., 2011. Parametric methods outperformed non-parametric methods in comparisons of discrete numerical variables. Bmc Med Res Methodol 11.
Fontaine, C.M., Dendoncker, N., De Vreese, R., Jacquemin, I., Marek, A., Van Herzele, A., Devillet, G., Mortelmans, D., François, L., 2014. Towards participatory integrated valuation and modelling of ecosystem services under land-use change. . Journal of Land Use Science 9, 278-303.
Gómez-Baggethun, E., Barton, D.N., 2013. Classifying and valuing ecosystem services for urban planning. Ecol Econ 86, 235 – 245.
Gómez-Baggethun, E., Martín-López, B., Barton, D.N., Braat, L., Saarikoski, H., Kelemen, E., García-Llorente, M., E., J.v.d.B., Arias, P., Berry, P., Potschin, M., Keene, H., Dunford, R., Schröter-Schlaack, C., Harrison, P., 2014. EU FP7 OpenNESS Project Deliverable 4.1. . State-of-the-art report on integrated valuation of ecosystem services. European Commission FP7, 2014.
Grêt-Regamey, A., Altwegg, J., Sirén, E.A., van Strien, M.J., Weibel, B., 2016. Integrating ecosystem services into spatial planning—A spatial decision support tool. . Landscape and Urban Planning, (online version) http://dx.doi.org/10.1016/j.landurbplan.2016.05.003.
Haines-Young, R., Potschin, M., 2010. The links between biodiversity, ecosystem services and human well-being'. In: Raffaelli, D., Frid, C. (Eds.), Ecosystem Ecology: a new synthesis. Cambridge University Press, Cambridge, pp. 110–139.
Harrison, P.A.e.a.c.-a., 2018. Selecting methods for ecosystem service assessment: A decision tree approach. Ecosystem Services (in press).
Hauck, J., Görg, C., Varjopuro, R., Ratamäki, O., Maes, J., Wittmer, H., Jax, K., 2013. 2013. "Maps have an air of authority": Potential benefits and challenges of ecosystem service maps at different levels of decision making. Ecosystem Services, 25-32.
Hausknost, D., Grima, N., Singh, S.J., 2017. The political dimensions of Payments for Ecosystem Services (PES): Cascade or stairway? Ecol Econ 131, 109-118.
IPBES, 2015. Preliminary guide regarding diverse conceptualization of multiple values of nature and its benefits, including biodiversity and ecosystem functions and services (deliverable 3 (d)). Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.
Jacobs, S., Dendoncker, N., Martin-Lopez, B., Barton, D.N., Gomez-Baggethun, E., Boeraeve, F., McGrath, F.L., Vierikkoh, K., Geneletti, D., Sevecke, K.J., Pipart, N., Primmer, E., Mederly, P., Schmidt, S., Aragao, A., Baral, H., Bark, R.H., Briceno, T., Brogna, D., Cabral, P., De Vreese, R., Liquete, C., Mueller, H., Peh, K.S.H., Phelan, A., Rincon, A.R., Rogers, S.H., Turkelboom, F., Van Reeth, W., van Zanten, B.T., Wam, H.K., Washbourne, C.L., 2016. A new valuation school: Integrating diverse values of nature in resource and land use decisions. Ecosystem Services 22, 213-220.
Jacobs, S.e.a., 2016. A new valuation school: Integrating diverse values of nature in resource and land use decisions. Ecosystem Services (in press).
Jacobs, S.e.a.c.-a., 2018. The means determine the end – pursuing integrated valuation in practice. Ecosystem Services (in press).

Laurans, Y., Rankovic, A., Billé, R., Pirard, R., Mermet, L., 2013. Use of ecosystem services economic valuation for decision making: Questioning a literature blindspot. Journal of Environmental Management 119, 208-219.

Martin-Lopez, B., Gomez-Baggethun, E., Garcia-Llorente, M., Montes, C., 2014. Trade-offs across value-domains in ecosystem services assessment. Ecol Indic 37, 220-228.

Martín-López, B., Gómez-Baggethun, E., García-Llorente, M., Montes, C., 2014. Trade-offs across value-domains in ecosystem services assessment. Ecological Indicators 37 (2014) 220– 228.

Martinez-Harms, M.J., Bryan, B.A., Balvanera, P., Law, E.A., Rhodes, J.R., Possingham, H.P., Wilson, K.A., 2015. Making decisions for managing ecosystem services. Biological Conservation 184, 229-238.

Nassl, M., Loffler, J., 2015. Ecosystem services in coupled social-ecological systems: Closing the cycle of service provision and societal feedback. Ambio 44, 737-749.

Pascual, U.e.a.c.-a., 2017. Valuing nature's contributions to people: the IPBES approach. . Current Opinion in Environmental Sustainability. https://doi.org/10.1016/j.cosust.2016.12.006 26-27, 7-16.

Peri, P., López, D., Rusch, V.E., Rusch, G.M., Martínez-Pastur, G., 2017. State and transition model approach in native forests of Southern Patagonia: Linking ecosystem services, thresholds and resilience. . International Journal of Biodiversity Science, Ecosystem Services & Management, 105-118.

