natureresearch
https://doi.org/10.1038/s41893-020-00625-y

Supplementary information

Increasing decision relevance of ecosystem service science

In the format provided by the authors and unedited

Increasing decision-relevance of ecosystem service science

Supplementary Information

Contents

Supplementary Information
Appendix A. Self-Organizing Map

Methods

Article abstracts from a corpus of more than 12,000 documents, containing the terms "ecosystem or environmental service", were inserted into the Abstract Map Open (AMO) software. A Porter stemming algorithm was then applied to remove suffixes from words and count the frequency of each term. We removed terms that did not appear more than 100 times in the collection due to computational limitations of the training software in handling the appropriate size of the artificial neural network (ANN). This resulted in a data abstraction table representing each document as an n-dimensional vector of terms ($n = 645$).

Self-organizing map (SOM) is a method that trains an ANN to produce a representation of low-dimensional space, similar to multidimensional scaling (Kohonen 1982). Our 2D representation of training samples uses unsupervised learning to preserve the topology of the input based on similarities and difference across a neighborhood network. A version of the Self-Organizing Map Program Package (SOM PAK) built for ESRI's ArcGIS software was applied to train the SOM in two rounds using a Euclidean distance metric (Lacayo-Emery 2011). Each round had a different length of training, initial learning rate (α), and initial neighborhood radius (r). Following best practices that consider both the size and number of dimensions of the initialized SOM, we used the following training parameters:

- Round 1: 10 times through original length; $\alpha = 0.04$; r = length
- Round 2: 100 times through original length; $\alpha = 0.03$; r = 1/10 of length

Once the SOM is trained the resulting output dot symbols, called *best matching units (BMU)*, indicate where the more than 12,000 abstracts fall in 2D space. Each hotspot map, called a *component plane*, represent the strength of each term and the likelihood that each abstract contains this term. Using a surface that depicts the maximum value for all 645 terms we applied the ArcGIS Hillshade function to generate a term dominance landscape. This relief map visualizes term dominance using a series of cartographic design techniques first described by Skupin (2004). Future research will build on these techniques to quantify and map how the ecosystem services knowledge space has changed over time, including potential implications for policy and practice (Skupin et al. 2013).

References

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, *43*(1), 59-69.

Lacayo-Emery, M. A. (2011). An integrated toolset for exploration of spatio-temporal data using self-organizing maps and GIS. San Diego State University Graduate Thesis. Available from: https://www.researchgate.net/publication/275240068_An_Integrated_Toolset_for_Exploration_of_Spatio-_Temporal_Data_Using_Self-organizing_Maps_and_GIS

Skupin, André. (2004). "The World of Geography: Visualizing a Knowledge Domain with Cartographic Means." *PNAS* 101 (Suppl. 1): 5274-5278.

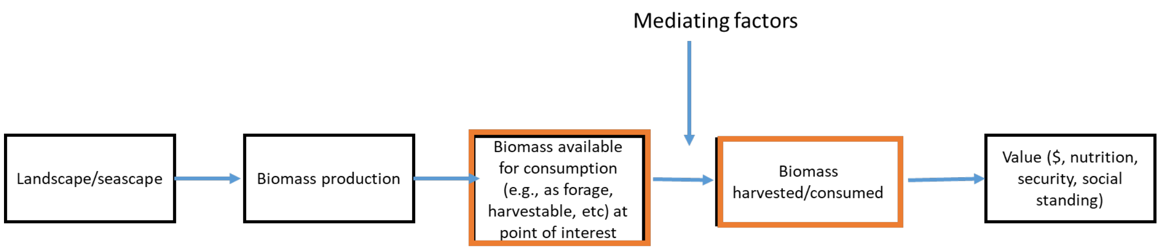
Skupin, A., Biberstine, J. R., & Börner, K. (2013). Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PloS one*, *8*(3), e58779.

Verutes, G. and Fisher, B. (In prep.) Using self-organizing maps to visualize the changing focus of ecosystem services research over time: implications for policy and practice.

