Meta-analysis reveals that the provision of multiple ecosystem services requires a diversity of land covers

Carla Gómez-Creutzberg[a,1], **Malgorzata Lagisz**[b], **Shinichi Nakagawa**[b], **Eckehard G. Brockerhoff**[c,d] and **Jason M. Tylianakis**[a,e,1]

[a] Centre for Integrative Ecology, University of Canterbury, Christchurch 8140, New Zealand; [b] Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia; [c] Scion (New Zealand Forest Research Institute), Christchurch 8440, New Zealand; [d] Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland (Present address); [e] Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, Berkshire SL5 7PY, United Kingdom.

[1] To whom correspondence may be addressed. Email: cgomezcre@gmail.com or jason.tylianakis@canterbury.ac.nz

Author contributions

- Conception and development of original project concept - JMT, EGB, CGC
- Securing funding for the project - JMT, EGB
- Study design - CGC, SN, ML, EGB, JMT
- Data acquisition - CGC, with input from JMT and EGB
- Data analysis - CGC, JMT, SN, ML
- Drafting of manuscript - CGC
- Revision and critical appraisal of the manuscript: JMT, SN, ML, EGB

Classifications: Biological Sciences, Social Sciences

Abstract

Land-use change creates acute trade-offs among ecosystem services that support wellbeing. We comprehensively assess trade-offs and synergies in ecosystem service provisioning across land covers. We systematically surveyed published literature (1970 – 2015) for New Zealand, to quantify 1137 individual land cover - ecosystem services relationships for 473 service provision indicators across 17 services. For each service, we used a network meta-analysis to obtain quantitative estimates of provision from each land cover. Multivariate analyses of these estimates allowed us to compare 1) land covers in the provision of multiple services, and 2) services in terms of the different land covers that provide them. We found a significant trade-off in service provision between land covers with a high production intensity against those with extensive or no production; the former providing only a limited range of services. However, our results also indicate that optimal provision of multiple services is unlikely to be met by a single land cover, but requires a combination of multiple land covers in the landscape. When applied to land-use/land-cover planning, our approach reveals: 1) land covers that cluster together, and thus provide redundancy (and potentially resilience) in service production, and 2) land covers that are likely to exhibit complementary roles in service provision because they occur at opposite extremes of the multivariate service space. It also allows identification of service bundles that respond similarly to land cover. Actively incorporating findings from different disciplines into ecosystem services research can guide practitioners in shaping land systems that sustainably support human welfare.

Keywords: Ecosystem services, land-use planning, trade-off, network meta-analysis, land cover

Introduction

Human transformation of the Earth's surface through land-use activities has reached an unprecedented magnitude, and constitutes a major driver of global environmental change (1–3). Humans rely on resources appropriated through land use, however most of these practices affect the earth´s ecosystems in ways that undermine human welfare (1).

Continued population growth (4) and growing per capita consumption of resources (5) make it critical to reconcile production and sustainability in land systems.

Ecosystem services offer a framework for addressing these complex issues in land-use planning and management by linking human welfare with ecosystems. They explicitly account for the benefits that ecosystems bring to society, and prioritize long-term welfare over immediate economic reward (6–9). This perspective encourages strategies that optimize service provision across different land uses or enhance the provision of multiple services within a single type of land use (10). To this end, important efforts have been made to map and quantify services and estimate their economic value (11–14) or to examine synergies and trade-offs in their spatial occurrence (15–21).

Despite these advances, our understanding of service trade-offs and synergies has traditionally been impaired by the lack of, and costliness of obtaining, spatial data on multiple ecosystem services from multiple land uses across landscapes. Recent efforts have addressed these problems, for example, by using satellite earth observation to assess service supply (22) or by examining ecosystem service bundles and their spatial distribution in relation to socio-ecological systems (23–26). However, these approaches are still limited by data availability, which generally constrains the assessment of each service to measurements of individual indicators that are often location-specific. Furthermore, in many cases these measurements may not necessarily convey information that can be readily used by decision-makers or the public, because, for example, they do not take full account of the differences in the spatial and temporal scales at which ecological and social systems operate (27) or their link with human welfare issues is not stated or communicated effectively (9, 28, 29). Thus, consistent generalizations regarding trade-offs among broad categories of land use (e.g., natural vs. production systems) or services (e.g., those with local vs. global beneficiaries) remain elusive.

Data limitations also constrain the use of monetary valuation of ecosystem services as a means to assess synergies and trade-offs. Monetary values are particularly difficult to define for services that provide non-market benefits including those related to fair distribution and sustainability (28). Moreover, monetary values are not necessarily relevant to all decisions (30, 31) and can have limited extrapolation to other locations (32) or be highly contested among certain groups (33, 34). This necessitates the development of alternative methods to provide decision makers with evidence that is well-suited to their needs, is based on readily-available data and can be extrapolated (9).

Here we present an approach to inform land-use decisions by assessing trade-offs and synergies in the provision of multiple ecosystem services across land covers (as a proxy for land use). We use New Zealand as a case study because the high levels of endemic flora and fauna and relatively recent introduction of large-scale intensive agriculture make conservation-production tensions particularly acute, and necessitate conservation strategies that go beyond protected areas (35). We derive quantitative evidence of land cover effects on ecosystem service provision from existing literature through a meta-analysis (SI Appendix 7 and SI Dataset 1). Unlike existing reviews and meta-analyses on ecosystem services (36–42), our work does not collate existing ecosystem service assessments. Rather, we synthesize primary biophysical research that compares land covers in relation to a large variety of measures that indicate the provision of a service, regardless of whether service terminology was used.

We use this comprehensive evidence base to detect land covers that may operate as 'generalists' (i.e. providing many services) or 'specialists' (i.e. providing just a few services). This allows us to determine whether landscapes require multiple land covers to provide a comprehensive suite of non-production services, or whether a single land cover (e.g., a natural habitat) can provide the majority. We also identify ecosystem service bundles (i.e. services that are provided by the same land cover type), and potential trade-offs among services that are best provided by different land covers. Subsequently, we test whether there is a systematic difference between exotic-species-dominated production and native non-production systems in their provisioning of service bundles. Due to potential tensions between service beneficiaries and providers (43, 44), we also test whether services with localized (e.g., soil formation) vs. global (e.g., climate regulation) benefits tend to flow from different land covers. If they do, then this could exacerbate the disconnect between beneficiaries and those who enable, maintain and restrict services (45). Finally, we present an example of how this information can be used to readily examine the effects of land use or land cover trajectories or contrasting decisions on landscape-scale management of ecosystem service trade-offs.

Methods

Our systematic review was structured according to the "Guidelines for Systematic Review in Environmental Management" developed by the Collaboration for Environmental Evidence (46). We searched the literature for quantitative comparisons of two or more land uses in the provision of one or more ecosystem services within New Zealand. Our ecosystem service definitions were adapted from the Millennium Ecosystem Assessment (47), with a total of 35 services spanning across the provisioning, regulating, cultural and supporting categories (SI Appendix 1). Land uses, formally defined as the purposes to which humans put land into use (48), were captured in our research as land covers (SI Appendix 1), since these include units that are not directly used by humans and, consequently, correspond more closely with the actual experimental or sampling units of many of the documents in our search.

Data collection, aggregation and calculation of effect sizes

Full details of the search and screening process are described in SI Appendix 2; here we present a brief outline. We searched the Scopus database for titles, abstracts and keywords with at least one match in each of the 3 components that structured our search: 1) "New Zealand", 2) land cover and land use terms and 3) ecosystem service terms (see SI Appendix 3 for the full search phrase). Land cover terms included all possible variations of "land use" and "land cover" as well as terms on specific land use and land covers (both generic and specific to New Zealand). The ecosystem services component drew upon the names of each service (and possible variations of these) but also included vocabulary describing processes and conditions that could reflect their provision at the site scale akin to individual land cover units. The search was finalized in December 2014, and was constrained to include documents published from 1970 onward, to be comparable with current land use regimes in New Zealand (49).

Our keyword search yielded 9,741 references, and an initial automated screening reduced these to 4,373 publications by removing references that only mentioned a single type of land cover or land use in their title, abstract and keywords. Publications with 2 or more land cover terms were scanned using Abstrackr, an interactive machine learning system for semi-automated abstract screening, often used in medical meta-analyses (50). By learning from the abstracts or words a user identifies as relevant during the screening process, Abstrackr can predict the likely relevance of unscreened abstracts and effectively assist in the exclusion of irrelevant ones (more details in SI Appendix 2).

Abstract screening yielded 914 relevant papers, which were passed on to a team of four reviewers for full-text assessment and data extraction. Studies that did not have replicated observations (as defined in SI Appendix 2) for any land covers were discarded, whereas studies that contained replication on some, but not all, of the land covers were kept and only data on the replicated land covers were extracted. Although we only included terrestrial land covers, ecosystem services provided by land but linked to a water body were included in our analysis. Full details of how the full-text selection criteria were applied can be found in SI Appendix 4. In total, we extracted data from 136 studies that passed all inclusion criteria (see SI Appendix 5 for bibliographic details of each study).

Information on the land covers, quantitative measures of ecosystem service provision, experimental design and bibliographic details for each study was collated in a database. To allow for comparability across studies, individual land covers described in each study were matched to the nearest category in New Zealand's Land Cover Database - LCDB (51). This classification system includes forest, shrubland and grassland areas of either predominantly native or exotic vegetation, as well as cropland and more artificial surfaces such as built-up surfaces and mining areas (SI Appendix 1).

Often, the same quantitative measure of service provision obtained from a study (indicators, presented in SI Dataset 1) would be relevant to more than one ecosystem service. We therefore assigned each indicator to one or more ecosystem services and defined the general direction of each indicator - ecosystem service relationship (i.e. by determining whether larger values of the indicator would generally reflect an increase or decrease in service provision). This was done because the majority of the studies in our meta-analysis did not explicitly use 'ecosystem services' terminology. Instead, they measured environmental or ecological variables that could be used as indicators of ecosystem service provision, provided a conceptual link was defined between the indicator (e.g., annual water discharge of a catchment) and the corresponding service (provision of freshwater). Thematic experts (see Acknowledgements) were consulted for assigning indicators to services and determining the direction of the indicator - ecosystem service relationship when we could not readily define this. Although we recognize that the relationship between an indicator and a service may be non-linear (e.g., pollination services may saturate with large numbers of pollinators), in most cases it was not possible to establish a clearly defined non-linear function, so we assumed a linear relationship for all indicators.

Unique identifiers allowed us to define individual studies, regardless of whether they were published within a publication that spanned more than one study or across different publications (SI Appendix 2). Multiple measures from within the same replicate site were aggregated into a single value per replicate (see SI Appendix 2 for details). Methods for standardizing measures of variance are provided in SI Appendix 6.

We obtained a final database with information on 473 ecosystem service indicators among 3105 pairwise comparisons of two land covers from 136 studies. A log Response Ratio was used as the effect measure for comparing pairs of land covers within each study and was standardized such that larger values always represented greater service provision in the numerator land cover relative to the denominator one (see SI Appendix 2 for this standardization and log Response Ratio variance calculations).

Studies with more than one indicator of a given service were aggregated to have the same weight as studies with only a single indicator (see SI Appendix 2). Subsequently, the total number of land cover comparisons in our final dataset of 136 studies was reduced from 3105 to 1003 comparisons for individual services within single studies (See SI Dataset 2 for an overview of the final data).

Data analysis

Data analysis was conducted as a two stage process: we first examined the provision of each ecosystem service by different land covers, and then assessed the relationships among land covers in terms of multiple ecosystem services. For the first stage, we conducted a separate network meta-analysis (52) for each ecosystem service. Network meta-analysis allowed us to compare, for each ecosystem service, a wide array of land covers across different studies, even though we did not have data for direct comparisons among all combinations of land covers. Conventional meta-analysis compares 2 treatments at a time, using direct comparisons from each study. In contrast, a network meta-analysis can compare multiple (i.e. 3 or more) treatments simultaneously by using both direct evidence (studies comparing pairs of treatments) and indirect evidence derived from linking common treatments across different studies in a network of evidence (52). For example, if some studies show that land cover A is better than B in the provision of a service, and others provide direct evidence that B is better than C, then a network meta-analysis allows us to make the inference that A will also be better than C.

We conducted our network meta-analyses with the R package *Netmeta* (53), which offered a frequentist approach to calculate point estimates (and their corresponding 95% confidence intervals) of the effect of the different land covers on the provision of each ecosystem service. We used a random effects meta-analytic model to generate these estimates and their confidence intervals, both of which were then used to define probability scores (54) and examine how different land covers ranked in the provision of each ecosystem service. We used both rankings and point estimates to construct forest plots comparing all land covers to the high producing exotic grassland (which we defined as a baseline reference) in the provision of each service. We selected high producing exotic grasslands as the reference because it was the only land cover that was represented across all ecosystem services in our dataset.

Trade-offs and synergies in land cover effects across the whole suite of ecosystem services were then examined using hierarchical clustering of the network meta-analytic estimates. For this, we constructed a land cover by ecosystem services matrix (found in SI Appendix

7) using the estimated log Response Ratios of each land cover (relative to the high producing exotic grassland reference) in each service, as determined with the individual network meta-analyses (see forest plots in SI Appendix 10 for estimated ratios). Missing values in this matrix resulted from sets of land covers for which we had no information on a given service or could not infer the corresponding ratios.

For analysis, we selected the largest possible subsets of this matrix with no gaps. This resulted in two data subsets: a matrix of ten land covers by eight ecosystem services and another matrix with nine ecosystem services by eight land covers. The matrix with ten land covers was used to compare land covers in their provision of the eight services. This allowed us to explore how land-cover differences influence suites of services. The matrix with nine services was rotated to have services as rows (land covers as columns) and used to compare services in terms of the land covers that provide them. This allowed us to identify services that tended to be provided by different land covers and thus would likely be traded off with one another in land decisions. A dissimilarity matrix was then calculated from each of these matrices using the *daisy* function of the *cluster* package for R (55) with Euclidean distances. For the matrix with ten land covers, distances were based on land cover observations for each service, while for the rotated matrix with nine distances were based on service observations for each land cover. Each of the distance matrices was then used to perform hierarchical clustering (56) to identify groups of land covers and services exhibiting similar behavior.

Finally, we tested hypotheses on whether characteristics of the land covers and ecosystem services in our distance matrices explained the trends observed in each of the corresponding clusterings. Land covers were grouped under two categorical variables, one denoting the presence/absence of forest cover and another separating production land covers, dominated by exotic vegetation cover, from those with no production activities. Originally, we intended to compare land covers with a native vs. exotic vegetation cover separately from production vs. natural. However, we omitted the former category because, except for one, all land covers with exotic vegetation were production and all native covers had little or no production. We used a permutational multivariate analysis of variance

(PERMANOVA) to test whether these variables or their interaction explained between-land-cover differences in the provision of multiple ecosystem services.