Potschin, M., Haines-Young, R., 2016. Defining and Measuring Ecosystem Services. In: Potschin, M., Haines-Young, R., Fish, R., Turner, R.K. (Eds.), Routledge Handbook of Ecosystem Services. Routledge, p. 640.

Priess, J.A., Mimler, M., Klein, A.M., Schwarze, S., Tscharntke, T., Steffan-Dewenter, I., 2007. Linking deforestation scenarios to pollination services and economic returns in coffee agroforestry systems. Ecol Appl 17, 407-417.

Primmer, E., Jokinen, P., Blicharska , M., Barton, D.N., Bugter, R., Potschin, M., 2015. Governance of Ecosystem Services: A framework for empirical analysis. Ecosystem Services http://dx.doi.org/10.1016/j.ecoser.2015.05.002i.

Rasch, D., Teuscher, F., Guiard, V., 2007. How robust are tests for two independent samples? J Stat Plan Infer 137, 2706-2720.

Ruckelshaus, M., McKenzie, E., Tallis, H., Guerry, A., Daily, G., Kareiva, P., Polasky, S., Ricketts, T., Bhagabati, N., S.A.Wood, J.Bernhardt, 2015. Notes from the field: Lessons learned from using ecosystem service approaches to inform real-world decisions. Ecol Econ 115, 11–21.

Saarikoski, H., Mustajoki, J., Barton, D.N., Geneletti, D., Langemeyer, J., Gomez-Baggethun, E., Marttunen, M., Antunes, P., Keune, H., Santos, R., 2016. Multi-Criteria Decision Analysis and Cost-Benefit Analysis: Comparing alternative frameworks for integrated valuation of ecosystem services. Ecosystem Services.

Schröter, M., Barton, D.N., Remme, R.P., Hein, L., 2014. Accounting for capacity and flow of ecosystem services: A conceptual model and a case study for Telemark, Norway. Ecol Indic 36, 539-551.

Soler, R.M., Schindler, S., Lencinas, M.V., Peri, P.L., Martínez Pastur, G., 2016. Why biodiversity increases after variable retention harvesting: A metaanalysis for southern Patagonian forests. Forest Ecology and Management, 161–169.

Spangenberg, J.H., von Haaren, C., Settele, J., 2014. The ecosystem service cascade: Further developing the metaphor. Integrating societal processes to accommodate social processes and planning, and the case of bioenergy. Ecol Econ 104, 22-32.

Stata, 2013. Stata graphics reference manual. Release 13. https://www.stata.com/manuals13/g.pdf Stata Press Publication, StataCorp LP, College Station, Texas.

TEEB, 2010. The Economics of Ecosystems and biodiversity: mainstreaming the economics of nature: a synthesis of the approach, conclusions and recommendations of TEEB. TEEB ocotber 2010.

van Oudenhoven, A.P.E., Siahainenia, A.J., Sualia, I., Tonneijck, F.H., van der Ploeg, S., de Groot, R.S., Alkemade, R., Leemans, R., 2015. Effects of different management regimes on mangrove ecosystem services in Java, Indonesia. Ocean & Coastal Management 116, 353–367.

Vatn, A., 2005. Institutions and the environment. Edward Elgar Pub., Cheltenham, UK ; Northampton, MA.
Vatn, A., 2009. An institutional analysis of methods for environmental appraisal. Ecological Economics, 68, 2207–2215.
Verburg, R., Selnes, T., Verweij, P., 2016. Governing ecosystem services: National and local lessons from policy appraisal and implementation. Ecosystem Services 18, 186-197.
Yee, S.H., Carriger, J.F., Bradley, P., Fisher, W.S., Dyson, B., 2015. Developing scientific information to support decisions for sustainable coral reef ecosystem services. Ecol Econ 115, 39-50.

Supplementary Material

S1 Conceptual basis for the value of information for decisive purposes

Hypothesis testing outcomes of an action situation		Reality *(ex post)*	
		H0 true Costs > Benefits	H1 true Benefits > Costs
Valuation estimates *(ex ante)*	H0 true Costs > Benefits	«Status quo correct»	**Type II error** «Missed opportunity»
	H1 true Benefits > Costs	**Type I error** «Costly action»	«Action correct»

Table S1. The value of information for decisive purposes – how much does it contribute to reducing the likelihood of "costly actions" or "missed opportunities"

Uncertainty – poor accuracy and reliability of ES appraisal - can lead to "costly actions" or "missed opportunities". Table S1 defines type I and type II errors, respectively. A 'well-defined' decision is one in which action alternatives, their benefits and costs, can be defined sufficiently to test a null-hypothesis about the net benefits of an action (e.g. landuse change). Information value can be measured in terms of reducing the likelihood of making a type I or II error, multiplied by the foregone benefits or avoided costs of those errors. Appropriateness of integrating further information (appraisals) is then a question about whether this value outweighs additional information costs.