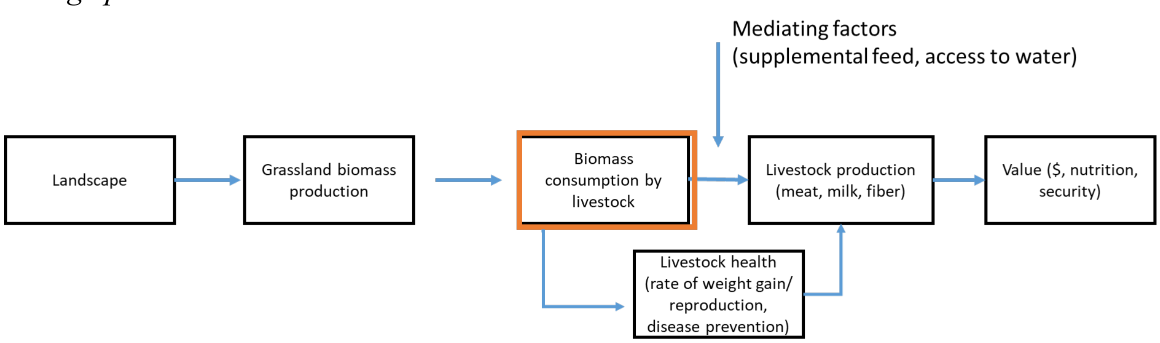
Appendix B. Service-Specific Chains

Conceptual chains for the production of a range of ecosystem services. Boxes outlined in orange indicate the point at which a biophysical benefit is delivered. For chains in which multiple metrics of biophysical service delivery were possible, multiple orange boxes are present. Boxes to the left of the orange box(es) make up possible components of ecological production functions. Boxes to the left of the orange box(es) represent possible components of socio-economic utility functions (SUFs). Examples of service-specific mediating factors are indicated for some chains. These chains are meant as illustrative examples rather than a comprehensive catalog of all possible chains and their components.

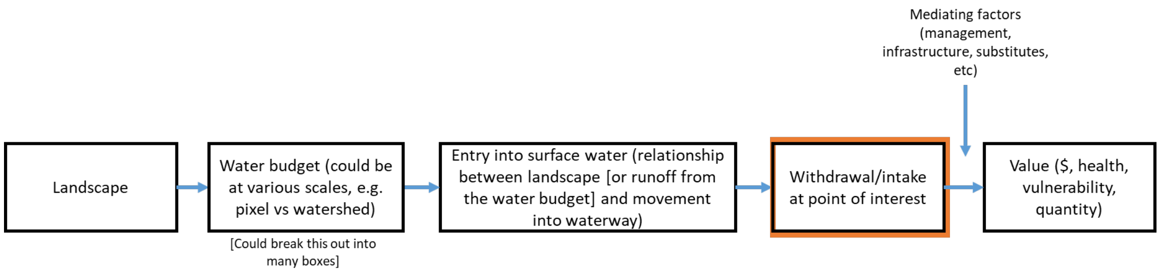
Food, fiber and fuel production:



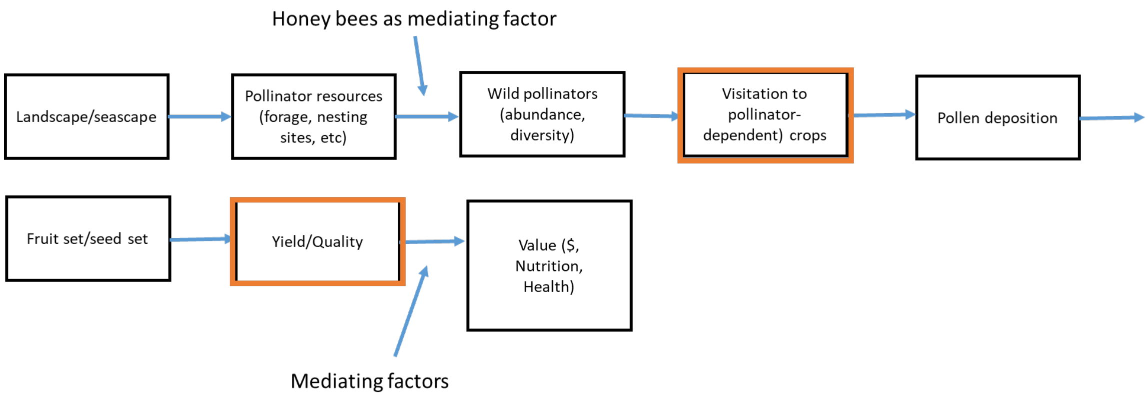
Forage production:



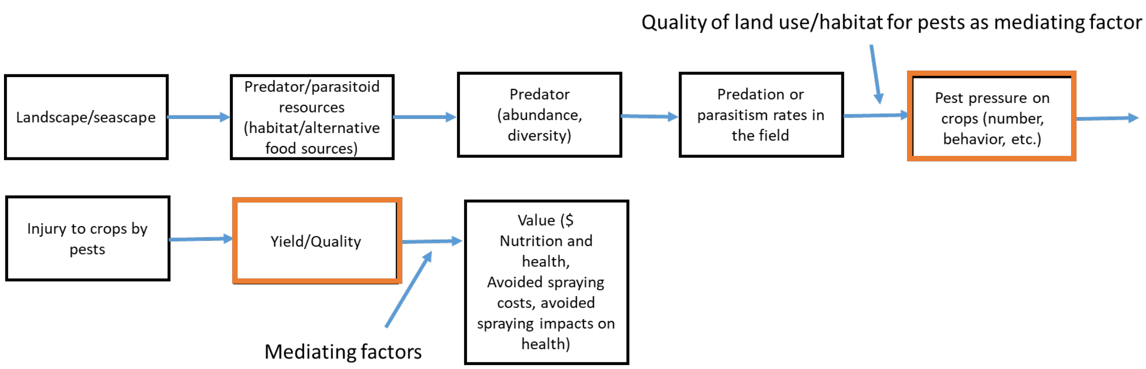
Freshwater provision:



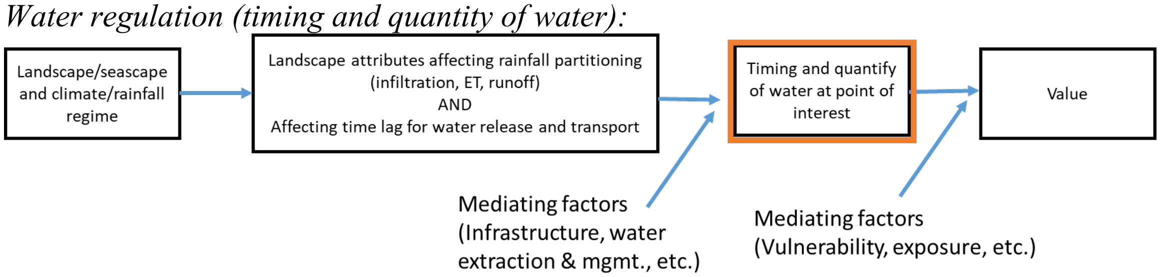
Crop pollination:



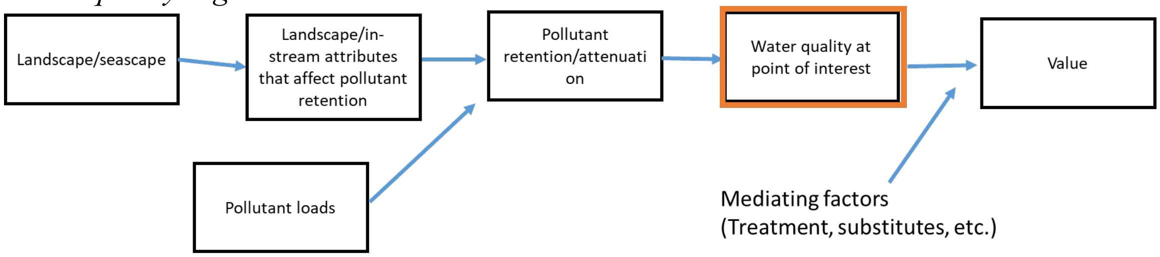
Pest control:



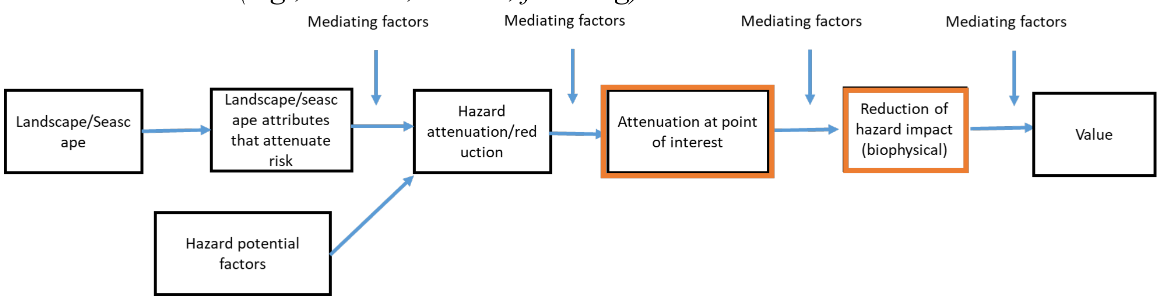
Water regulation (timing and quantity of water):



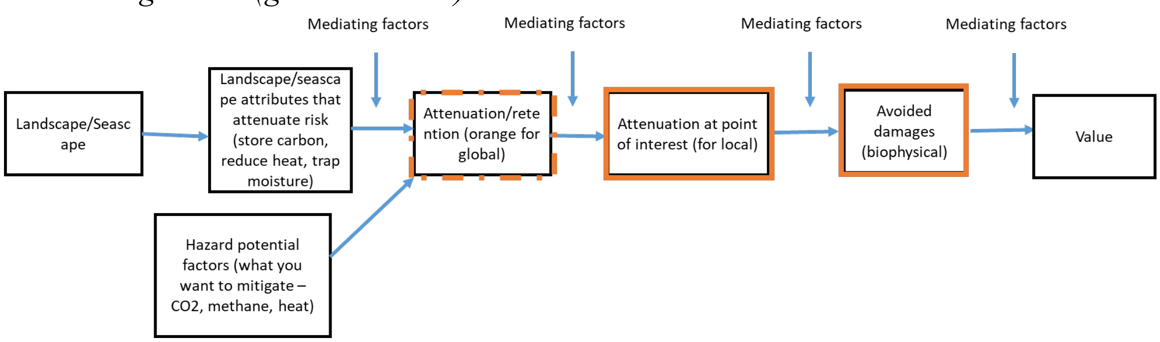
Water quality regulation:



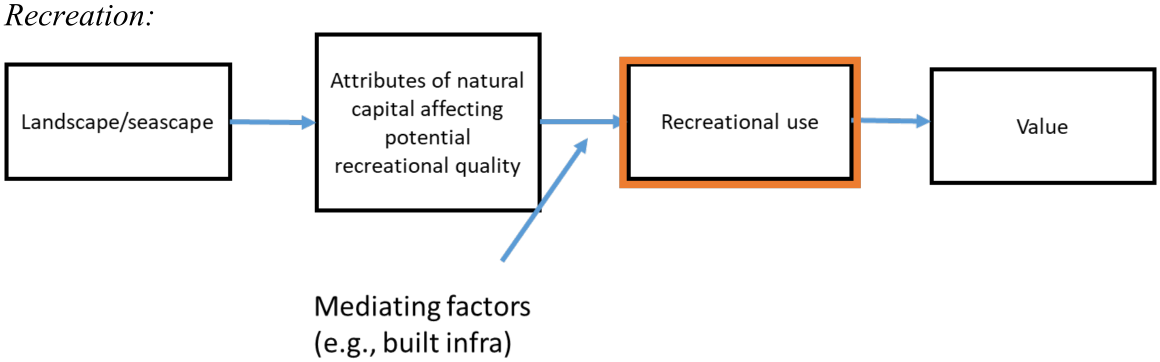
Hazard reduction (e.g., erosion, storms, flooding):