Similarly, ecosystem services were classified into three categories according to the scale at which their benefits were perceived (locally within each land cover, regionally across neighboring land covers, or globally) and into three categories representing the biophysical domain to which the majority of indicators used to quantify the provision of each the service belonged to (biotic, hydrologic or edaphic). A PERMANOVA was applied to test whether scale or domain explained groupings of ecosystem services within the multidimensional space of service provisioning across land covers. We did not test for an interaction between these two variables because some combinations of factor levels were absent and we lacked sufficient degrees of freedom.

PERMANOVA analyses were conducted using the *adonis* function of the *vegan* package in R (57). Since variables are added sequentially in adonis, to be conservative we performed each PERMANOVA twice and swapped the order of the variables in the second iteration so that each variable was tested second, after controlling for any collinearity with the other predictor (i.e. adjusted sums of squares). The *betadisper* function of the *vegan* package was used to test the assumption of multivariate homogeneity of group dispersions, and all tests met this assumption. SI Appendix 8 presents the land covers and ecosystem service categories used in these analyses.

Results

Data coverage

The 136 studies in our database contributed data on 17 different ecosystem services and 25 land covers (SI Appendix 9). All four categories of ecosystem services (47) were represented within our dataset. However, most studies examined regulating or supporting services, with 116 and 115 studies, respectively. Only 47 studies presented data on provisioning services and five on cultural ones. All of the services in the supporting category (habitat provision, nutrient cycling, soil formation, water cycling and primary production) are represented in our database. Only four land cover comparisons had more than 20 studies (high producing exotic grassland vs. exotic forest, indigenous forest vs. high producing exotic grassland, short-rotation cropland vs. high producing exotic grassland and exotic forest vs. indigenous forest); whereas the remaining land cover pairs were represented by 10 or fewer studies each.

Land cover effects on individual ecosystem services

SI Appendix 1 presents an overview of the evidence network and individual network meta-analyses for each of the 17 ecosystem services in our database. The results of the meta-analyses are expressed as forest plots (SI Appendix 10). For several services, the narrow confidence intervals in these forest plots reveal that land covers with native shrub, grass and forest vegetation (i.e. broadleaved indigenous hardwoods, indigenous forest, manuka/kanuka, matagouri or grey scrub and, in many cases, tall tussock grassland) tended to rank higher in their provision than the more intensive high-value production land covers (particularly short-rotation cropland and high-producing exotic grassland). Regulation of water timing and flows, water purification, freshwater provision and disease mitigation conformed to this general pattern. In these services, low producing grasslands (which comprise a mix of exotic and native vegetation) and exotic forests also perform relatively well and always rank within the top half of all land covers.

For habitat provision the difference between native vegetation and production systems was less important than the presence of open vs. woody vegetation cover. For this service, land covers with woody vegetation (sub-alpine shrubland, matagouri or grey scrub, exotic forest, broadleaved indigenous hardwoods and forest) ranked higher in their estimates of service provision than those with open covers (short-rotation cropland, depleted, tussock, low and high producing grasslands) or deciduous hardwoods. Meanwhile, primary production tended to be highest under production systems (e.g., exotic forest, cropland and high-intensity grassland) and lower in natural systems (e.g., low producing and tussock grassland, indigenous forest), rather than differing between forested and open covers; however, these trends were not statistically significant due to the wide and overlapping confidence intervals. Importantly, these results indicate that no single land cover provides all services at a maximal level. Rather, in order to ensure flows of multiple services, multiple land covers will be required within the landscape.

The forest plots in SI Appendix 10 for primary production, erosion control, pest regulation, waste treatment, capture fisheries, ethical & spiritual values, pollination and regional & local climate regulation all present wide, overlapping confidence intervals for all or most of their estimates. This suggests non-significant differences in land-cover provision of these services. For some services, this could be due to small evidence bases, either in terms of few studies or few comparisons for specific land cover pairs within a network (which can be seen as large differences in link weights in the corresponding evidence networks, SI Appendix 10). In the case of erosion control, where the evidence base is formed by 22 studies (SI Appendix 10), overlapping confidence intervals in the land covers with the greatest number of comparisons express high variability in service provision from these land covers and suggest that other factors besides land cover (e.g., slope, soil type) likely account for the differences in erosion control across the sites in all 22 studies.

Land cover effects across multiple ecosystem services

To explore how the above trends in the provision of individual services translate into trade-offs and synergies among suites of services and the land covers providing them, we conducted multivariate analyses on service provision across land covers. First, we examined whether groups of land covers played a similar role in the average provision of services and, conversely, whether groups of services responded similarly to differences in land cover.

Differences among land covers in their provision of services

When we focused on the subset of data with values for the greatest number of land covers across the maximum number of ecosystem services, we observed a gradient of land covers that separates those with lower production (and, generally, forest cover) from the high value production systems (Fig. 1). More specifically we can identify the following clusters:

- Cluster A - Fruit and vegetable production systems

- Cluster B - Intensive production systems: exotic forests (harvested and unharvested) and high production grasslands
- Cluster C - Indigenous forests (well-established or in advanced succession) and low production grasslands
- Cluster D - Tall tussock grasslands and deciduous hardwoods

The gradient from clusters A to C is consistent with the aforementioned trend of native land covers performing better in the services for which production land covers perform poorly and vice versa (see SI Appendix 10). These clusters provide an approach to identify the strongest trade-offs in service provision, such as that between the land covers in cluster C and those in clusters A and D (Fig. 1) with the latter specializing in biomass production, the formation of soil suitable for plant growth and fast water cycling rates, and the former being better suited for providing habitat, cycling nutrients and purifying, providing and regulating the flow of water. Larger differences in the height at which clusters separate from each other indicate greater differences in service provision. Consequently, in Fig. 1, clusters B and C (both with grasslands and forests) are more similar to each other in their provision of services than they are to clusters A or D. Likewise, B and C are more similar to each other (indicated by the lower branch point) than A and D are to each other.

The trade-off in service provision between production and non-production land covers was statistically significant (PERMANOVA, Pseudo $F_{1,6} = 2.927$, partial $R^2 = 0.259$, $p < 0.01$; see detailed results in SI Appendix 11). The assumption of homogeneous dispersion between both groups was met ($F_{1,8} = 0.15$, $p > 0.05$), suggesting that neither provides a greater range of ecosystem services among its different land covers. Conversely, the separation between forested and non-forested land covers did not significantly explain the distribution of land covers in service space (Pseudo $F_{1,6} = 1.226$, partial $R^2 = 0.109$, $p > 0.05$; see also SI Appendix 11) nor did the interaction between forested/non-forested and production/non-production (Pseudo $F_{1,6} = 1.141$, partial $R^2 = 0.101$, $p > 0.05$; SI Appendix 11).

Differences among ecosystem services in the land covers that provide them

By clustering ecosystem services based on their provisioning in each land cover, we identified some services that tend to perform differently from each other and, consequently, have their provision traded-off across land covers. This trade-off is acute for water-related services; most of these tend to occupy distinct spaces within the dendrogram with water cycling standing apart from all other services, water purification and freshwater provision in a separate cluster, and regulation of water timing and flows in a single branch close to global climate regulation and nutrient cycling (Fig. 2). The trade-off between water cycling and regulation of water timing and flows is likely because land covers that allow increased runoff and present low water retention (such as harvested forests, croplands and built-up areas) deliver more of the water cycling service than the land covers that promote soil water storage and, consequently, perform better under the regulation of water timing and flows service (e.g., broadleaved indigenous hardwoods, indigenous forests and low producing grasslands). Freshwater provision and water purification form a cluster because the water quality aspect of their provision was assessed with common indicators (SI Dataset 1) and in both cases greater service provision came from land covers contributing to enhanced water quality (such as tall tussock grassland and indigenous forest, SI Appendix 10).

In contrast to the water-related services, those more closely linked to the soil system (nutrient cycling and soil regulation) are found closer to each other in Fig. 2, and appear to be delivered similarly across land covers (see forest plots in SI Appendix 10). In our analysis, global climate regulation falls under this broad group of services and forms a tight cluster with nutrient cycling (Fig. 2). This is likely due to the indicators shared by both and a gap in our database with respect to the contribution of vegetation and livestock in greenhouse gas fluxes. Further research on these aspects should therefore allow for a more comprehensive quantification of the provision of this service across land covers and uses.

Neither the spatial scale of service provision (local, regional, global) nor the main biophysical domain of services (edaphic, hydrologic and vegetation) significantly affect the distribution of services in multidimensional 'land cover space'. When tested together as additive main effects in a model, we observed changes in the importance of each variable that depended on their order in the model (SI Appendix 11), suggesting collinearity between the two predictors. When each was tested after removing the partial effect of the other (i.e. as second in the model), neither the biophysical domain of services

(PERMANOVA, Pseudo $F_{2,4}$ = 2.253, partial R^2= 0.312, p > 0.05; see also SI Appendix 11) or spatial scale of service provision (PERMANOVA, Pseudo $F_{2,4}$ = 2.337, partial R^2= 0.323, p > 0.05; SI Appendix 11) independently explained significant variation in the distribution of services in multidimensional space. In Fig. 2, the only services of similar scale that cluster together are freshwater provision and water purification, which deliver benefits at regional scales and, as mentioned before, share some of their indicators and responses to these indicators. The two other services which form a tight cluster in Fig. 2 (nutrient cycling and global climate regulation) also have common indicators (pertaining to microbial respiration and carbon dioxide fluxes in the soil) to which they respond similarly, despite their benefits being experienced at different scales.

Discussion

Our meta-analysis revealed a clear difference between high - value production land covers, which specialize mainly in services relating to primary productivity, and all the land covers with native shrub or forest vegetation. Together, land covers with native vegetation outperformed production ones in the provision of most supporting and regulating services. However, in New Zealand production land covers are dominant, with exotic forests, high producing exotic grasslands, croplands, and orchards/vineyards occupying 42% of the country´s terrestrial area in 2012 (58). Ecosystem service assessments were conceived, in part, to make explicit how decisions on ecosystem management reflect preferences for different, competing sets of services (7, 59). The trade-off we find between production and native land covers illustrates how the provision of services with a high market value and short-term returns occurs at the expense of services that have a non-market value but are essential for sustained, long-term human welfare (60).

The above findings resonate with the recommendations of Foley and colleagues (61) with respect to halting indiscriminate expansion of agriculture into sensitive ecosystems. However, our findings also suggest that, at the landscape scale, the dichotomy between production and non-production land covers is not solved with a single ‘generalist’ native land cover. Even for the services that were best delivered by land covers with native vegetation, we did not find evidence of a single land cover consistently performing better than the rest in providing all of the services. Instead, our findings show that while indigenous forests tend to perform well in the provision of most services (particularly erosion control, waste treatment, disease mitigation and global, regional and local climate regulation), in many services they are outperformed by other land covers such as native shrubland (manuka and or kanuka, which contribute very well to soil formation and regulation of water timing and flows), tall tussock grasslands (which are well suited to freshwater provision, water purification and pest regulation) and even advanced successional forest (broadleaved indigenous hardwoods, which rank high in regulation of water timing and flows, nutrient cycling and habitat provision). Therefore, a landscape with a mosaic of these land covers is more likely to offer a broader suite of services than one dominated by large extents of any single native land cover (62–64).

Thus, we support earlier recommendations to extend beyond the dichotomy of conservation vs. production land into a more a comprehensive management (65–67). Such management could, for example, contemplate the extension or restoration of under-represented native land uses at strategic sites where intensive use is not matched by increased production yield, to promote provision of critical services or broaden the existing suite. To this end, management will need to be informed by a comprehensive understanding of how services can scale up from individual land use units and how the relative sizes of different land use units within a landscape can affect service provision.

Our analysis shows that low intensity-production land covers that retain some native vegetation (i.e. the low producing grasslands in our dataset) can approach native land covers (broadleaved indigenous hardwoods and indigenous forests) in terms of overall service provision. These low-intensity production land covers demonstrate that production and a suite of other services can be jointly delivered, providing empirical support to the notion of managed ecosystems with "restored" services proposed by Foley et al. (1). Importantly, we identified great variability in the effect of some land covers on the provision of certain services, despite there being high replication in our evidence base for these effects (e.g., water purification by short-rotation croplands and erosion control by high producing exotic grasslands, indigenous and exotic forests). This suggests that local environmental conditions and management practices can significantly alter how a given land use affects service provision (68), which implies some potential to improve service provision by adjusting management practices of specific land uses (69–71). Within individual land uses, decisions on which practices to adopt will require detailed research on the effects of different management regimes on service provision, as well as an understanding of the extent to which the plasticity in service provision is constrained (or favored) by environmental factors.

A critical challenge in applying the ecosystem services framework to spatial and environmental planning is understanding the extent to which different land uses affect the provision of services (72). The uneven coverage of different services that we observed in the literature reflects both the variable ease of quantifying different services and the likely relevance of comparing provision of a given service among land uses. Within our dataset, supporting and regulating services are best represented. In the global literature, regulating services are also the most commonly quantified and mapped category, however, they are usually followed by provisioning services while the evidence on supporting services is scarce (11, 13, 14, 36, 38, 73). The limited representation of provisioning services in our dataset possibly occurred because most provisioning services (e.g., milk, timber) are linked to single or few land covers and, consequently, are unlikely to be compared across land covers. Such services, however, enter the market directly and can be more readily quantified in monetary terms. In contrast, the supporting and regulating services that predominate in our dataset usually translate to externalities in the context of production systems and are likely more readily quantified through biophysical indicators than monetary units (36, 74).

Cultural services are poorly represented in our database, with the few indicators for this category all being shared with the capture fisheries provisioning service, because they pertain to eels, which are of cultural significance to Māori in New Zealand. As has been argued elsewhere (75), cultural ecosystem services have non-material and ideological dimensions that are not readily quantified and, thus, are not well represented even within the emerging body of specialized literature on ecosystem service provision assessment (76). Moreover, it has been suggested that cultural services escape the instrumental value domain present in the ecosystem services framework. Instead, they fall under the relational domain, whereby value is not solely defined in terms of the direct benefits we can derive from an ecosystem, but also in terms of the social webs of desired and actual relationships we construct around that ecosystem or its components (33). Consequently, for these services, a quantitative approach like ours should be complemented with assessments that address the relational dimensions of the values people hold for the natural elements in different land uses to better represent their importance in a cultural context (77).

Individual services are defined to encompass distinct processes and values, but these are often quantified by overlapping sets of indicators (74). For example, in our dataset indicators of water and soil pertained to more than one service (e.g., water purification and provision of freshwater both share indicators of water quality, while erosion control and soil formation share indicators on soil stability). Since the MEA was released, there have been initiatives to redefine services and their categories (78, 79); here we argue that future work in determining how to best quantify services and their spatio-temporal variation will be at least as important as refining their taxonomy. Furthermore, if a focus on quantifying ecosystem services should reveal aspects of services that are best left unquantified (such as the relational domain of cultural services), this could also lead to the development of alternative ways of assessing those services, which could then be applied in combination with quantitative approaches like the one we have developed here. Recent developments, like the concept of nature's contributions to people and the framework for their assessment proposed by Díaz and colleagues (80), provide an opportunity for reconciling these aspects.