S2 Definition of method types

Mapping (n=24)	**Biophysical Modelling** (n=10)	**Monetary valuation** (n=12)	**Socio-cultural valuation** (n=25)	**Synthetising methods** (n=9)
ESTIMAP is a set of spatially-explicit models each of which can be run separately for the assessment of different ecosystem services at the European or regional scale (Zulian et al. 2013a,b); **Matrix approaches** Simple matrix mapping links a spreadsheet of ecosystem service supply/demand indicators by land cover category to a GIS map, to generate maps of ecosystem service supply, demand and balance (supply minus demand). The indicators can be derived from scientific data or can be scores based on local or expert knowledge (e.g. Burkhard et al. 2012). Advanced matrix mapping approaches build on simple matrix mapping approaches through incorporating multiple sources of spatial datasets. An example of such an approach used in OpenNESS is GreenFrame which was developed to assess spatial variation in ecosystem service provision potential of green infrastructure in spatial planning (Kopperoinen et al. 2014). The method utilises an extensive set of spatial datasets grouped into themes combined with both scientific experts' and local actors' scorings. **QUICKScan,** an interactive GIS-based modelling tool designed to be used in a facilitated workshop to enable policy-makers, experts and stakeholders to jointly explore the	**SITE Landuse model** (SImulation of Terrestrial Environments mode is based on a cellular automata approach, in which the biophysical environment is represented by a grid. Simulated land-use decisions are based on a multicriteria suitability analysis of (i) the natural environment and (ii) the socio-economic conditions (Priess *et al.*, 2007) **Bayesian belief network** BBNs are based on simple diagrams consisting of nodes representing processes or factors, and links showing how the processes are connected, typically derived using expert knowledge. For ecosystem service assessment, the nodes may represent factors determining the supply or demand of sevices, such as land cover or soil type, as well as outcomes such as water flow, and information on values and preferences. Each link is assigned a weight to indicate the probability that the link is true, or the strength of the causal relationship, so that uncertainty is explicitly taken into account (see Smith et al. this issue). BBNs can be linked to GIS to undertake spatial analysis. **Climate envelope modelling** Species distribution models (also known as climate envelope or niche models) are used to project the potential impacts of climate changes on species. They are correlative models, based on the statistical relationship between a species' distribution and the climate,	**Time use value** Visitors time use at recreation sites is obtained through a survey, and multiplied by a monetary estimate of time value, for example average wage rate after taxes. **Value transfer** Benefits transfer (BT), or more generally - value transfer (VT) - refers to applying quantitative estimates of ecosystem service values from existing studies to another context (see Barton, 1999). **Cost-based,** Mitigation cost-based valuation methods are a group of 'exchange-based' techniques that use the cost of actual measures to maintain ecosystem service provision as a proxy for the value of actions undertaken in the mitigation hierarchy (BBOP, 2009), including actions to avoid, minimise, restore or replace ecosystems and their services that are potentially at risk in connection with a development. As a valuation technique, the costs of actions are taken as proxies for the value of the ecosystem services lost. This group of methods therefore includes: (i) restoration cost; (ii) replacement cost; and (iii) clean-up cost. **Revealed preference** Values of ecosystem services are revealed indirectly through purchases (e.g. house prices) or behaviour (travel costs). Examples used in OpenNESS include: (i) hedonic pricing, which is the study of multi-correlation between environmental characteristics of a good and its sales price; and (ii) travel cost methods (TCM), which are based on the observation that	**Questionnaire & narrative analysis** Narrative methods aim to capture the importance of ecosystem services to people through their own stories and direct actions (both verbally and visually) (see de Oliviera & Berkes 2014). **Photoseries analysis** Photo-sharing websites such as Flickr, Panoramio and Instagram are used to provide revealed preferences for cultural ecosystem services, assuming that visitors are attracted by the location where they take photographs e.g. Richards & Friess, 2015). **Preference assessment** Preference assessment is a direct and quantitative consultative method for analysing perceptions, knowledge and associated values of ecosystem service demand or use (or even social motivations for maintaining the service) without using economic metrics. Data is collected through surveys using a consultative approach with different variations, such as free-listing exercises, ecosystem service ranking, rating or ecosystem service selection (e.g. Martín-López et al. 2012). **Time use** This method is an innovation of the conventional stated preference techniques taken from the contingent valuation approach. Surveys are used to estimate the value of ecosystem services by asking people how much time they would be willing to invest for a change in the quantity or quality of a given service (García-Llorente et al. 2016). **Q-method** a research method used in psychology and in social sciences to study people's "subjectivity" using a form of factor	**Integrated mapping-modelling (INVEST, EcoServ)** InVEST, a set of models for mapping and valuing the ecological or economic value of multiple ecosystem services at a local to regional scale (Sharp et al. 2016). EcoServ-GIS is a Geographic Information System (GIS) toolkit for mapping ecosystem services at a county or regional scale. It uses input GIS/map data to generate fine-scale maps that illustrate human need or demand for ecosystem services as well as the capacity of the natural environment to provide them. http://ecosystemsknowledge.net/ **Scenario development** Scenarios are defined within the OpenNESS project as 'plausible, simplified description(s) of how the future may develop, based on a coherent and internally consistent set of assumptions about key driving forces'. Engaging with stakeholders helps to formulate scenarios which are consistent with the stakeholder perspectives (Priess & Hauck 2015). **Multi-criteria decision analysis** MCDA is an umbrella term to describe a collection of formal approaches which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter. Spatial MCDA are carried out in GIS in order to enable a visualization of the multiple criteria. **Benefit-cost analysis** Benefit-cost analysis is a decision-support tool for screening alternatives