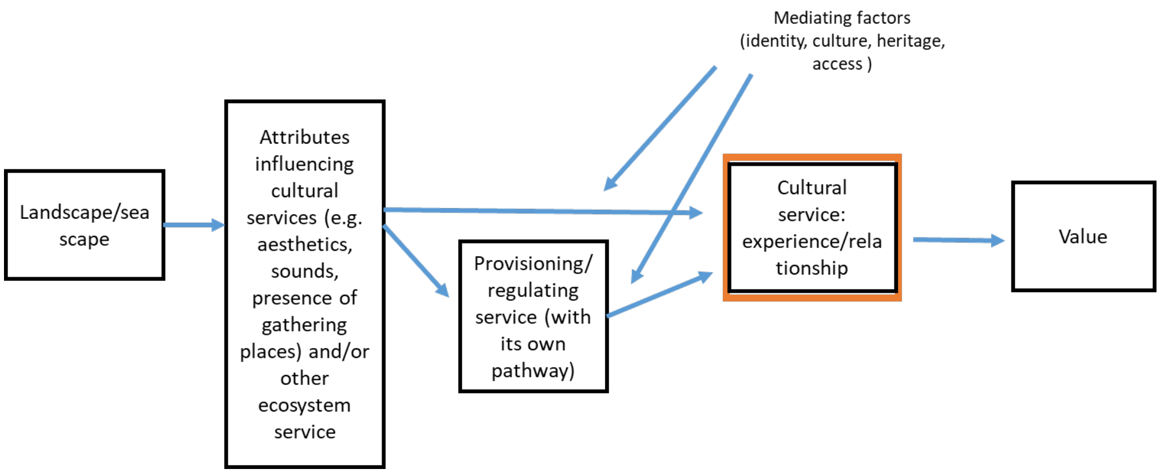
Climate regulation (global or local):



Recreation:



General cultural services:



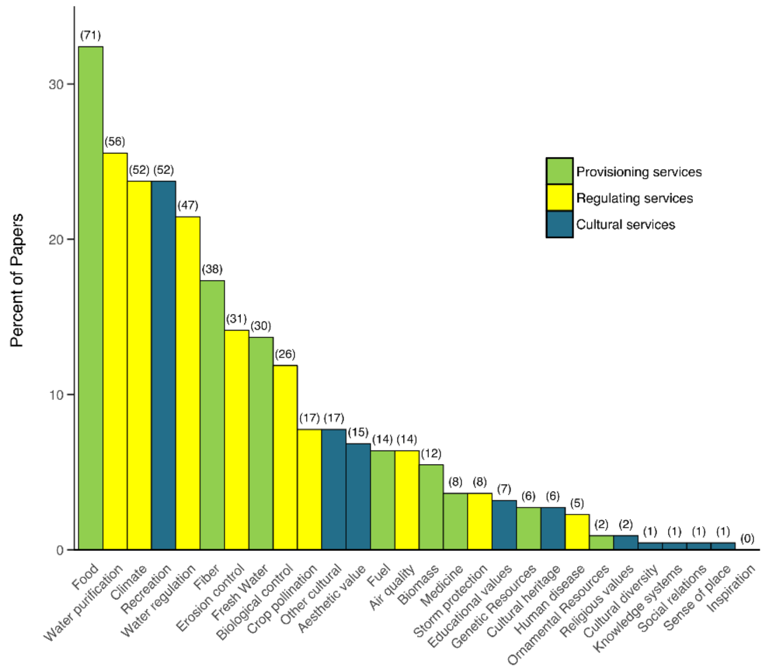


Figure S1: Allocation of research effort across provisioning, regulating and cultural ecosystem service categories among the 223 papers assessing a specific provisioning, regulating or cultural ecosystem service. The number in parentheses on top of each bar is the number of papers assessing each service. The percentages sum to more than 100, as a single paper may assess multiple services. Nearly half (47%, 261/551) of these assessments were regulating services, one-third (34%, 186/551) were provisioning services, and the remaining 19% were cultural services (104/551).