Our work suggests that there is great potential in using existing data for assessing trade-offs in ecosystem service provision more cost-efficiently than through direct field observation and, moreover, that conceptual improvements in service quantification could greatly improve our ability to exploit this potential. An important caveat in our approach stems from underlying factors that are correlated with land cover and use and impact the provision of certain services. For example, steep slopes are frequently found in some land covers and land uses (like forestry and natural habitats) and would influence erosion control. Within our work the effect of these factors on service provision has not been separated from that of land cover. In fact, one could argue that land use is not selected independently from the local environment, so these factors are a component of any land use and its influence on services. Nevertheless, future approaches may benefit from examining how these factors affect the between- or within-land-use differences in service provision. This distinction would allow a shift from comparisons across locations (as we examined here) to the predicted impacts of land use change on services at any location. However, such predictions would need to incorporate legacy effects of past land uses, as these can have enduring consequences on ecosystem functioning (81, 82).

Likewise, we do not examine whether it is land covers or land uses that best capture the full range of benefits from each ecosystem service. Identifying whether there are any differences in the benefits captured by land cover over land use categorizations, may help inform the selection of either in future assessments. Moreover, some ecosystem service benefits may not be captured by neither land cover nor land use, but by other spatially variable factors (e.g., slope influences housing views which informs aesthetic values). These factors were beyond the scope of our work, however assessments that aim to capture the full range of ecosystem service benefits will likely need to include these factors in addition to land cover / land use.

Overall, we have presented a method for using existing data to assess trade-offs and synergies in service provision across land covers. This approach can facilitate the comparison of entire landscapes in the provision of multiple ecosystem services. Quantitative measures of how multiple land covers provide ecosystem services generated with our review (or land uses from equivalent exercises) can be used to map land covers or land uses into the multidimensional service space (Fig. 3). This mapping could reveal two key characteristics for land-use planning: 1) land covers/uses that cluster together, and thus exhibit redundancy (and potentially resilience) in service provision, or 2) land covers/uses that occur at opposite extremes of service space, and are therefore likely to exhibit complementary roles in their service provision (as services are traded off between them). In addition, the total hyper-volume occupied by all land covers/uses in this multidimensional service space (ordination plots in Fig. 3) can indicate the diversity of services provided by the full set of land covers/uses within a given landscape, which could be used in comparisons of existing landscapes or future scenarios.

As an example, the cases in Fig. 3 show how increasing the number of land covers within a landscape results in a corresponding increase in the diversity of services provided by that landscape. Case 3, with the greatest diversity of land covers, occupies the greatest hyper-volume in multidimensional service space. However, most of the land covers at the edge of this volume exhibit low redundancy, in contrast to the land covers in Case 2 that cluster around one portion of the ordination plot (and thus provide redundancy in the delivery of that set of services). Note that no landscape is likely to ever occupy a hyper-volume that extends through the entire service space, since some areas of this space may not correspond to any actual configuration of existing services. Therefore this approach is best applied for comparing existing and/or potential land cover (or land use) configurations against each other, rather than for assessing individual cases with no reference of the diversity of land covers that could actually be achieved in that landscape.

Similarly, mapping ecosystem services in multidimensional land-cover or land-use space (e.g., Fig. 2) allows the identification of bundles of services that respond similarly to land cover / land use. These bundles can then be used to identify management decisions that minimize disruption of service flows. Our approach opens the way for actively incorporating existing sources of information into ecosystem services research and informing practitioners to shape land systems that sustainably support human welfare.

Acknowledgements

We thank Melanie Hamzah for her assistance in the abstract screening and Sol Heber, Sophie Hunt, Jessica Furlong & Matthew Scott for their help with the full-text assessment and data extraction to collate our dataset. We also thank thematic experts Karen L. Adair (soil microbiology), Catherine M. Febria (freshwater ecology), Leo Condron (soil biogeochemistry), Matthew Turnbull (plant physiological ecology) & Angus McIntosh (freshwater ecology) for help with interpreting potential ecosystem service indicators. We are grateful to Daniel Stouffer and Melissa Ann Broussard for their valuable technical advice and members of the Stouffer and Tylianakis lab for comments and useful discussions. This work was funded by the Ministry of Business, Innovation and Employment, NZ programme BEST: Building biodiversity into an ecosystem service-based approach for resource management C09X1307. We thank Suzie Greenhalgh for coordinating this programme and providing helpful discussions.

References

1. Foley JA, et al. (2005) Global Consequences of Land Use. *Science* 8(July):570–574.

2. Steffen W, et al. (2015) Planetary boundaries: Guiding human development on a changing planet. *Science* 347(6223):1259855–1259855.

3. Turner B, Lambin EF, Reenberg A (2008) The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences* 105(7):20666–20671.

4. UNADES (2017) *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables* (United Nations).

5. Godfray HCJ, et al. (2010) Food Security: The Challenge of Feeding 9 Billion People. *Science* 327(February):812–818.

6. Costanza R, et al. (1997) The value of the world's ecosystem services and natural capital. *Nature* 387(6630):253–260.

7. Costanza R, et al. (2014) Changes in the global value of ecosystem services. *Global Environmental Change* 26:152–158.

8. Daily GC, Alexander S, Ehrlich P (1997) *Ecosystem services: benefits supplied to human societies by natural ecosystems* (Ecological Society of America, Washington DC) Available at: http://www.esa.org/esa/wp-content/uploads/2013/03/issue2.pdf.

9. Guerry AD, et al. (2015) Natural capital and ecosystem services informing decisions: From promise to practice. *Proceedings of the National Academy of Sciences* 112(24):7348–7355.

10. Lambin EF, Meyfroidt P (2011) Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences* 108(9):3465–3472.

11. Crossman ND, et al. (2013) A blueprint for mapping and modelling ecosystem services. *Ecosystem Services* 4:4–14.

12. Groot RS de, et al. (2012) Global estimates of the value of ecosystems and their services in monetary units. *Ecosystem Services* 1(1):50–61.

13. Egoh B, et al. (2012) *Indicators for mapping ecoystem services: a review* (Publications Office of the European Union, Luxembourg).

14. Martínez-Harms MJ, Balvanera P (2012) Methods for mapping ecosystem service supply: a review. *International Journal of Biodiversity Science, Ecosystem Services & Management* 8(1-2):17–25.

15. Bateman IJ, et al. (2013) Bringing ecosystem services into economic decision-making: land use in the United Kingdom. *Science* 341(6141):45–50.

16. Chan KMA, Shaw MR, Cameron DR, Underwood EC, Daily GC (2006) Conservation planning for ecosystem services. *PLoS biology* 4(11):e379.

17. Naidoo R, et al. (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences of the United States of America* 105(28):9495–500.

18. Queiroz C, et al. (2015) Mapping bundles of ecosystem services reveals distinct types of multifunctionality within a Swedish landscape. *Ambio* 44(1):89–101.

19. Qiu J, Turner MG (2013) Spatial interactions among ecosystem services in an urbanizing agricultural watershed. *Proceedings of the National Academy of Sciences of the United States of America* 110(29):12149–12154.

20. Remme RP, Schröter M, Hein L (2014) Developing spatial biophysical accounting for multiple ecosystem services. *Ecosystem Services* 10:6–18.

21. Willemen L, Hein L, Mensvoort ME van, Verburg PH (2010) Space for people, plants, and livestock? Quantifying interactions among multiple landscape functions in a Dutch rural region. *Ecological Indicators* 10(1):62–73.

22. Cord AF, et al. (2017) Priorities to Advance Monitoring of Ecosystem Services Using Earth Observation. *Trends in Ecology & Evolution* 32(6):416–428.

23. Raudsepp-Hearne C, Peterson GD, Bennett EM (2010) Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. *Proceedings of the National Academy of Sciences of the United States of America* 107(11):5242–5247.

24. Turner KG, Odgaard MV, Bøcher PK, Dalgaard T, Svenning JC (2014) Bundling ecosystem services in Denmark: Trade-offs and synergies in a cultural landscape. *Landscape and Urban Planning* 125:89–104.

25. Van der Biest K, et al. (2013) EBI-An Index for Delivery of Ecosystem Service Bundles Ecological Indicators. *Ecosystem Services: Global Issues, Local Practices* 37:263–272.

26. Yang G, et al. (2015) Using ecosystem service bundles to detect trade-offs and synergies across urban-rural complexes. *Landscape and Urban Planning* 136:110–121.

27. Birkhofer K, et al. (2015) Ecosystem services - current challenges and opportunities for ecological research. *Frontiers in Ecology and Evolution* 2(January):1–12.

28. Costanza R, et al. (2017) Twenty years of ecosystem services: How far have we come and how far do we still need to go? *Ecosystem Services* 28:1–16.

29. Olander L, et al. (2017) So you want your research to be relevant? Building the bridge between ecosystem services research and practice. *Ecosystem Services* 26:170–182.

30. Parks S, Gowdy J (2013) What have economists learned about valuing nature? A review essay. *Ecosystem Services* 3:e1–e10.

31. Gómez-Baggethun E, Ruiz-Pérez M (2011) Economic valuation and the commodification of ecosystem services. *Progress in Physical Geography* 35(5):613–628.

32. Schmidt S, Manceur AM, Seppelt R (2016) Uncertainty of monetary valued ecosystem services - value transfer functions for global mapping. *PLoS ONE* 11(3):e0148524.

33. Chan KMA, et al. (2016) Opinion: Why protect nature? Rethinking values and the environment. *Proceedings of the National Academy of Sciences* 113(6):1462–1465.

34. McCauley DJ (2006) Selling out on nature. *Nature* 443(7107):27–28.

35. Craig J, et al. (2000) Conservation Issues in New Zealand. *Annual Review of Ecology, Evolution and Systematics* 31:61–78.

36. Howe C, Suich H, Vira B, Mace GM (2014) Creating win-wins from trade-offs? Ecosystem services for human well-being: A meta-analysis of ecosystem service trade-offs and synergies in the real world. *Global Environmental Change* 28(1):263–275.

37. Lee H, Lautenbach S (2016) A quantitative review of relationships between ecosystem services. *Ecological Indicators* 66:340–351.

38. Malinga R, Gordon LJ, Jewitt G, Lindborg R (2015) Mapping ecosystem services across scales and continents - A review. *Ecosystem Services* 13:57–63.

39. Nieto-Romero M, Oteros-Rozas E, González Ja, Martín-López B (2014) Exploring the knowledge landscape of ecosystem services assessments in Mediterranean agroecosystems: Insights for future research. *Environmental Science & Policy* 37:121–133.

40. Seppelt R, Dormann CF, Eppink FV, Lautenbach S, Schmidt S (2011) A quantitative review of ecosystem service studies: Approaches, shortcomings and the road ahead. *Journal of Applied Ecology* 48(3):630–636.

41. Quintas-Soriano C, et al. (2016) Ecosystem services values in Spain: A meta-analysis. *Environmental Science and Policy* 55:186–195.

42. Reynaud A, Lanzanova D (2017) A Global Meta-Analysis of the Value of Ecosystem Services Provided by Lakes. *Ecological Economics* 137:184–194.

43. Leimona B, Noordwijk M van, Groot RS de, Leemans R (2015) Fairly efficient, efficiently fair: Lessons from designing and testing payment schemes for ecosystem services in Asia. *Ecosystem Services* 12:16–28.

44. Vatn A (2010) An institutional analysis of payments for environmental services. *Ecological Economics* 69(6):1245–1252.

45. Syrbe R-U, Grunewald K (2017) Ecosystem service supply and demand – the challenge to balance spatial mismatches. *International Journal of Biodiversity Science, Ecosystem Services & Management* 13(2):148–161.

46. CEE (2013) *Guidelines for Systematic Review and Evidence Synthesis in Environmental Management - Version 4.2* (Centre for Evidence Based Conservation).

47. MEA (2005) *Ecosystems and human well-being: Synthesis* (Island Press, Washington) Available at: http://www.who.int/entity/globalchange/ecosystems/ecosys.pdf.

48. Dale VH, et al. (2000) Ecological Principles and Guidelines for Managing the Use of Land. *Ecological Applications* 10(3):639–670.

49. MacLeod CJ, Moller H (2006) Intensification and diversification of New Zealand agriculture since 1960: An evaluation of current indicators of land use change. *Agriculture, Ecosystems and Environment* 115(1-4):201–218.

50. Wallace BC, Small K, Brodley CE, Lau J, Trikalinos TA (2012) Deploying an interactive machine learning system in an evidence-based practice center. *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics - IHI '12*:819.

51. Thompson S, Grüner I, Gapare N (2003) *Illustrated Guide to Target Classes of the New Zealand Land Cover Database Version 2* (New Zealand Ministry for the Environment).

52. Salanti G (2012) Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* 3(2):80–97.

53. Schwarzer G, Rücker G, Krahn U, König J (2017) Package 'netmeta' - A comprehensive set of functions providing frequentist methods for network meta-analysis. Available at: https://cran.r-project.org/web/packages/netmeta/index.html.

54. Rücker G, Schwarzer G (2015) Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Medical Research Methodology* 15(1):58.

55. Maechler M, et al. (2017) Package Cluster - "Finding Groups in Data": Cluster Analysis Extended. Available at: https://cran.r-project.org/web/packages/cluster/index.html.

56. R Core Team (2018) R: A language and environment for statistical computing. Available at: https://www.r-project.org/.

57. Oksanen J, et al. (2017) Package 'vegan'. Available at: https://cran.r-project.org/web/packages/vegan/index.html.

58. Landcare Research (2015) Land Cover Database version 4.1, Mainland New Zealand. Available at: http://www.lcdb.scinfo.org.nz/home.

59. Balvanera P, et al. (2012) Ecosystem services research in Latin America: The state of the art. *Ecosystem Services* 2:56–70.

60. Rodríguez-Loinaz G, Alday JG, Onaindia M (2015) Multiple ecosystem services landscape index: A tool for multifunctional landscapes conservation. *Journal of Environmental Management* 147:152–163.

61. Foley JA, et al. (2011) Solutions for a cultivated planet. *Nature* 478(7369):337–342.

62. Eastburn DJ, O'Geen AT, Tate KW, Roche LM (2017) Multiple ecosystem services in a working landscape. *PLoS One* 12(3):e0166595.

63. Fischer J, Lindenmayer DB, Manning AD (2006) Biodiversity, ecosystem function and resilience: Ten guiding principles for off-reserve conservation. *Frontiers in Ecology and the Environment* 4:80–86.

64. Law EA, et al. (2015) Better land-use allocation outperforms land sparing and land sharing approaches to conservation in Central Kalimantan, Indonesia. *Biological Conservation* 186:276–286.