impacts of different policy options on ecosystem services (Verweij et al. 2012); **Smarthphone Apps** Blue-green factor app let's a smarphone user map and scores blue-green structures at property level and compare with minimum zoning requirements. ESM-App/MapNat lets smartphone users map important places, paths and polygons for ecosystem services **PPGIS** Participatory GIS (PGIS) or Public Participation GIS (PPGIS) is a form of deliberative mapping. It uses web-based surveys, face to face interviews and workshops to integrate perceptions, knowledge (local-based or technical) and values of different stakeholders and presents the outputs in the form of a map of ecosystem services (see Brown and Fagerholm, 2015).	with the assumption that the species' distribution is in equilibrium with the climatic conditions. **Hydrological models,** hydrological models, such as the Soil and Water Assessment Tool (SWAT; Francesconi et al. 2016); **Meta-analysis** based on systematic literature review advantages and trade-offs of ES and biodiversity from similar studies were applied to a number of silivcultural research plots and current management prescriptions(Soler *et al.*, 2016)	recreational services can only be realised through physical access to nature. **Stated preference** Stated preference valuation is a family of techniques which use individual respondents' stated hypothetical choices to estimate change in the utility associated with a proposed increase in quality or quantity of an ecosystem service or bundle of services (Bateman *et al.*, 2002).	analysis to look for correlations between subjects across a sample of variables **Deliberative valuation** Deliberative valuation is not one particular valuation method, but it is a valuation paradigm providing a framework to combine various tools and techniques that bridge citizens and academia, as well as different disciplines within science. Such methods invite stakeholders and citizens (the general public) to form their preferences for ecosystem services together through an open dialogue with others (see Wilson & Howarth, 2002). **Stakeholder analysis** Ecosystem services scoring using card game **ES cascade focus group**	by their internal rate of return, or ranking alternatives by their discounted benefit/cost ratio or net present value (see Boardman et al. 2006).

S3 Selected stakeholder survey questions

Question wording for stakeholder survey questions reported in this article are given below where not reported in the article text itself.

1.Please rate your level of participation in the OpenNESS Oslo case study*

1.1 I participated in problem framing of the research conducted
1.2. I participated in the selection of research methods/approaches used
1.3 I participated in co-production of knowledge (i.e. attended workshops/meetings/ stakeholder engagement activities)
1.4. I was fully informed about results
1.5 I am a member of the Case Study Advisory Board
1.6. I participated in another way. Please specify:

3.Please rate your role in relation to the issue investigated: mapping and evaluation of ecosystem services from urban blue-green infrastructure*

3.1. I make decisions related to the issue investigated
3.2 I contribute to decision making related to the issue investigated
3.3. I am affected by the issue investigated
3.4. I am interested in the issue investigated
3.5. I have another interest not mentioned above (Please specify)

*Responses scored level of applicability as Not applicable -Little bit applicable -Somewhat applicable – Applicable -Very applicable

8 Please rate your assessment of the intended or already realized use of the OpenNESS research

The OpenNESS research on "operationalising the ecosystem service and natural capital concepts" resulted in the following*:
8.1. The OpenNESS research resulted in a change in future vision in the area (e.g. vision document on the future landscape, policy etc.) (e.g. vision document on the future landscape, policy etc.)
8.2. The OpenNESS research resulted in a change in the way information and tools are used to support decisions
8.3. The OpenNESS research resulted in a change in decision making
8.4. The OpenNESS research resulted in a change in actions
8.5. The OpenNESS research resulted in another positive or negative impact(s).
Please specify:

*Responses scored as It is very unlikely to take place - Probably not take place -Not sure - It will probably take place - It already took place