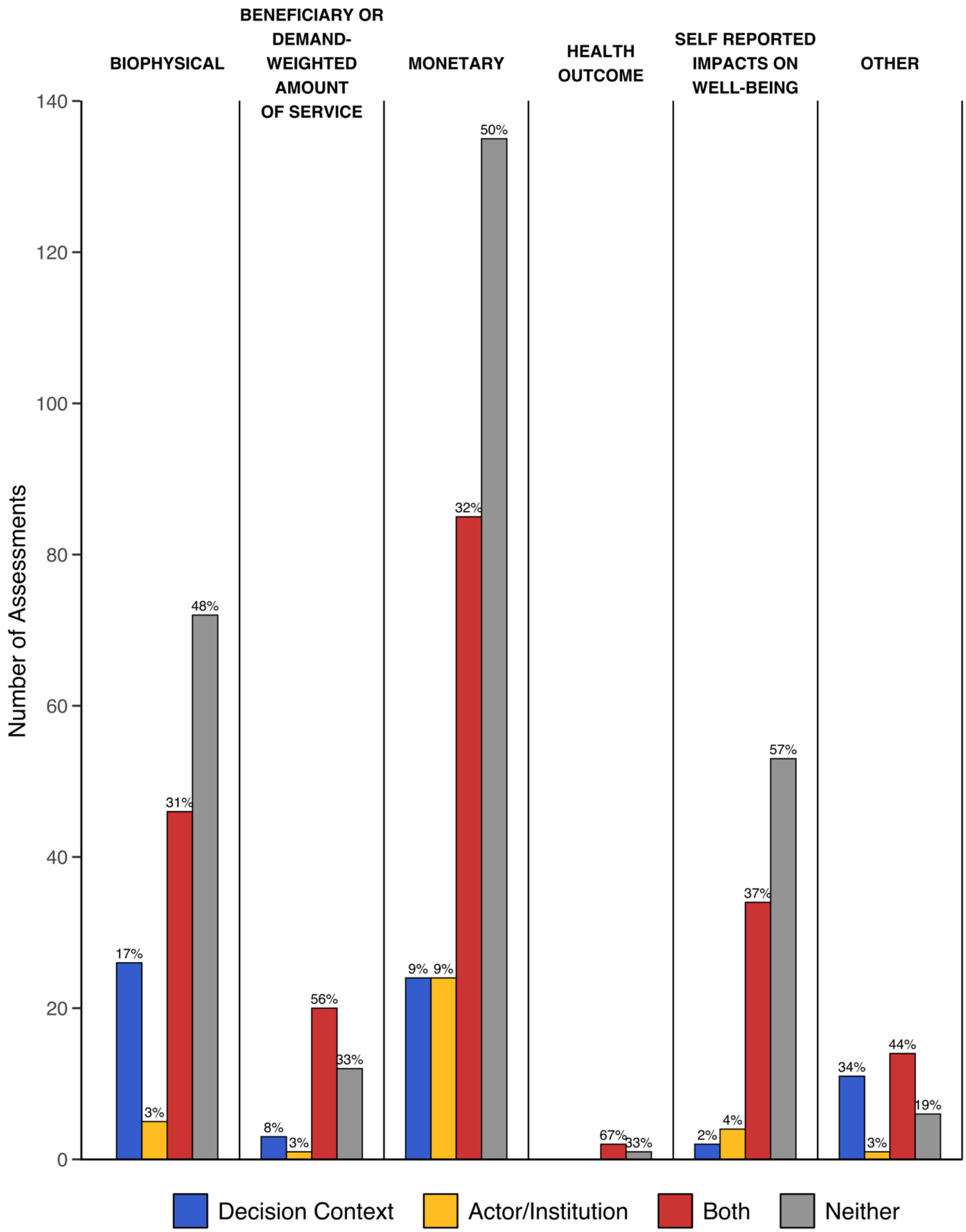


Figure S2: Connection of assessments that measured service to a decision context and/or actor/institution (see Figure 1), by the final metric of ecosystem service benefit (supply, service delivery, or value, depending on chain type). Percentages sum to 100% within final metric categories.