65. Grau R, Kuemmerle T, Macchi L (2013) Beyond 'land sparing versus land sharing': Environmental heterogeneity, globalization and the balance between agricultural production and nature conservation. *Current Opinion in Environmental Sustainability* 5(5):477–483.

66. Mora C, Sale PF (2011) Ongoing global biodiversity loss and the need to move beyond protected areas: A review of the technical and practical shortcomings of protected areas on land and sea. *Marine Ecology Progress Series* 434:251–266.

67. Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity - ecosystem service management. *Ecology Letters* 8(8):857–874.

68. Holland RA, et al. (2011) The influence of temporal variation on relationships between ecosystem services. *Biodiversity and Conservation* 20(14):3285–3294.

69. Guerra CA, Pinto-Correia T (2016) Linking farm management and ecosystem service provision: Challenges and opportunities for soil erosion prevention in Mediterranean silvo-pastoral systems. *Land Use Policy* 51:54–65.

70. Maseyk FJ, Mackay AD, Possingham HP, Dominati EJ, Buckley YM (2017) Managing Natural Capital Stocks for the Provision of Ecosystem Services. *Conservation Letters* 10(2):211–220.

71. Pang X, Nordström EM, Böttcher H, Trubins R, Mörtberg U (2017) Trade-offs and synergies among ecosystem services under different forest management scenarios – The LEcA tool. *Ecosystem Services* 28:67–79.

72. Braat LC, Groot RS de (2012) The ecosystem services agenda: bridging the worlds of natural science and economics, conservation and development, and public and private policy. *Ecosystem Services* 1(1):4–15.

73. Englund O, Berndes G, Cederberg C (2017) How to analyse ecosystem services in landscapes—A systematic review. *Ecological Indicators* 73:492–504.

74. Czúcz B, et al. (2018) Where concepts meet the real world: A systematic review of ecosystem service indicators and their classification using CICES. *Ecosystem Services* 29:145–157.

75. Chan KMA, et al. (2012) Where are Cultural and Social in Ecosystem Services? A Framework for Constructive Engagement. *BioScience* 62(8):744–756.

76. Hernández-Morcillo M, Plieninger T, Bieling C (2013) An empirical review of cultural ecosystem service indicators. *Ecological Indicators* 29:434–444.

77. Lyver POB, et al. (2017) An indigenous community-based monitoring system for assessing forest health in New Zealand. *Biodiversity and Conservation* 26(13):3183–3212.

78. TEEB (2010) *The Economics of Ecosystems and Biodiversity: Mainstreaming the economics of Nature: A synthesis of the approach, conclusions and recommendations of TEEB.* (United Nations Environment Programme).

79. CICES (2018) Common International Classification of Ecosystem Services (CICES) V5.1 (January 2018). Available at: https://cices.eu/ [Accessed February 2, 2018].

80. Díaz S, et al. (2018) Assessing nature's contributions to people. *Science* 359(6373):270–272.

81. Dallimer M, et al. (2015) Historical influences on the current provision of multiple ecosystem services. *Global Environmental Change* 31:307–317.

82. Perring MP, et al. (2016) Global environmental change effects on ecosystems: The importance of land-use legacies. *Global Change Biology* 22(4):1361–1371.

Figure legends

Fig. 1. Hierarchical clustering of land covers. Land covers exhibiting a greater similarity in their provision of 8 ecosystem services are clustered closer to each other than covers with contrasting trends in service provision. Likewise, distances along the height axis indicate dissimilarity among clusters of land covers, with clusters that merge at a greater height exhibiting greater dissimilarity in service provision.

Fig. 2. Hierarchical clustering of ecosystem services. Ecosystem services that cluster together tend to be provided similarly across eight land covers. A greater separation between the branching points for clusters along the height axis indicates greater dissimilarity among clusters in the extent to which they deliver each of the services included in the analysis.

Fig. 3. Example visualizations for exploring land use trade-offs in the provision of multiple ecosystem services from entire landscapes. Quantitative measures of the provision of multiple ecosystem services by different land uses or land covers (such as those obtained from our meta-analysis) can be used to generate ordinations that 'map' land covers or land uses into the multidimensional space of ecosystem service provision (ordination graphs). Distribution of land covers (or uses) within that space can assist with identification of redundancies in service provision (among land covers/uses that map close together) and trade-offs among land covers/uses that provide contrasting sets of services and, consequently, occupy opposite extremes of the ordination space. Furthermore, the hypervolume enclosed by the total set of land covers/uses from a given landscape expresses the diversity of services provided by that landscape. As an example, our data can be used to compare multi-service provision for: a landscape with few, undifferentiated production land covers (Case 1); a landscape with a combination of some production and non-production land covers (Case 2) and a landscape with a broad range of production and non-production land covers that provide a diverse range of services (Case 3). In each case, the size of the points is proportional to the areal extent of the land cover.

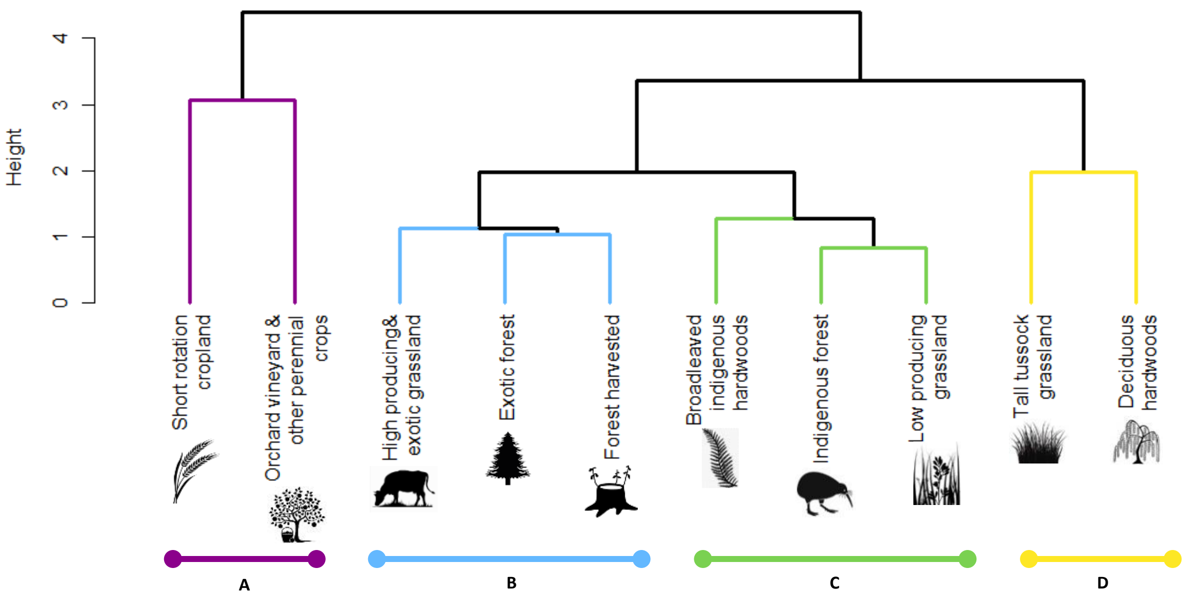


Fig. 1. *Hierarchical clustering of land covers. Land covers exhibiting a greater similarity in their provision of 8 ecosystem services are clustered closer to each other than covers with contrasting trends in service provision. Likewise, distances along the height axis indicate dissimilarity among clusters of land covers, with clusters that merge at a greater height exhibiting greater dissimilarity in service provision.*

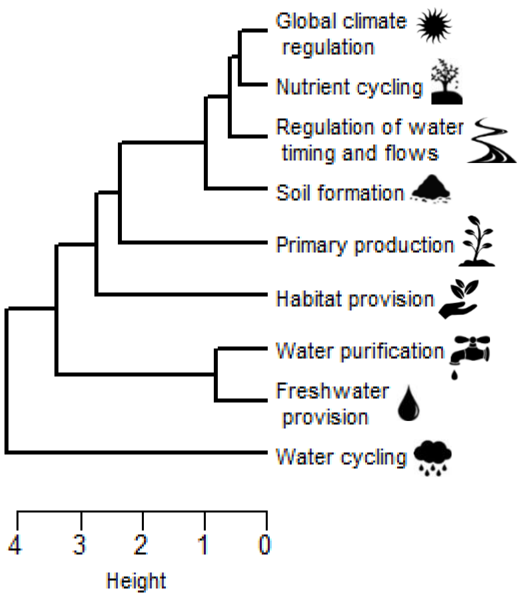


Fig. 2. *Hierarchical clustering of ecosystem services. Ecosystem services that cluster together tend to be provided similarly across eight land covers. A greater separation between the branching points for clusters along the height axis indicates greater dissimilarity among clusters.*

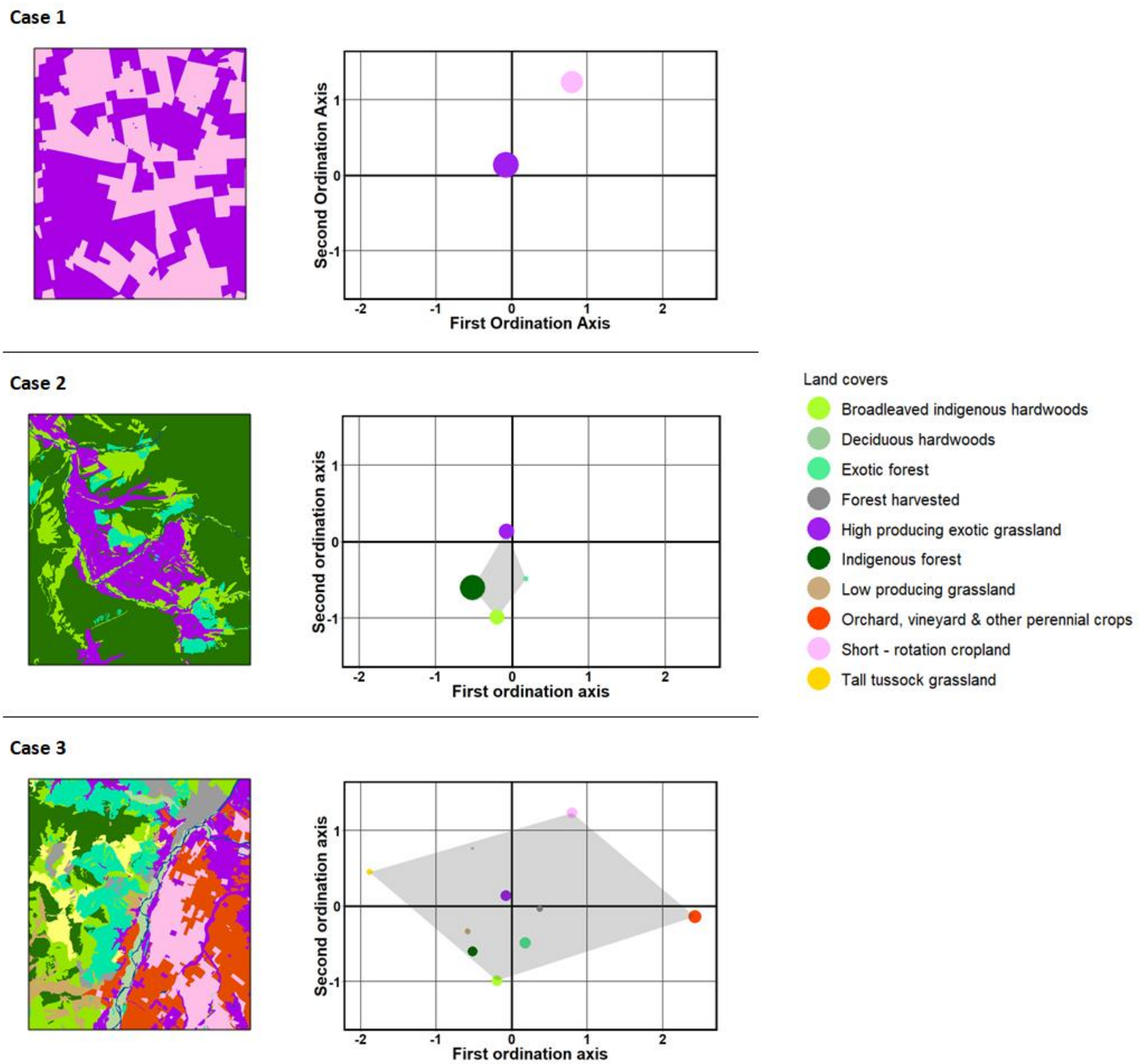


Fig. 3. *Example visualizations for exploring land use trade-offs in the provision of multiple ecosystem services from entire landscapes. Quantitative measures of the provision of multiple ecosystem services by different land uses or land covers (such as those obtained from our meta-analysis) can be used to generate ordinations that 'map' land covers or land uses into the multidimensional space of ecosystem service provision (ordination graphs). Distribution of land covers (or uses) within that space can assist with identification of redundancies in service provision (among land covers/uses that map close together) and trade-offs among land covers/uses that provide contrasting sets of services and, consequently, occupy opposite extremes of the ordination space. Furthermore, the hypervolume enclosed by the total set of land covers/uses from a given landscape expresses the diversity of services provided by that landscape. As an example, our data can be used to compare multi-service provision for: a landscape with few, undifferentiated production land covers (Case 1); a landscape with a combination of some production and non-production land covers (Case 2) and a landscape with a broad range of production and non-production land covers that provide a diverse range of services (Case 3). In each case, the size of the points is proportional to the areal extent of the land cover.*

Supplementary Information for: Meta-analysis reveals that the provision of multiple ecosystem services requires a diversity of land covers

Carla Gómez-Creutzberg[a,1], **Malgorzata Lagisz**[b], **Shinichi Nakagawa**[b], **Eckehard G. Brockerhoff**[c,d] and **Jason M. Tylianakis**[a,e,1]

[a] Centre for Integrative Ecology, University of Canterbury, Christchurch 8140, New Zealand; [b] Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia; [c] Scion (New Zealand Forest Research Institute), Christchurch 8440, New Zealand; [d] Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland (Present address); [e] Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, Berkshire SL5 7PY, United Kingdom.

[1] To whom correspondence may be addressed. Email: cgomezcre@gmail.com or jason.tylianakis@canterbury.ac.nz

This PDF file includes: SI Appendices 1 to 11

Other supplementary materials for this manuscript include SI Datasets 1 to 2

Appendix 1. Definitions

Table S1. Overview of Ecosystem services

Category	Ecosystem Service	Description
Provisioning	Food - crops	Cultivated plants for use by people or animals.
	Food - milk	Animals raised for domestic or commercial consumption or use.
	Food - meat	Animals raised for domestic or commercial consumption or use.
	Food - capture fisheries	Wild fish captured through trawling and other non-farming methods.