S4 Additional descriptive data

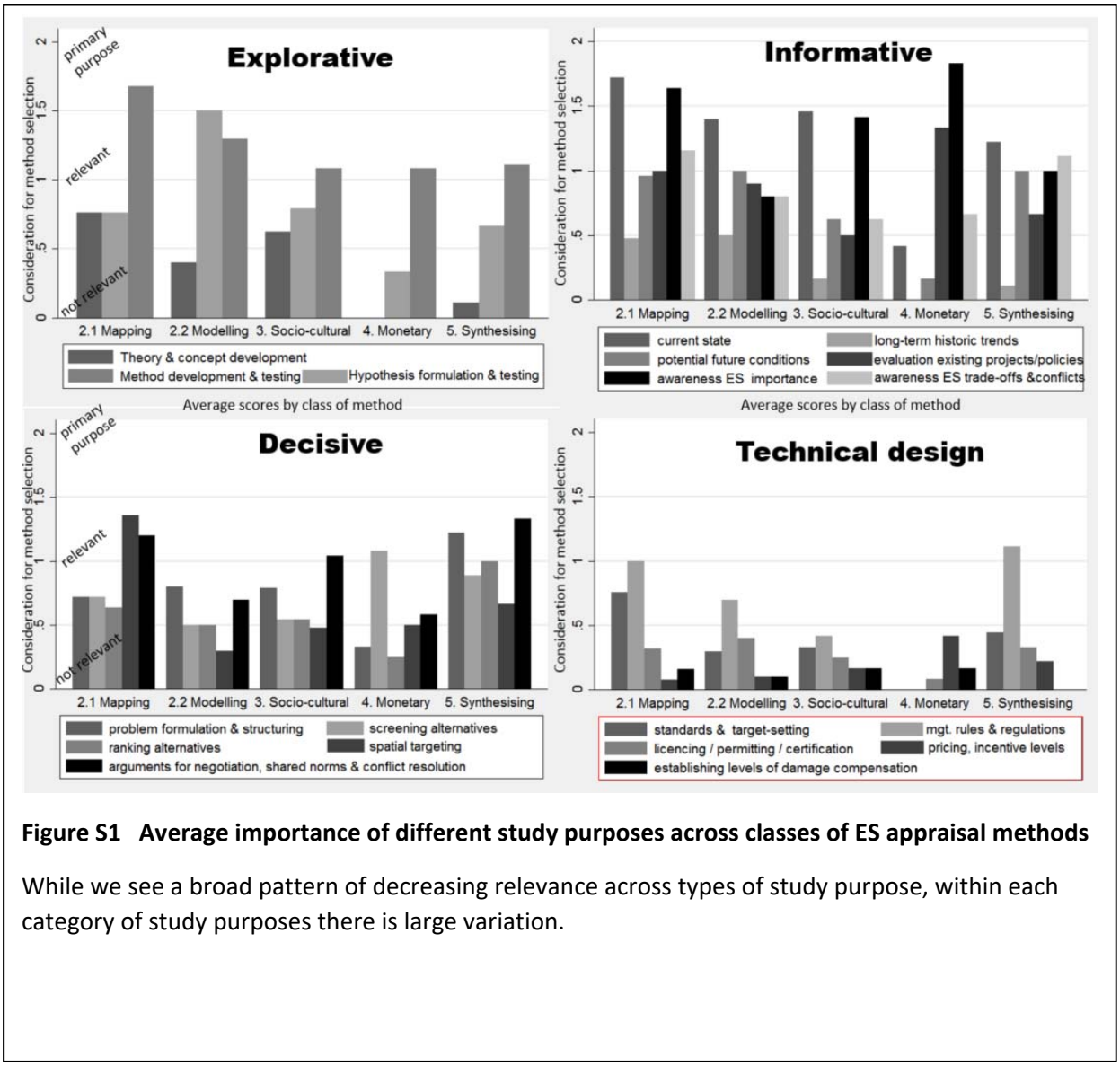

Figure S1 Average importance of different study purposes across classes of ES appraisal methods

While we see a broad pattern of decreasing relevance across types of study purpose, within each category of study purposes there is large variation.