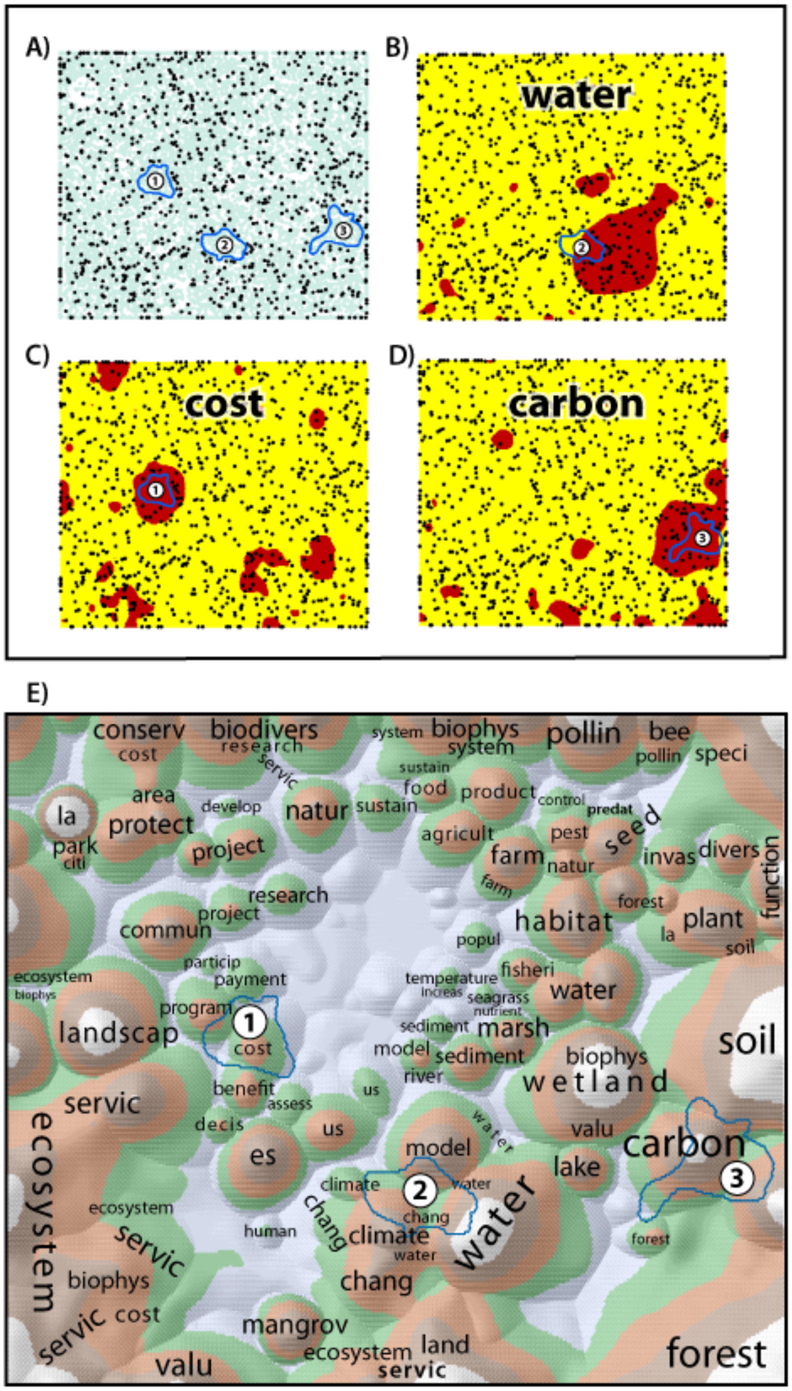


Figure S3. Self-organizing map representation of a trained artificial neural network consisting of a corpus of more than 12,000 documents. A) Black dot symbols indicate the location of our random sample in 2D space. B-D) Hotspots (in red) for the terms “water”, “cost”, and “carbon” denote areas where the 90th percentile of distribution of values within each dimension is exceeded. These three terms were chosen because they dominate the largest gaps of our sample (blue polygons, labeled 1-3). The black dot symbols that overlap red areas confirm that abstracts in our sample include these terms. E) Term dominance landscape where the size of each label indicates cluster strength.