Category	Ecosystem Service	Description
	Food - aquaculture	Fish, shellfish, and/or plants that are bred and reared in ponds, enclosures.
	Food - wild plant and animal products	Plant and animal food sources gathered or caught in the wild.
	Fiber - timber and wood	Products made from trees harvested from forest ecosystems, plantations, or non-forested lands.
	Fiber - other Fibers	Non-wood and non-fuel based fibers sourced from the environment.
	Biomass fuel	Sources of fuel derived from plants and animals (wood, biofuel production, dung).
	Genetic resources	Genes and genetic information used for animal breeding, plant improvement, and biotechnology (MEA, 2005).
	Biochemicals, natural medicines and pharmaceuticals	Medicines, biocides, food additives, and other biological materials derived from ecosystems for commercial or domestic use.
	Ornamental resources	Products from nature that serve aesthetic purposes.
	Freshwater	Inland bodies of freshwater, groundwater, rainwater, and surface waters for household, industrial, and agricultural uses.
Regulating	Maintenance of air quality	Influence ecosystems have on air quality by either emitting chemicals to the atmosphere (reducing air quality) or extracting chemicals from the atmosphere (increasing air quality).
	Global climate regulation	Influence ecosystems have on the global climate by emitting greenhouse gases or aerosols to the

Category	Ecosystem Service	Description
		atmosphere, or by absorbing greenhouse gases or aerosols from the atmosphere.
	Regional and local climate regulation	Influence ecosystems have on local and regional climatic systems (expressed in local temperatures, rains, winter, frost frequency and other climatic factors).
	Regulation of water timing and flows	Influence ecosystems have on the timing and magnitude of water runoff, flooding, and aquifer recharge (particularly in terms of the water storage potential of the ecosystem or landscape).
	Erosion control	Role plants play in soil retention and the prevention of landslides.
	Water purification	Role ecosystems play in filtering nutrients, heavy metals and pollutants in water.
	Waste treatment	Role ecosystems play in decomposing organic wastes and recycling them; taking up and detoxifying compounds through soil and subsoil processes.
	Disease mitigation	Influence that ecosystems have on the incidence and abundance of human pathogens. Bio-control agents and pathogens limit the need for chemical interventions.
	Pest regulation	Influence ecosystems have on the amount of crop and livestock pests and diseases. Bio-control agents and pathogens limit the need for chemical interventions.
	Pollination	Role ecosystems play in transferring pollen between male and female plants.

Category	Ecosystem Service	Description
	Natural hazard regulation	Degree to which ecosystems reduce damage caused by natural hazards.
Cultural	Ethical and spiritual values	Spiritual, religious, aesthetic, intrinsic or other values people attach to ecosystems, landscapes or species.
	Educational and inspirational values	Information people get from ecosystems that is used for intellectual development, culture, art, design and innovation.
	Recreation and ecotourism	Recreation undertaken in nature, including tourism sector business and tourist activities that rely on natural or managed ecosystems.
Supporting	Habitat provision	Natural or semi-natural environments that provide all the necessary elements for the survival and reproduction of animal and plant populations and their capacity to recover after disturbances.
	Nutrient cycling	Cycling of essential nutrients for life (20 in total, includes nitrogen and phosphorous) (MEA, 2005).
	Soil formation	Rate of soil formation (MEA, 2005).
	Primary production	As a measure of the assimilation or accumulation of energy and nutrients by organisms (MEA, 2005).
	Water cycling	Different from freshwater provision in that it involves the cycling of water through ecosystems as a benefit for living organisms (MEA, 2005).

Table S2. Overview of land cover classes as defined by Thompson et al. (2003) for New Zealand's Land Cover Database – LCDB.

Group	Class	Description
Artificial surfaces	Built-up area (settlement)	Central business districts, suburban dwellings, commercial and industrial areas, horticultural sites dominated by structures and sealed surfaces. Includes associated hard surfaces and infrastructures (e.g., roads, carparks, paved areas), low density residential areas.
	Urban parkland/ Open space	Open, typically mown, within or associated with bolt-up areas. Includes parks (with scattered trees), playing fields, cemeteries, airports, golf courses, river berms). Hard surfaces, trees or scrub exceeding one hectare are classified separately.
	Surface mine & dump	Gravel pits and other open quarries or disposal areas for solid waste material.
	Transport infrastructure	Roads, railroads, airport runways, skid sites from forest logging that are discernible in the satellite imagery used for the classification (generally should exceed one hectare).
Bare or lightly vegetated surfaces	Sand, gravel and rock	Coastal strip, landward side of the coastline.
	Gravel and rock	Areas of gravel, sand and rock areas adjacent to rivers, streams and lakes. Also includes: bare ground associated with thermal activity; scree slopes, glacial debris, rock tor areas in hills and high lands and recently formed surfaces with little or no biomass.
	Landslide	Subsoil and parent material exposed due to localized erosion.

Group	Class	Description
	Permanent snow and ice	Areas greater than one hectare with perennial cover of snow and ice, includes glacier areas.
	Alpine grass / herbfield	Vegetation areas above the tree line, dominated by low growing and mat forming herbs and grasses. High bare ground components. Sites don't have a history of pastoral use.
Water bodies	Lake or pond	Permanently or intermittently standing open fresh water without emerging vegetation. Water bodies can be natural or artificial (oxidation ponds, fire control ponds and reservoirs).
	River	Flowing open freshwater without emerging vegetation. Includes natural and modified systems with a width greater than 30 meters.
	Estuarine open water	Standing or flowing open water areas without emerging vegetation and in which saline waters are occasionally or periodically diluted by freshwater (or freshwater is made saline). Includes estuaries of rivers, lagoons and dune swales.
Cropland	Short - rotation cropland	Areas where soil is exposed to cultivation regularly or at least annually. Includes land for growing cereal crops, root crops, annual seed crops, annual vegetable crops, hops, strawberry fields, annual flower crops and open ground nurseries.
	Orchard, vineyard & other perennial crops	Orchards and areas cultivated less than annually used for producing tree crops as well as crops grown on shrubs or climbing plants. Includes areas with perennial vines supporting grape crops.
Grassland,	High producing	Exotic grasslands with vigorous vegetation cover

Group	Class	Description
Sedgeland, Saltmarsh	exotic grassland	(clovers - *Trifolium* spp., ryegrass -*Lolium perenne*, cocksfoot - *Dactylis glomerata*). Usually comprises intensively managed grasslands, rotationally grazed for wool, lamb, beef, dairy and deer production. Management involves pasture renewal every five to ten years, application of fertilizers and in some cases irrigation. Class also includes extensively managed grasslands with low-producing grasses (browntop - *Agrostis capillaris* and sweet vernal - *Anthoxanthum odoratum*) that happen to show lush growth due to high soil fertility or annual rainfall.
	Low producing grassland	Exotic and indigenous grasslands with low plant vigor (seasonally varying) and biomass (which may be due to environmental conditions or management practices). Dominant species include less productive exotic grasses (browntop - *Agrostis capillaris* and sweet vernal - *Anthoxanthum odoratum*) usually mixed with short tussock species. Also includes areas of short tussock grassland hard tussock (*Festuca novaezelandiae*), blue tussock (*Poa colensoi*), and / or silver tussock (*Poa cita*).
	Tall tussock grassland	Highland areas that have not been under intensive farm management and with the presence of *Chionochloa* spp. (usually accompanied by short tussock and herbs - particularly *Celmisia* spp.). Can support extensive summer grazing which usually accounts for the presence of exotic grasses.
	Depleted	Grassland/herbfield areas with very low herbaceous

Group	Class	Description
	grassland	vegetation cover. Short tussock grassland species are present but have less than 10% cover. *Hieracium* spp. and/or exotic grassland species are conspicuous as is the bare ground. Low plant vigor is due to soil nutrient loss from repeated burning and overgrazing.
	Herbaceous freshwater vegetation	Permanent or periodical freshwater wetland areas with emergent herbaceous aquatic vegetation dominated by sedges (*Cyperaceae*), rushes (*Juncaceae*) or tall erect herbs (*Poaceae*, *Restionaceae*, *Typhaceae*). Includes areas with low growing dicotyledon herbs and with *Sphagnum* moss.
	Herbaceous saline vegetation	Estuarine or coastal wetland areas (with saline/brackish water or saltwater saturated soils) dominated by herbaceous aquatic vegetation. Dominance of salt-tolerant plants (*Schoenoplectus* spp., *Apodasmia similis*, or glasswort - *Sarcocornia quinqueflora*). Most areas are subject to tidal changes in water level.
	Flaxland	Areas dominated by lowland flax (*Phormium tenax*), usually moist and as part of wetland systems .
Scrub and/or Shrubland	Fernland	Areas where bracken fern (*Pteridium esculentum*), umbrella fern (*Gleichenia* spp.), and ring fern (*Paesia scaberula*) dominate. It represents a successional vegetation type in previously forested land and encompasses sites with low fertility and which have been recently burnt.
	Gorse and/or broom	Disturbed areas where low fertility, extensive grazing and fire have facilitated the spread and establishment

Group	Class	Description
		of gorse and/or broom.
	Manuka and/or kanuka	Indigenous shrubland, usually found as an early successional scrub type on previously forested land. Fires are used to maintain area in scrub stage and prevent and prevent succession into mature stands and forest. Both manuka and kanuka can occur but the former is more common in the South Island and the latter in the North Island.
	Matagouri or grey scrub	Matagouri (*Discaria toumatou*) is a thorny divaricate shrub found in open shrubland or thickets among montane areas of the South Island. It is usually associated with freely drained recent soils (river terraces, outwash fans) and can occur in areas under extensive farm management as a response to practices such as phosphate application. Grey scrub areas are dominated with small-leaved indigenous shrubs with divaricate growth (entangled fine branches at almost right angles to each other), and native climbers (e.g., *Muehlenbeckia* and *Parsonsia*), however, the dominant feature is the woody component (hence grey scrub).
	Broadleaved indigenous hardwoods	Areas usually in an advanced successional stage back to indigenous forest. Vegetation cover involves a mix of broad-leaved, usually seral (successional) broadleaved species (wineberry - *Aristotelia serrata*, mahoe - *Melicytus ramiflorus*, *Pseudopanax* spp., *Pittosporum* spp., *Fuchsia* spp., ngaio - *Myoporum laetum*, and titoki - *Alectryon excelsus*), tutu (*Coriaria* spp.) and tree ferns. Usually found in areas with high rainfall. Class also

Group	Class	Description
		includes low-growing coastal broadleaved forest.
	Sub-alpine Shrubland	Diverse range of communities occurring in the 900 - 1200 m.a.s.l range between indigenous forest and tall tussock grassland, alpine grass/herbfields and alpine gravel and rock. Can also be found at lower altitudes as secondary vegetation after forest clearance. Community composition and height strongly influenced by rainfall and exposure. Class includes frost flats (old tephra plains - plains formed by alluvial deposition of pumice stone).
	Mixed exotic shrubland	Single-species or mixed communities of introduced shrubs and climbers boxthorn, hawthorn (*Crataegus* spp.), elderberry (*Sambucus* spp.), brier (*Rosa rubiginosa*), buddleja (*Buddleja davidii*), blackberry (*Rubus* spp.), and old man's beard (*Clematis vitalba*). Includes areas of amenity planting where shrublands are larger than one hectare.
Forest	Exotic forest	Areas with *Pinus radiata* as well as exotic forests of conifers (douglas fir, Monterey cypress, larch) and broadleaved species (*Acacia*, *Eucalyptus*). Includes wilding pines (i.e. those that are growing spontaneously outside of plantations), when identifiable in the imagery. Also includes linear features of evergreen or deciduous trees extending for more than 150 meters.
	Forest - harvested	Areas with evidence of harvesting since the previous LCDB survey. Includes canopy openings, skidder tracks, new roads, log landings. It is assumed that these sites will become replanted if they are occuring inside

Group	Class	Description
		plantations, they are therefore marked as plantations and checked in the next LCDB survey iteration (which should occur every 5 years) by then the forest should be identifiable. Cleared areas of native forest are also assigned to this class unless the loss is due to localized erosion (in which case they are classified as landslides). This class is used to check the extent of harvested forest that is replanted and the indigenous forest that is converted to another land cover.
	Deciduous hardwoods	Willows and poplar species growing adjacent to inland water and rivers as well as planted deciduous hardwoods (oak - *Quercus* spp., ash - *Fraxinus excelsior* and elm - *Ulmus* spp.).
	Indigenous forest	Forest dominated by indigenous tall forest canopy species.
	Mangrove	*Avicenna officinalis* communities occurring in estuarine mudflats and tidal creeks. Distribution is confined to North Island (up to 38 degrees South).

Appendix 2. Detailed data collection and processing methods

Literature search

To assess the viability of our project, an initial scoping search was conducted using Google Scholar and some general terms to capture our review subject (essentially, "New Zealand", "land use" or variations of this, and keywords on some of the ecosystem services, see SI Appendix 3). After manually screening the results of the scoping search, 201 potentially relevant documents were identified, suggesting that the project would not be limited by insufficient evidence/data.

The multidisciplinary database Scopus was then selected for our formal search since it provided uniform access (i.e. independent of institutional subscription categories) to a comprehensive collection of abstracts and citations from international, peer-reviewed journals and serial books. We searched for titles, abstracts and keywords that contained at least one match in each of the 3 components that structured our search: 1) "New Zealand", 2) land cover and land use terms, and 3) ecosystem service terms (see SI Appendix 3 for the full search phrase). Land cover terms included all possible variations of "land use" and "land cover" as well as terms on specific land covers (both generic and specific to New Zealand. The ecosystem services component drew upon the names of each service (and possible variations of these) but also included vocabulary describing processes and conditions that could reflect their provision at the site scale akin to an individual land cover unit. To distill the final set of land-use and ecosystem service terms, 69 trial searches were conducted. This ensured that the final phrase, with approximately 840 terms, was sufficiently comprehensive. The search was finalized in December 2014, and was constrained to include documents published from 1970 onward, to be comparable with current land use regimes in New Zealand (1).

Document screening and assessment

In total, 9,741 citations matched our search criteria. The titles, abstracts and keywords of these citations were subjected to an automatic screening that removed any duplicates and selected only those that mentioned at least two different land cover terms. We conducted the latter step because our search returned studies with at least one land cover term whereas we required studies that compared two or more land covers. The abstracts of the 4,373 citations marked as relevant after this screening were then scanned to check whether they pertained to research that could potentially allow for the quantitative comparison of two or more different terrestrial land covers, in New Zealand, for the provision of any of the 35 ecosystem services we defined for the project.