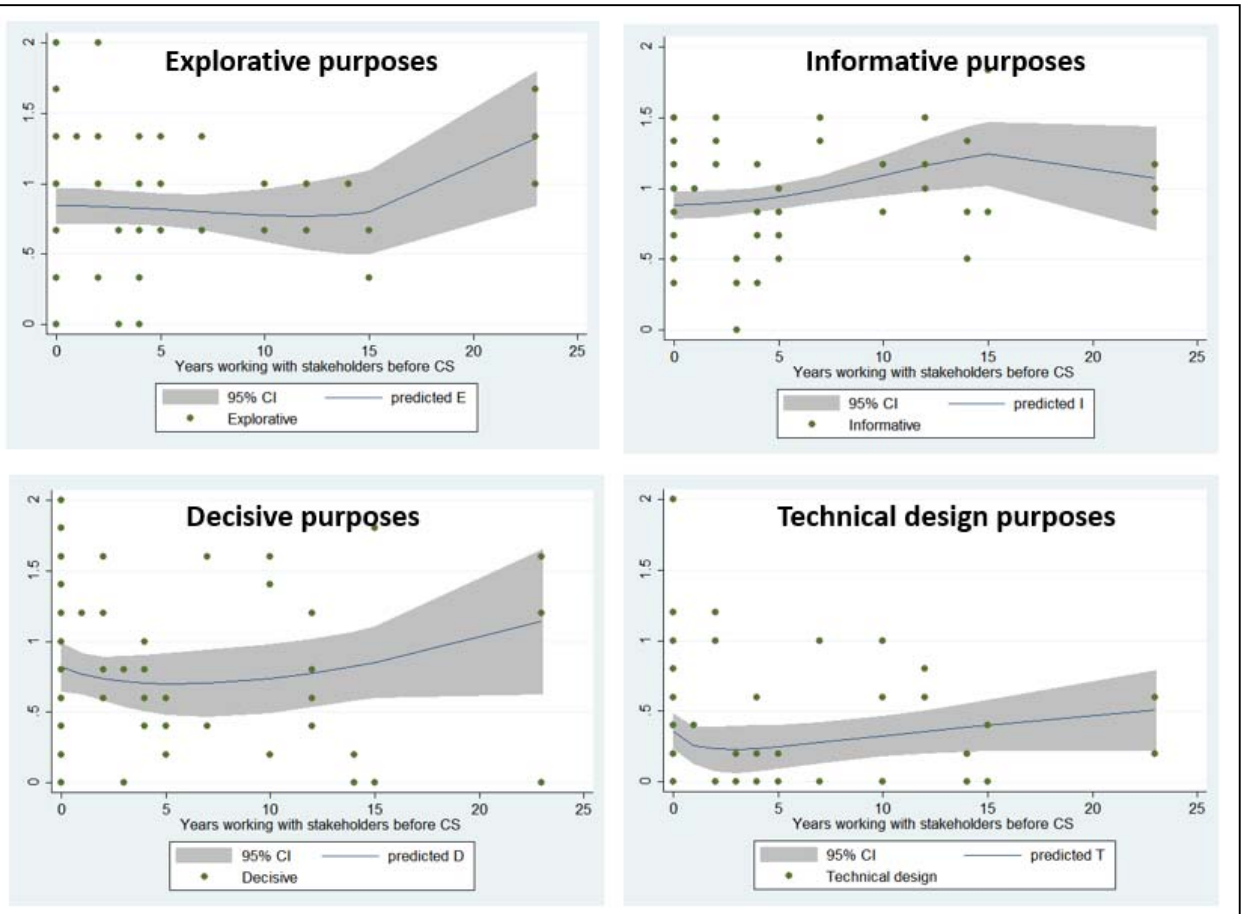

Figure S2 Do more challenging purposes of ES appraisal increase with stakeholder-researcher familiarity ? (multinomial function fit with confidence intervals 95%)

There are no clear correlations between decisive or technical design use of ES appraisal methods and years researchers have worked with stakeholders before the OpenNESS case study started.

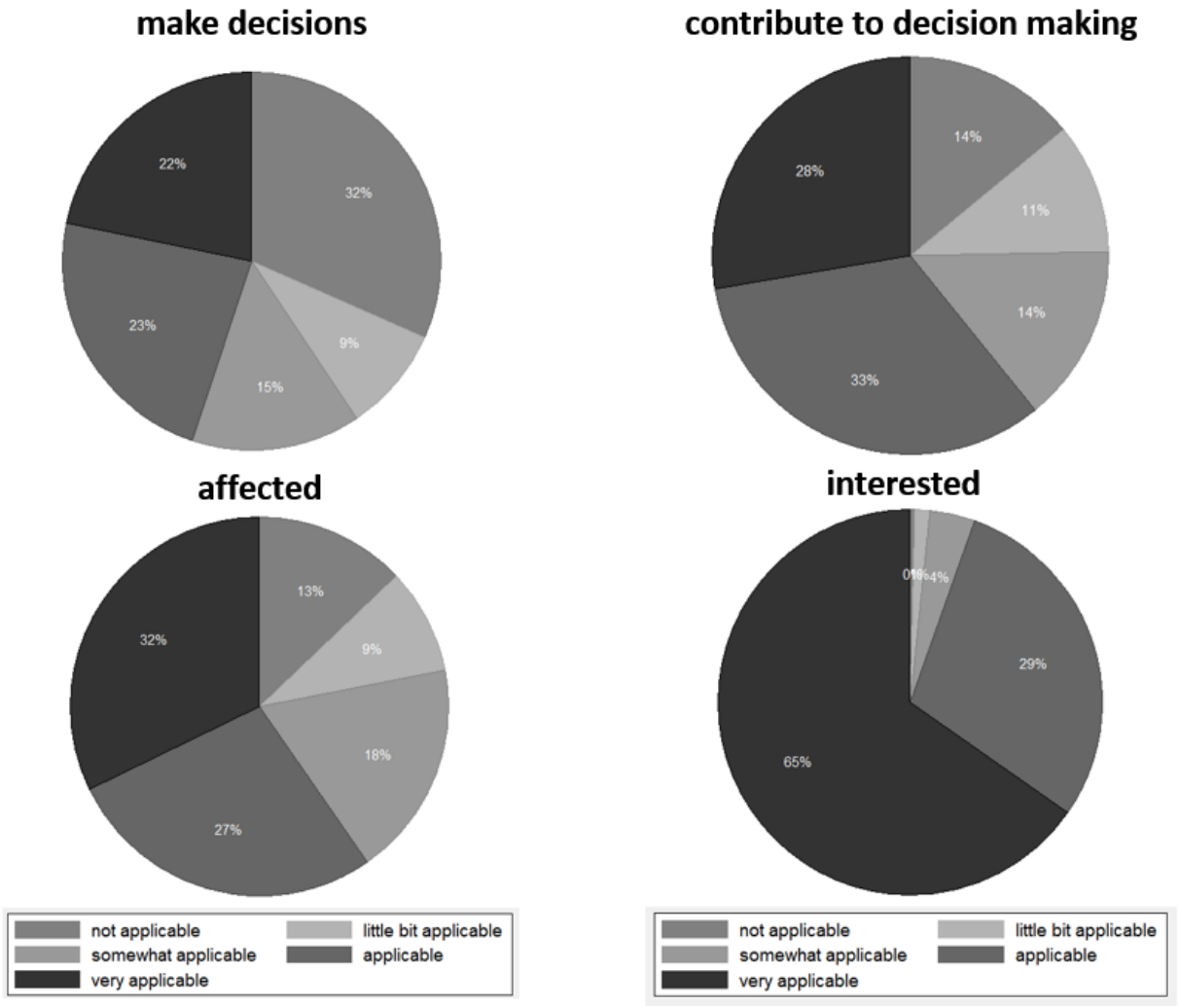

Figure S3 Stakeholder participation in decision-making in the case study area

The stakeholders who interacted with researchers in case studies had diverse roles with regards to decision-making on ecosystem services management. Figure S3 shows the proportion of respondents in the stakeholder survey who "make decisions", "contribute to decision-making", are "affected" by or "interested" in ecosystem services issues assessed by the case studies. A bit less than half of the stakeholders interviewed "make decisions" about the ecosystem services investigated, while over half "contribute to decision-making". The majority of stakeholders were "affected" by or "interested" in the ecosystem services issues assessed in the case studies.