Abstracts were screened using an interactive machine learning system for semi-automated abstract screening, Abstrackr, which is often used in medical meta-analyses (2). By using an active learning approach and a dual supervision classification algorithm, Abstrackr draws from the selection decisions and relevant/irrelevant words reviewers find in a sample of their abstracts to estimate the likelihood of the unscreened abstracts being relevant. The screening order of the remaining abstracts is subsequently prioritized according to this likelihood. Abstracts in our study were screened by two reviewers who, after checking for agreement in their decisions for a common pilot set of 500 abstracts, independently reviewed the remaining abstracts until a stopping point of 50 consecutive non-relevant citations was reached. This stopping point was reached after 2,957 abstracts were screened, leaving 1,416 unscreened ones, which the machine-learning algorithm deemed to be less relevant than the 50 that comprised our stopping point.

The abstract screening and the initial pilot search yielded 914 relevant abstracts, which were passed on to a team of 4 reviewers for full-text assessment and data extraction. The full-text assessment included the following inclusion criteria:

1. Selected studies had to present quantitative data derived from original research conducted in 1970 or later, and which did not duplicate data that had been published in other studies already in our analysis.

2. Only quantitative measures that could be taken as indicators of the provision of one (or more) ecosystem services (according to the service definitions in SI Appendix 1) were extracted from each study.

3. The data for each land cover had to come from at least two replicate observations. For any given study, a replicate observation was defined as one taken from a land cover unit that could be identified as a distinct spatial feature and which had sufficient separation from other units of the same land cover included in the same study so as to ascertain spatial independence of the observations. For the cases where the spatial separation of two units of the same land cover could not be readily ascertained, we applied a distance criterion. Instead of defining a fixed minimum distance between replicate land cover units (which could vary depending on the scale of a service), we defined as separate replicates any two units of the same land cover that were separated such that the distance between them was larger than the distance between any of them and any neighboring units of a different land cover.

Given the diversity of ecosystem services in our review, the documents we assessed comprised a very heterogeneous set of sampling and experimental designs, which meant that a land cover unit could range in size from whole forests and catchments to forest fragments, fields and crop plots; all of which were included as long as we could verify that the unit was spatially independent from other units of the same land cover and dominated in at least 80% by the same land cover type. Studies that did not have replicated observations (as defined above) for any land covers were discarded whereas studies that contained replication on some, but not all, of the land covers were kept and only data on the replicated land covers were extracted. If in the original studies' data were corrected for the potential effect of confounding covariates (i.e. slope), the corrected values were extracted instead of the uncorrected ones. Finally, within our review, the units of interest comprised only terrestrial land covers such that any comparisons between water bodies and any terrestrial land covers were not extracted. However, information on how different terrestrial land covers affected ecosystem services linked to a water body was included in our analysis. Full details of how the full-text selection criteria were applied can be found in SI Appendix 4.

Weekly meetings of the reviewers were held throughout the assessment and data extraction processes to ensure consistent implementation of these criteria. Authors were contacted whenever the data presented in a study were not readily extractable (due to legibility or formatting issues) or when further clarification was needed on the types of land covers involved in the study or the methods used to sample them. Authors of 96 studies were contacted with a 50% success response rate. In total, data from 260 studies were extracted.

Data aggregation and calculation of effect sizes

Extracted data from all studies were recorded in a database. As described in the main text, we assigned the quantitative measures of ecosystem service provision (indicators) reported by each study to one or more ecosystem services (SI Dataset 1) and defined whether service provision would generally increase or decrease with larger values of the indicator. Thematic experts (see Acknowledgements) were consulted when there was uncertainty in the allocation of an indicator to a service or the direction of the relation between both.

Our database included some cases where either a single document contained the results of multiple experiments (each with a unique method or indicator) or, conversely, different results of a single experiment were published in separate documents. The latter case included studies with partial duplication of the results in different publications and a case were the results for different land covers for the same experiment were published in separate documents. To bring these studies into comparable terms with those that had the results of experiments published in only one article, we generated new unique study identifiers such that, for all effects in our review, these cases would either be treated as separate studies (i.e. when a single publication presented the outcomes of multiple, separate experiments) or merged into a single study (i.e. where the outcomes of a single experiment was published in multiple studies with no or partial overlap). For cases where two or more publications contained duplicated results from a single experiment, only data from the publication with the most comprehensive set of results was kept in our dataset.

In addition, several studies in our database contained multiple or repeated measures of the same indicator within a single land cover replicate. To allow for a standardized comparison across all studies, these were summarized to a single value per land cover replicate. In cases where the multiple measures were taken at different soil or water depths, the measurement of the topmost layer that occurred across all the land covers in the study was taken as the summarized value. For all other cases, repeated or multiple measures were summarized into a single mean value of the service indicator per replicate. This is equivalent to aggregating data to one response value per individual patient in medical meta-analyses. Studies that did not provide enough information to allow for the aggregation of multiple/repeated measurements to the standard replication level had to be excluded from the analysis.

Some studies reported a summary for all replicates of the same land cover, for the remaining studies, the mean and variance across replicates of the same land cover (and ecosystem service indicator) was calculated. For studies where data were already summarized across land cover replicates, but were presented as medians with either absolute and/or interquartile ranges, conversions to means and variances were made following the methods defined by (3). Similarly, conversions from standard deviation, standard error and 95% confidence intervals to variance were also applied for summarized data that presented these measures of variation (see SI Appendix 6 for the equations to convert 95% confidence intervals). For studies that reported a mean value for the indicator per land cover but no measure of variance, Taylor's power law (4) was used to impute variances with estimates from a linear regression model of all reported means and variances in our dataset in log-log space (the full regression model can be found in SI Appendix 6). Imputed variances accounted for 11% of the records in the final dataset found in SI Dataset 2. Cases where data could not be converted into means with variances across replicate land covers (including data on only the maximum and minimum values, geometric means with no variation and medians with standard errors or lacking variation or sample size) were not included in the dataset.

We adopted a log Response Ratio as the standard effect measure for comparing pairs of land covers within each study. For each pair of land covers (A and B) in a study, this measure was estimated as ln(A) - ln(B) for the indicators in which larger values corresponded to an increase in service provision. For the indicators in which larger values reflected a decrease in service provision (e.g., when the amount of soil eroded was a measure of erosion control) the inverse of the log Response Ratio was used i.e. - (ln(A) - ln(B)). For all cases the variance of the log Response Ratio was estimated as (5):

$$v_{RR} = \frac{s_A^2}{n_1 \overline{Y}_A^2} + \frac{s_B^2}{n_2 \overline{Y}_B^2}$$

Where v_{RR} is the variance for the log Response Ratio, $\overline{Y}$ denotes the mean value of the ecosystem service indicator for land covers A and B, with variance *s* and sampling size *n*. For cases where both land covers had negative numbers in the indicator, we based the ratio and variance calculations on the absolute indicator values and took the inverse of the ratio calculated with absolute values as the final value. Similarly, in cases where at least one of the land covers had a value of zero for the ecosystem service indicator, we added a small value (3 orders of magnitude smaller than the smallest value in the dataset) to the zero to allow for the ratio and variance calculations.

Studies that presented multiple indicators for a single ecosystem service from each site (e.g., measures of the cycling rates of various different nutrients), or which presented data on the same indicator expressed in two different units, required further aggregation of their log Response Ratios to avoid giving these studies disproportionate weight in the analysis. For some studies, not all the indicators of a service (or different units for the same indicator) had information from all of the land covers for the study. Thus, for each study, we only aggregated the indicators of the same ecosystem service (or sets of data with different units on the same indicator) that were measured from all of the land covers present in the study for that service, and excluded from the analysis any indicators that were not presented for the full range of land covers. If only a single indicator (or unit) contained data on all of the land covers then the other indicators of that service were excluded for that study. If two or more indicators (or units) contained data for the full set of land covers, then data were aggregated by taking a mean of the log Response Ratios and corresponding variances across the different indicators for each pairwise combination of land covers of that service in that study. The final dataset with aggregated log Response Ratios for each study and ecosystem service can be found in SI Dataset 2.

Appendix 4. Full search phrase for pilot and formal searches

A. Phrase used in formal search

Database used: Scopus

Please consult https://dev.elsevier.com/tips/ScopusSearchTips.htm for an explanation of booloean operators and wildcards used

Search was conducted on the 26th January, 2015.

Phrase:

TOPIC: ("New Zealand")

AND

TOPIC: ("land use" OR "land cover" OR habitat OR "vegetation type" OR ecosystem* OR forest* OR plantation OR scrub* OR shrub* OR pasture* OR grass* OR crop* OR "tussock grassland" OR "grey scrub" OR bush OR herbfield OR catchment OR "drainage basin" OR watershed* OR wetland* OR river OR lake OR peatland OR marsh OR bog OR fernland OR flaxland OR matagouri OR mangrove OR orchard OR estuary OR urban OR mine OR town OR city OR residential OR park OR garden)

AND

TOPIC: ("ecosystem service*" OR "habitat quality" OR "habitat provision" OR "nursery provision" OR "habitat diversity" OR "habitat complexity" OR "habitat feature" OR "habitat character*" OR ((feeding OR resting OR roosting OR nesting OR brood OR foraging OR mating) NEAR (site* OR cover)) OR "vegetation cover" OR "food availability" OR nurser* OR nitrogen OR phosphorus OR potassium OR calcium OR magnesium OR sulphur OR nitrification OR "soil organic matter" OR nitrification OR fixation OR "nutrient cycl*" OR "chemical cycl*" OR "decomposition" OR "nutrient uptake" OR "nutrient export" OR detritus OR bacteria OR microorganism OR "biogeochemical cycl*" OR microbial OR decomposition OR "soil formation" OR weathering OR "humification" OR "mineralization" OR pedogen* OR "soil quality" OR "soil fertility" OR "soil nutrients" OR "nutrient leaching" OR "microbial biomass" OR "nutrient storage" OR "soil structure" OR "nutrient assimilation" OR biomass OR "primary production" OR "primary productivity" OR litter* accumulation OR aboveground OR belowground OR NPP OR "carbon allocation" OR "productivity allocation"

OR "air quality" OR pollut* OR "nitrogen oxide*" OR "sulphur and oxide*" OR aerosol* OR "atmospheric cleansing capacity" OR "tropospheric oxidizing capacity" OR "acid rain" OR particulate OR "volatile organic compounds" OR "carbon stock" OR "total carbon" OR "total C" OR "carbon storage" OR emission* OR "carbon loss" OR ((carbon OR methane OR "tropospheric ozone" OR aerosol* OR "greenhouse gas*" OR "nitrous oxide") SAME (emission* OR sink OR sequestration)) OR albedo OR "heat flux" OR evapotranspiration OR precipitation OR rainfall OR temperature OR wind OR humidity OR "climate regulation" OR "climatic variability" OR runoff OR interception OR infiltration OR "water flow" OR discharge OR "water retention" OR "lag time" OR "water storage" OR "aquifer recharge" OR "stream*flow" OR "water yield" OR "water balance" OR "base*flow" OR percolation OR "flow regime" OR "flow regulation" OR erosion OR gullying OR gully OR "soil cover" OR "vegetation cover" OR rill OR "soil loss" OR "sediment yield" OR "sediment retention" OR "soil stability" OR "soil compaction" OR "aquatic or pollution" OR "water quality" OR "water purification" OR "water filtration" OR "filtration" OR "dissolved organic carbon" OR "heavy metals" OR "dissolved oxygen" OR "nutrient retention" OR "microbial degradation" OR "benthic indicators" OR "nutrient removal" OR "maximum daily loads" OR "load" OR (waste SAME (regulat* or treat* or assimilat* or decompos* or process* or degrad*)) OR pollut* OR toxic* OR contaminant* OR "detoxification" OR "soil pollut*" OR "nutrient retention" OR remineralisation OR "human AND pathogen*" OR " human AND disease*" OR "infectious disease*" OR (propagation SAME (disease OR vector OR pathogen)) OR "disease vector" OR "pathogen infect*" OR "disease risk" OR "disease incidence" OR "ecology of disease" OR "disease ecology" OR "vector control" OR "invasion resistance" OR "pest control" OR " pest management" OR biocontrol OR "biological control" OR "biological pest control" OR "natural pest control" OR weed* OR "pest predat*" OR ("natural enemy" SAME (conservation OR augmentation)) OR "seed set" OR pollinat* OR "flower visit*" OR zoophilus OR ornithophilous OR melittophilous OR entomophilous OR fruit OR "crop plants" OR "hazard mitigation" OR "disaster reduction" OR "disaster risk reduction" OR buffer* OR ((storm OR flood OR drought OR fire OR landslide OR avalanche OR "mass movement" OR hurricane OR windstorm) SAME (protect* OR buffer* OR mitigate* OR attenuat* OR defen*)) OR "flood storage" OR "storm flow" OR "peak flow" OR "extreme event*" OR "storm peak" OR ((timber OR round*wood OR pulp*wood OR wood) NEAR (harvest* OR yield OR extraction OR production)) OR “forest product” OR “Non-timber forest product” OR “non-wood forest product” OR ((fiber OR leather OR hemp OR hide OR “merino wool” OR yarn OR “alpaca wool” OR “merino wool” OR “possum wool” OR “possum fur” OR “harakeke flax” OR “flax fibre” OR wool OR fur) SAME (production OR supply OR provision OR yieldOR extraction)) OR ((“fuel wood” OR “wood fuel” OR firewood) SAME (production OR extraction))OR biofuel OR “biomass energy” OR biogas OR biodiesel OR “woody biomass” OR “cultural identity” OR “maori” OR “livelihood” OR ((sacred OR spiritual) SAME (site OR landscape OR place OR plant OR animal* OR ecosystem)) OR “spiritual inspiration” OR “ritual site” OR “sacred grove” OR “sense of belonging” OR aesthetics OR “scenic value” OR “environmental attribute*” OR “site attribute*” OR “aesthetic enjoyment” OR “aesthetic preference” OR “environmental aesthetics” OR “scenic beauty” OR “aesthetic pleasure” OR “environmental perception” OR wilderness OR “landscape preference*” OR “visual landscape” OR hedonic OR “cultural importance” OR “archaeological site*” OR “historic site” OR “heritage site” OR “ancestral site” OR “cultural landscape” OR “cultural heritage” OR “cultural site” OR “cultural attribute* OR”traditional landscape" OR landmark* OR “ritual site” OR “burial site” OR “tribal landmark” OR “natural heritage place*” OR “Maori site*” OR “intellectual development” OR “traditional knowledge system” OR ethnobotan* OR ethnobiolog* OR “maatauranga maori” OR “experimental farm” OR “educational forest” OR “educational farm” OR “distribution research project” OR “distribution research locations” OR “distribution research site*” OR “didactic farm” OR “didactic forest” OR “educational visit*” OR “school visit*” OR “field trip*” OR “field station” OR “research site*” OR “research location*” OR “research project location*” OR inspirational site OR ((movie OR film OR photograph* OR painting) SAME (setting OR location)) OR “maori art” OR craft* OR “inspiration from nature” OR “nature in art” OR “nature in film” OR “nature in literature” OR biomimicry OR bionics OR biomimet* OR recreation OR visit* OR tourism OR “nature tourism” OR ecotourism OR “adventure tourism” OR “rural tourism” OR “agri*tourism” OR “cultural tourism” OR “nature*based tourism” OR “nature*based recreation” OR angling OR hiking OR tramping OR birding OR hunting OR fishing OR mountaineering OR alpinism OR walking OR kayaking OR rowing OR surfing OR sailing OR rafting OR canoeing OR skiing OR “snow sport*” OR “winter sport*”

OR windsurfing OR “kites surfing” OR horse riding" OR caving OR “outdoor sport*” OR rappelling OR abseiling)

B. Phrases used in pilot search

Database used: Google Scholar

Searches were conducted between Feburary - April , 2014.