1

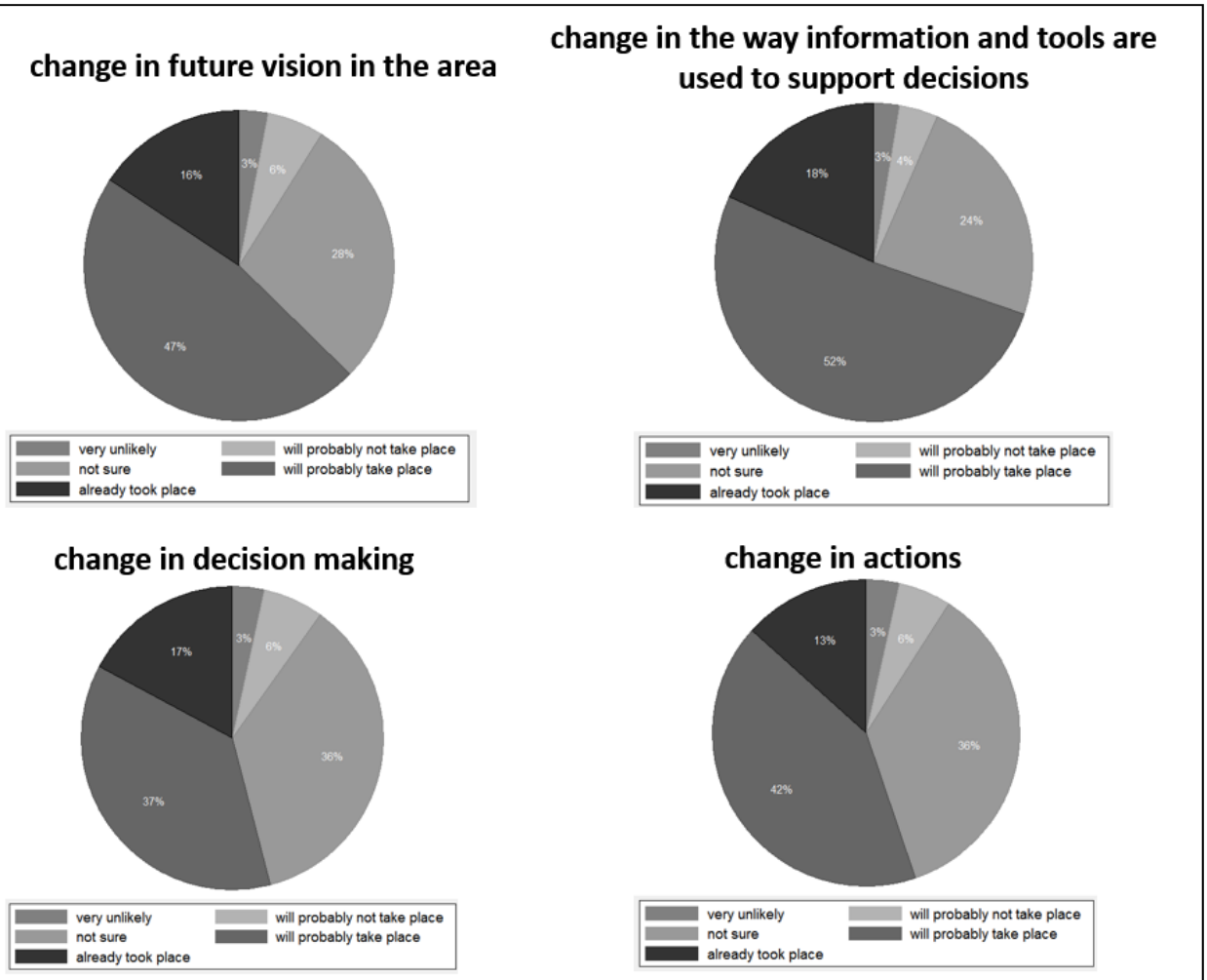

Figure S4 ES appraisal methods' impact on decision-making and actions Figure S4 shows how the stakeholders as a group assessed the impact of the ES appraisal methods applied in the case studies. A majority of stakeholders found that appraisal methods lead to a "change in future vision in the area", "change in the way information and tools are used to support decisions", "change in decision-making" and "change in actions". The proportions were slightly lower when it came to actual decision-making or action taking place.

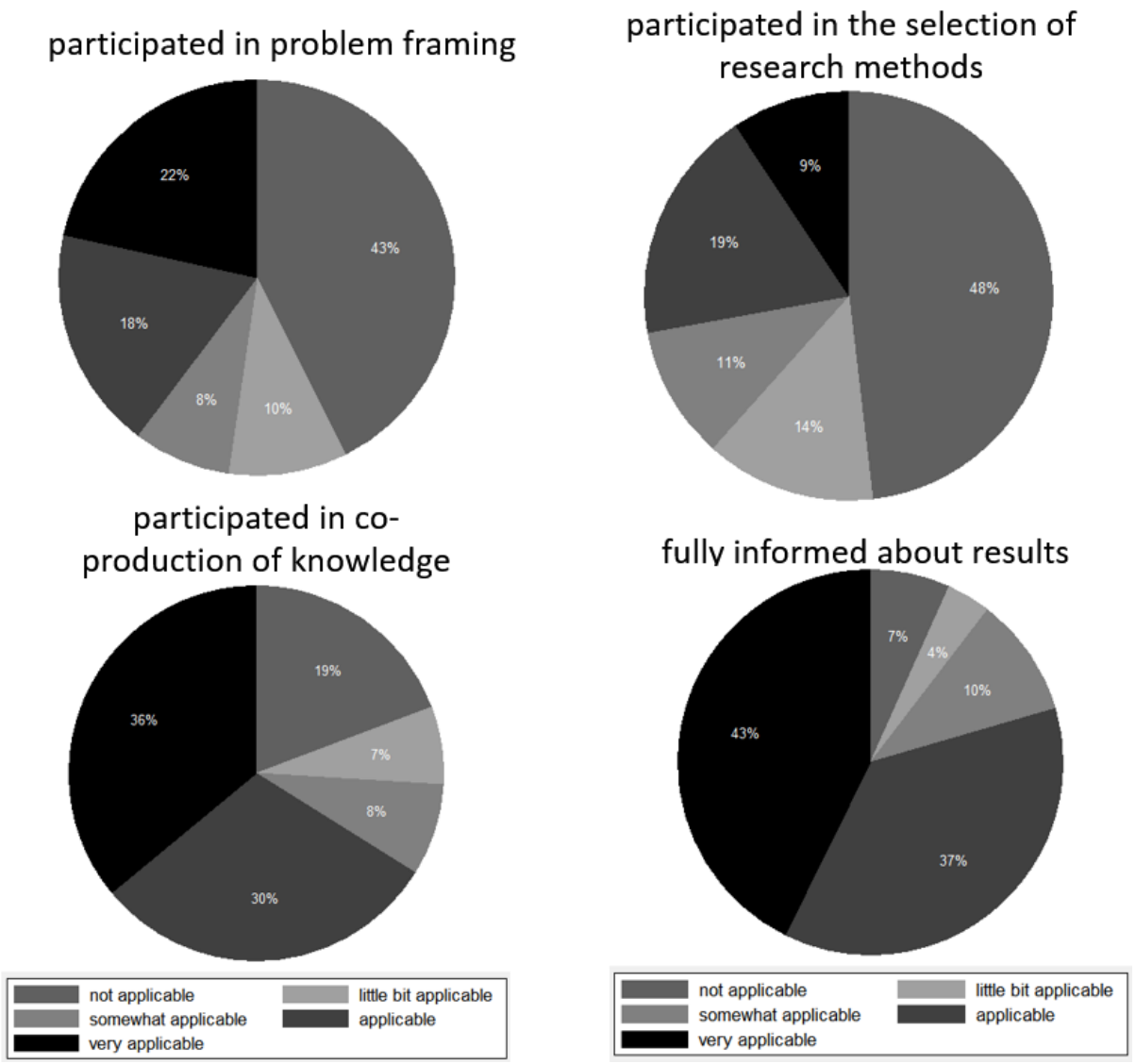

Figure S5 Stakeholder participation in study design and method selection

Figure S5 shows the extent to which stakeholders as a group partipated in “problem framing”, “method selection”, “co-production of knowledge” (i.e. attended workshops/meetings/ stakeholder engagement activities) or was simply ”fully informed about results”. A minority of stakeholders participated in problem framing or selection of methods.