1. “New Zealand” AND “land use” AND diversity
2. “New Zealand” AND “land use” AND biodiversity
3. “New Zealand” AND “land use” AND H’
4. “New Zealand” AND “land use” AND evenness
5. “New Zealand” AND “land use” AND “species richness”
6. “New Zealand” AND “land use” AND “species abundance”
7. “New Zealand” AND “land use” AND insect
8. “New Zealand” AND “land use” AND arthropod
9. “New Zealand” AND “land use” AND invertebrate
10. “New Zealand” AND “land use” AND Collembola
11. “New Zealand” AND “land use” AND Diptera
12. “New Zealand” AND “land use” AND beetle
13. “New Zealand” AND “land use” AND Coleoptera
14. “New Zealand” AND “land use” AND arachnid
15. “New Zealand” AND “land use” AND spider
16. “New Zealand” AND “land use” AND mite
17. “New Zealand” AND “land use” AND Acari
18. “New Zealand” AND “land use” AND bird
19. “New Zealand” AND “land use” AND avifauna
20. “New Zealand” AND “land use” AND plant
21. “New Zealand” AND “land use” AND vegetation
22. “New Zealand” AND “land use” AND aquatic
23. “New Zealand” AND “land use” AND stream

24. “New Zealand” AND “land use” AND “water yield”
25. “New Zealand” AND “land use” AND “water quality”
26. “New Zealand” AND “land use” AND soil
27. “New Zealand” AND “land use” AND “microbial biomass”
28. “New Zealand” AND “land use” AND microbes
29. “New Zealand” AND “land use” AND biota
30. “New Zealand” AND “land use” AND fungi
31. “New Zealand” AND “land use” AND mycorrhiza
32. “New Zealand” AND “land use” AND nutrient
33. “New Zealand” AND “land use” AND nitrogen
34. “New Zealand” AND “land use” AND phosphorus
35. “New Zealand” AND “land use” AND potassium
36. “New Zealand” AND “land use” AND carbon
37. “New Zealand” AND “land use” AND methane
38. “New Zealand” AND “land use” AND transpiration
39. “New Zealand” AND “land use” AND evapotranspiration
40. “New Zealand” AND “land use” AND photosynthesis

Searches also substituted “land use” with:

1. “catchment”
2. “paired catchment”
3. “vegetation type”
4. “site” and “comparison”
5. “forest” and “tussock”
6. “forest” and native bush"
7. “forest” and “pasture”
8. “plantation” and “native”
9. “plantation” and “pasture”
10. “plantation” and “tussock”

11. “native and pasture”
12. “native and tussock”
13. “tussock” and “bush”
14. “tussock” and “pasture”
15. “paired catchment” and “forest”
16. “paired catchment” and “tussock”
17. “paired catchment” and “pasture”
18. “paired catchment” and “bush”
19. “paired catchment” and “native”
20. Pinus
21. Podocarp
22. Broadlea*
23. Chionochloa
24. Nothofagus
25. Pasture
26. Hieracium
27. Scrub
28. Shrubland
29. “Grey shrub”

Appendix 4. Decision tree for full-text assessment

Fig. S1: Decision tree with the selection criteria used in the full-text assessment of publications with relevant abstracts for our review.

Appendix 5. List of studies included in our final dataset

Table S3. Reference list for the studies included in our meta-analysis. The Study ID values are not sequential because they correspond to the unique identifier that each study received at the start of the literature screening process. They can be used to link the values presented in SI Dataset 2 to the bibliographical reference of each study.

Study ID	Reference
S0008	Fahey, B. & Watson, A. (1991). Hydrological impacts of converting tussock grassland to pine plantation, Otago, New Zealand. New Zealand Journal of Hydrology, 30, 1–15.
S0010	Tate, K.R., Ross, D.J., Saggar, S., Hedley, C.B., Dando, J. & Singh, B.K. et al. (2007). Methane uptake in soils from *Pinus radiata* plantations, a reverting shrubland and adjacent pastures: Effects of land-use change, and soil texture, water and mineral nitrogen. Soil Biology and Biochemistry, 39, 1437–1449.
S0011	Yeates, G.W., Hawke, M.F. & Rijkse, W.C. (2000). Changes soil fauna and soil conditions under *Pinus radiata* agroforestry regimes during a 25-year tree rotation. Biology and Fertility of Soils, 31, 391–406.
S0013	Thompson, R.M. & Townsend, C.R. (2003). Impacts on Stream Food Webs of Native and Exotic Forest : An Intercontinental Comparison. Ecology, 84, 145–161.
S0014	Thompson, R.M. & Townsend, C.R. (2004b). Impacts of riparian afforestation on stream biofilms: An exotic forest-native grassland comparison. New Zealand Journal of Marine and Freshwater Research, 38, 895–902.
S0015	Giddens, K.M., Parfitt, R.L. & Percival, H.J. (1997). Comparison of some soil properties under *Pinus radiata* and improved pasture. New Zealand Journal of Agricultural Research, 40, 409–416.
S0017	Adams, M.L., Davis, M.R. & Powell, K. (2001). Effects of grassland

Study ID	Reference
	afforestation on exchangeable soil and soil solution aluminium. Australian Journal of Soil Research, 39, 1003–1014.
S0019	Quinn, J.M. & Stroud, M.J. (2002). Water quality and sediment and nutrient export from New Zealand hill-land catchments of contrasting land use. New Zealand Journal of Marine and Freshwater Research, 36, 409–429.
S0020	Boulton, A.J., Scarsbrook, M.R., Quinn, J.M. & Burrell, G.P. (1997). Land-use effects on the hyporheic ecology of five small streams near Hamilton, New Zealand. New Zealand Journal of Marine and Freshwater Research, 31, 609–622.
S0021	Broekhuizen, N. & Quinn, J.M. (1998). Influences of stream size and catchment land-use on fine particulate organic matter retention in streams. New Zealand Journal of Marine and Freshwater Research, 32, 581–590.
S0022	Thompson, R.M. & Townsend, C.R. (2004c). Land-use influences on New Zealand stream communities: Effects on species composition, functional organisation, and food-web structure. New Zealand Journal of Marine and Freshwater Research, 38, 595–608.
S0030	Singh, B.K., Tate, K.R., Ross, D.J., Singh, J., Dando, J. & Thomas, N. et al. (2009). Soil methane oxidation and methanotroph responses to afforestation of pastures with *Pinus radiata* stands. Soil Biology and Biochemistry, 41, 2196–2205.
S0035	Jacobsen, L.B. (2012). Interacting effects of land use and landscape context on wild bees (*Apoidea*) in Canterbury, New Zealand. PhD thesis. University of Copenhagen.
S0036	Hughes, A.O., Quinn, J.M. & McKergow, L.A. (2012). Land use influences on suspended sediment yields and event sediment dynamics within two headwater catchments, Waikato, New Zealand. New Zealand Journal of Marine and Freshwater Research, 46, 315–333

Study ID	Reference
S0047	Fahey, B., Marden, M. & Phillips, C. (2003b). Sediment yields from plantation forestry and pastoral farming, coastal Hawke's Bay, North Island, New Zealand. Journal of Hydrology New Zealand, 42, 27–38.
S0048	Sparling, G.P., Shepherd, T.G. & Schipper, L.A. (2000). Topsoil characteristics of three contrasting New Zealand soils under four long-term land uses. New Zealand Journal of Agricultural Research, 43, 569–583.
S0050	Quinn, J.M., Cooper, A.B., Davies-Colley, R.J., Rutherford, J.C. & Williamson, R.B. (1997). Land use effects on habitat, water quality, periphyton, and benthic invertebrates in Waikato, New Zealand, hill-country streams. New Zealand Journal of Marine and Freshwater Research, 31, 579–597.
S0051	Hicks, B.J. & McCaughan, H.M. (1997). Land use, associated eel production, and abundance of fish and crayfish in streams in Waikato, New Zealand. New Zealand Journal of Marine and Freshwater Research, 31, 635–650.
S0052	Warburton, B., Cowan, P. & Shepherd, J. (2009). How many possums are now in New Zealand following control and how many would there be without it? - Landcare Research Contract Report LC0910/060. Landcare Research.
S0056	Parkyn, S.M., Davies-Colley, R.J., Scarsbrook, M.R., Halliday, N.J., Nagels, J.W. & Marden, M. et al. (2006). Pine afforestation and stream health: a comparison of land-use in two soft rock catchments, East Cape, New Zealand. New Zealand Natural Sciences, 31, 113–135.
S0057	Duncan, M.J. (1995). Hydrological impacts of converting pasture and gorse to pine plantation, and forest harvesting, Nelson, New Zealand. Journal of Hydrology (NZ), 34, 15–41.
S0060	Mark, A.F. & Rowley, J. (1976). Water Yield of Low-Alpine Snow Tussock Grassland in Central Otago. Journal of Hydrology (NZ), 15, 59 - 79.

Study ID	Reference
S0061	Holdsworth, D.K. & Mark, A.F. (1990). Water and nutrient input:output budgets: effects of plant cover at seven sites in upland snow tussock grasslands of eastern and central Otago, New Zealand. Journal - Royal Society of New Zealand, 20, 1–24.
S10035	Selby, M.J. & Hosking, P.J. (1973). The erodibility of pumice soils of the North Island, New Zealand. Journal of Hydrology New Zealand, 12, 32–56.
S10168	Cotching, W.E., Allbrook, R.F. & Gibbs, H.S. (1979). Influence of maize cropping on the soil structure of two soils in the Waikato district, New Zealand. New Zealand Journal of Agricultural Research, 22, 431–438.
S10174	Mosley, M.P. (1979). Sediment sources in the Harper-Avoca Catchment - New Zealand Forest Service Technical Paper No.68. New Zealand Forest Service, Wellington.
S10286	O'Loughlin, C.L., Rowe, L.K. & Pearce, A.J. (1982). Exceptional Storm Influences on Slope Erosion and Sediment Yield in Small Forest Catchments, North Westland, New Zealand. In: The first national symp. on forest hydrology. Melbourne, pp. 84–91.
S10908	Moore, T.R. (1989). Dynamics of dissolved organic carbon in forested and disturbed catchments, Westland, New Zealand. Water Resources Research, 25, 1331–1339.
S11177	Harris, R.J., Thomas, C.D. & Moller, H. (1991). The influence of habitat use and foraging on the replacement of one introduced wasp species by another in New Zealand. Ecological Entomology, 16, 441–448.
S11224EX003	Sparling, G.P. (1992). Ratio of microbial biomass carbon to soil organic carbon as a sensitive indicator of changes in soil organic matter. Soil Research, 30, 195–207.
S11336	Sparling, G.P. & Searle, P. (1993). Dimethyl sulphoxide reduction as a sensitive indicator of microbial activity in soil: The relationship with

Study ID	Reference
	microbial biomass and mineralization of nitrogen and sulphur. Soil Biology and Biochemistry, 25, 251–256.
S11397	Linklater, W. & Winterbourn, M.J. (1993). Life histories and production of two trichopteran shredders in New Zealand streams with different riparian vegetation. New Zealand Journal of Marine and Freshwater Research, 27, 61–70.
S11528	Murphy, E.C. & Dowding, J.E. (1994). Range and diet of stoats (*Mustela erminea*) in a New Zealand beech forest. New Zealand Journal of Ecology, 18, 11–18.
S11694	Bergin, D.O., Kimberley, M.O. & Marden, M. (1995). Protective value of regenerating tea tree stands on erosion-prone hill country, East Coast, North Island, New Zealand. New Zealand Journal of Forestry Science, 25, 3–19.
S11817	Edwards, E. & Huryn, A.D. (1996). Effect of riparian land use on contributions of terrestrial invertebrates to streams. Hydrobiologia, 337, 151–159.
S11863	Young, R.G. & Huryn, A.D. (1996). Interannual variation in discharge controls ecosystem metabolism along a grassland river continuum. Canadian Journal of Fisheries and Aquatic Sciences, 53, 2199–2211.
S11864	Stankiewicz, M., Jowett, G.H., Roberts, M.G., Heath, D.D., Cowan, P. & Clark, J.M. et al. (1996). Internal and external parasites of possums (*Trichosurus vulpecula*) from forest and farmland, Wanganui, New Zealand. New Zealand Journal of Zoology, 23, 345–353.
S11912	Friberg, N., Winterbourn, M.J., Shearer, K.A. & Larsen, S.E. (1997). Benthic communities of forest streams in the South Island, New Zealand: effects of forest type and location. Archiv für Hydrobiologie, 138, 289–306.
S11997	Friberg, N. & Winterbourn, M.J. (1997). Effects of native and exotic forest

Study ID	Reference
	on benthic stream biota in New Zealand: A colonization study. Marine and Freshwater Research, 48, 267–75.
S12036	Yeates, G.W., Saggar, S. & Daly, B.K. (1997). Soil microbial C, N, and P, and microfaunal populations under *Pinus radiata* and grazed pasture land-use systems. Pedobiologia, 41, 549–565.
S12049	Findlay, S., Hickey, C.W. & Quinn, J.M. (1997). Microbial enzymatic response to catchment-scale variations in supply of dissolved organic carbon. New Zealand Journal of Marine and Freshwater Research, 31, 701–706.
S12051	Fahey, B. & Jackson, R.J. (1997). Environmental effects of forestry at Big Bush Forest, South Island, New Zealand: I. Changes in water chemistry. Journal of Hydrology (NZ), 36, 43–71.
S12058	Davies-Colley, R.J. (1997). Stream channels are narrower in pasture than in forest. New Zealand Journal of Marine and Freshwater Research, 31, 599–608.
S12125	Storey, R.G. & Cowley, D.R. (1997). Recovery of three New Zealand rural stream as they pass through native forest remnants. Hydrobiologia, 353, 63–76.
S12178	Wilcock, R.J., Nagels, J.W., McBride, G.B., Collier, K.J., Wilson, B.T. & Huser, B.A. (1998). Characterisation of lowland streams using a single-station diurnal curve analysis model with continuous monitoring data for dissolved oxygen and temperature. New Zealand Journal of Marine and Freshwater Research, 32, 67–79.
S12276	Francis, G.S., Bartley, K.M. & Tabley, F.J. (1998). The effect of winter cover crop management on nitrate leaching losses and crop growth. The Journal of Agricultural Science, 131, 299–308.
S12283	Davies-Colley, R.J. & Quinn, J.M. (1998). Stream lighting in five regions of

Study ID	Reference
	North Island, New Zealand: Control by channel size and riparian vegetation. New Zealand Journal of Marine and Freshwater Research, 32, 591–605.
S12425	Scott, N.A., Tate, K.R., Ford-Robertson, J., Giltrap, D.J. & Smith, C.T. (1999). Soil carbon storage in plantation forests and pastures: land-use change implications. Tellus B, 51, 326–335.
S12482	Aslam, T., Choudhary, M.A. & Saggar, S. (1999). Tillage impacts on soil microbial biomass C, N and P, earthworms and agronomy after two years of cropping following permanent pasture in New Zealand. Soil and Tillage Research, 51, 103–111.
S12669	Francis, G.S., Tabley, F.J. & White, K.M. (1999). Restorative crops for the amelioration of degraded soil conditions in New Zealand. Australian Journal of Soil Research Aust. J. Soil Res, 37, 1017–34.
S12681	Schipper, L.A. & Sparling, G.P. (2000). Performance of Soil Condition Indicators Across Taxonomic Groups and Land Uses. Soil Science Society of America Journal, 64, 300.
S12799	Murphy, C. & Robertson, A. (2000). Preliminary study of the effects of honey bees (*Apis mellifera*) in Tongariro National Park. Science for conservation, 139, 5–18.
S12815	Whitmore, N., Alexander, D., Huryn, A.D., Arbuckle, C. & Jansma, F. (2000). Ecology and distribution of the freshwater crayfish *Paranephrops zealandicus* in Otago implications for conservation. - Science for conservation 148. Department of Conservation.
S12818	White, J.D., Coops, N.C. & Scott, N.A. (2000). Estimates of New Zealand forest and scrub biomass from the 3-PG model. Ecological Modelling, 131, 175–190.
S12865	O'Donnell, C.F. (2000). Influence of season, habitat, temperature, and

Study ID	Reference
	invertebrate availability on nocturnal activity of the new zealand long-tailed bat (*Chalinolobus tuberculatus*). New Zealand Journal of Zoology, 27, 207–221.
S13002	Mahmood, B. & Wall, G.L. (2001). The environmental impact of sewage effluent irrigation onto land - A case study in New Zealand. International Agricultural Engineering Journal, 10, 209–230.
S13051	Innes, J.G., King, C.M., Flux, M. & Kimberley, M.O. (2001). Population biology of the ship rat and Norway rat in Pureora forest park, 1983–87. New Zealand Journal of Zoology, 28, 57–78.
S13161	McLay, C.D.A., Dragten, R., Sparling, G.P. & Selvarajah, N. (2001). Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: A comparison of three approaches. Environmental Pollution, 115, 191–204.
S13208	Hall, M.J., Closs, G.P. & Riley, R.H. (2001). Relationships between land use and stream invertebrate community structure in a South Island, New Zealand, coastal stream catchment. New Zealand Journal of Marine and Freshwater Research, 35, 591–603.
S13210	Scarsbrook, M.R., Quinn, J.M., Halliday, J. & Morse, R. (2001). Factors controlling litter input dynamics in streams draining pasture, pine, and native forest catchments. New Zealand Journal of Marine and Freshwater Research, 35, 751–762.
S13213	Broad, T.L., Townsend, C.R., Closs, G.P. & Jellyman, D.J. (2001). Microhabitat use by longfin eels in New Zealand streams with contrasting riparian vegetation. Journal of Fish Biology, 59, 1385–1400.
S13282	Baillie, B.R. & Davies, T.R. (2002). Effects of land use on the channel morphology of streams in the Moutere Gravels, Nelson, New Zealand. Journal of Hydrology New Zealand, 41, 19–45.

Study ID	Reference
S13283	Rowe, D.K., Smith, J., Quinn, J.M. & Boothroyd, I. (2002). Effects of logging with and without riparian strips on fish species abundance, mean size, and the structure of native fish assemblages in Coromandel, New Zealand, streams. New Zealand Journal of Marine and Freshwater Research, 36, 67–79.
S13336	McQueen, D.J. & Shepherd, T.G. (2002). Physical changes and compaction sensitivity of a fine-textured, poorly drained soil (Typic Endoaquept) under varying durations of cropping, Manawatu region, New Zealand. Soil and Tillage Research, 63, 93–107.
S13361	Parfitt, R.L., Parshotam, A. & Salt, G. (2002). Carbon turnover in two soils with contrasting mineralogy under long-term maize and pasture. Australian Journal of Soil Research, 40, 127–136.
S13422	Groenendijk, F.M., Condron, L.M. & Rijkse, W.C. (2002). Effects of afforestation on organic carbon, nitrogen and sulfur concentrations in New Zealand hill country soils. Geoderma, 108, 91–100.
S13524	Prebble, M., Schallenberg, M., Carterm, J. & Shulmeister, J. (2002). An analysis of phytolith assemblages for the quantitative reconstruction of late Quaternary environments of the Lower Taieri Plain,Otago, South Island, New Zealand I. Modern assemblages and transfer functions. Journal of Paleolimnology, 27, 393–413.
S13526	Sparling, G.P. & Schipper, L.A. (2002). Ecological risk assessment: Soil quality at a national scale in New Zealand. Journal of Environmental Quality; Madison, 31, 1848.
S13539	Broad, T.L., Townsend, C.R., Closs, G.P. & Jellyman, D.J. (2002). Riparian land use and accessibility to fishers influence size class composition and habitat use by longfin eels in a New Zealand river. Journal of Fish Biology, 61, 1489–1503.

Study ID	Reference
S13554	Parkyn, S.M., Collier, K.J. & Hicks, B.J. (2002). Growth and population dynamics of crayfish *Paranephrops planifrons* in streams within native forest and pastoral land uses. New Zealand Journal of Marine and Freshwater Research, 36, 847–862.
S13582	Choudhary, M., Akramkhanov, A. & Saggar, S. (2002). Nitrous oxide emissions from a New Zealand cropped soil: tillage effects, spatial and seasonal variability. Agriculture, ecosystems & environment, 93, 33–43.
S13628	Bellingham, P.J. & Coomes, D.A. (2003). Grazing and community structure as determinants of invasion success by Scotch broom in a New Zealand montane shrubland. Diversity and Distributions, 9, 19–28.
S13746	Riley, R.H., Townsend, C.R., Niyogi, D.K., Arbuckle, C.A. & Peacock, K.A. (2003). Headwater stream response to grassland agricultural development in New Zealand. New Zealand Journal of Marine and Freshwater Research, 37, 389–403.
S13757	Haase, M. (2003). Clinal variation in shell morphology of the freshwater gastropod *Potamopyrgus antipodarum* along two hill-country streams in New Zealand. Journal of the Royal Society of New Zealand, 33, 549–560.
S13792	McLaren, R.G., Clucas, L.M., Taylor, M.D. & Hendry, T. (2003). Leaching of macronutrients and metals from undisturbed soils treated with metal-spiked sewage sludge. 1. Leaching of macronutrients. Australian Journal of Soil Research, 41, 571–588.
S13799	Nyström, P., McIntosh, A.R. & Winterbourn, M.J. (2003). Top-down and bottom-up processes in grassland and forested streams. Oecologia, 136, 596–608.
S13841	Niyogi, D.K., Simon, K.S. & Townsend, C.R. (2003). Breakdown of tussock grass in streams along a gradient of agricultural development in New Zealand. Freshwater Biology, 48, 1698–1708.

Study ID	Reference
S13848	Francis, G.S., Trimmer, L.A., Tregurtha, C.S., Williams, P.H. & Butler, R.C. (2003). Winter nitrate leaching losses from three land uses in the Pukekohe area of New Zealand. New Zealand Journal of Agricultural Research, 46, 215–224.
S13920	Watts, L.F. & Hawke, R.M. (2003). The effects of urbanisation on hydrologic response: A study of two coastal catchments. Journal of Hydrology New Zealand, 42, 125–143.
S13924	Parfitt, R.L., Scott, N.A., Ross, D.J., Salt, G. & Tate, K.R. (2003). Land-use change effects on soil C and N transformations in soils of High N status: comparisions under indigenous forest, pasture, and pine plantation. Biogeochemistry, 66, 203–221.
S13944	McDowell, R.W., Derwry, J.J., Muirhead, R.W. & Paton, R.J. (2003). Cattle treading and phosphorus and sediment loss in overland flow from grazed cropland. Australian Journal of Soil Research, 41, 1521–1532.
S14107	Quinn, J.M., Boothroyd, I.K. & Smith, B.J. (2004). Riparian buffers mitigate effects of pine plantation logging on New Zealand streams: 2. Invertebrate communities. Forest Ecology and Management, 191, 129–146.
S14116	Schipper, L.A. & Lee, W.G. (2004). Microbial biomass, respiration and diversity in ultramafic soils of West Dome, New Zealand. Plant and Soil, 262, 151–158.
S14119	Stevenson, B.A. (2004). Changes in phosphorus availability and nutrient status of indigenous forest fragments in Pastoral New Zealand Hill country. Plant and Soil, 262, 317–325.
S14157	Vink, C.J., Teulon, D.A., McLachlan, A.R. & Stufkens, M.A. (2004). Spiders (*Araneae*) and harvestmen (*Opiliones*) in arable crops and grasses in Canterbury, New Zealand. New Zealand Journal of Zoology, 31, 149–159.
S14172	Boothroyd, I.K., Quinn, J.M., Langer, E.R., Costley, K.J. & Steward, G. (2004).

Study ID	Reference
	Riparian buffers mitigate effects of pine plantation logging on New Zealand streams: 1. Riparian vegetation structure, stream geomorphology and periphyton. Forest Ecology and Management, 194, 199–213.
S14225	McLaren, R.G., Clucas, L.M., Taylor, M.D. & Hendry, T. (2004). Leaching of macronutrients and metals from undisturbed soils treated with metal-spiked sewage sludge. 2. Leaching of metals. Australian Journal of Soil Research, 42, 459–471.
S14266	Harris, R.J., Toft, R.J., Dugdale, J.S., Williams, P.A. & Rees, J.S. (2004). Insect assemblages in a native (kanuka –*Kunzea ericoides*) and an invasive (gorse –*Ulex europaeus*) shrubland. New Zealand Journal of Ecology, 28, 35–47.
S14315	Sparling, G.P. & Schipper, L.A. (2004). Soil quality monitoring in New Zealand: Trends and issues arising from a broad-scale survey. Agriculture, Ecosystems and Environment, 104, 545–552.
S14320	Donnison, A., Ross, C. & Thorrold, B.S. (2004). Impact of land use on the faecal microbial quality of hill-country streams. New Zealand Journal of Marine and Freshwater Research, 38, 845–855.
S14528	Young, R.G., Quarterman, A.J., Eyles, R.F., Smith, R.A. & Bowden, W.B. (2005). Water quality and thermal regime of the Motueka River: Influences of land cover, geology and position in the catchment. New Zealand Journal of Marine and Freshwater Research, 39, 803–825.
S14626	Sullivan, J.J., Timmins, S.M. & Williams, P.A. (2005). Movement of exotic plants into costal native forests from gardens in northern New Zealand. New Zealand Journal of Ecology, 1, 1–10.
S14719	Death, R.G. & Zimmermann, E.M. (2005). Interaction between disturbance and primary productivity in determining stream invertebrate diversity. Oikos, 111, 392–402.
S14732	Leisnham, P.T., Slaney, D.P., Lester, P.J. & Weinstein, P. (2005). Increased

Study ID	Reference
	larval mosquito densities from modified landuses in the Kapiti region, New Zealand: Vegetation, water quality, and predators as associated environmental factors. EcoHealth, 2, 313–322.
S14840	McDowell, R.W. (2006). Phosphorus and Sediment Loss in a Catchment with Winter Forage Grazing of Cropland by Dairy Cattle. Journal of Environment Quality, 35, 575.
S14894	McDowell, R.W. & Stewart, I. (2006). The phosphorus composition of contrasting soils in pastoral, native and forest management in Otago, New Zealand: Sequential extraction and31P NMR. Geoderma, 130, 176–189.
S15119	Harding, J.S., Claassen, K. & Evers, N. (2006). Can forest fragments reset physical and water quality conditions in agricultural catchments and act as refugia for forest stream invertebrates? Hydrobiologia, 568, 391–402.
S15336	Stark, J.D. & Maxted, J.R. (2007). A biotic index for New Zealand's soft-bottomed streams. New Zealand Journal of Marine and Freshwater Research, 41, 43–61.
S15424	Simon, K.S., Niyogi, D.K., Frew, R.D. & Townsend, C.R. (2007).Nitrogen dynamics in grassland streams of agricultural Nitrogen dynamics along a gradient development. Limnology and Oceanography, 52, 1246–1257.
S15451	Ghani, A., Dexter, M., Carran, R.A. & Theobald, P.W. (2007). Dissolved organic nitrogen and carbon in pastoral soils: The New Zealand experience. European Journal of Soil Science, 58, 832–843.
S15500	Quinn, J.M., Phillips, N.R. & Parkyn, S.M. (2007). Factors influencing retention of coarse particulate organic matter in streams. Earth Surface Processes and Landforms, 32, 1186–1203.
S15534	Singh, B.K., Tate, K.R., Kolipaka, G., Hedley, C.B., Macdonald, C.A. & Millard, P. et al. (2007). Effect of afforestation and reforestation of pastures on the activity and population dynamics of methanotrophic bacteria. Applied and

Study ID	Reference
	Environmental Microbiology, 73, 5153–5161.
S15577	Dewson, Z.S., James, A.B.W. & Death, R.G. (2007). Stream ecosystem functioning under reduced flow conditions. Ecological Applications, 17, 1797–1808.
S15817	MacLeod, C.J., Parish, D.M. & Robinson, R.A. (2007). Niche Opportunities and Introduced Birds Temporal Variation in Resource Abundance. In: Temporal dimensions of landscape ecology (eds. Bissonette, J.A. & Storch, I.). Springer US, New York, New York, USA, pp. 252–268.
S15831	Cadbury, S.L., Hannah, D.M., Milner, A.M., Pearson, C.P. & Brown, L.E. (2008). Stream temperature dynamics within a New Zealand glacierized river basin. River Research and Applications, 24, 68–89.
S16020	Dodd, M.B., Quinn, J.M., Thorrold, B.S., Parminter, T.G. & Wedderburn, M.E. (2008). Improving the economic and environmental performance of a New Zealand hill country farm catchment: 3. Short-term outcomes of land-use change. New Zealand Journal of Agricultural Research, 51, 155–169.
S16437	Morgan, D.K., Waas, J.R. & Innes, J. (2009). An inventory of mammalian pests in a New Zealand city. New Zealand Journal of Zoology, 36, 23–33.
S16511	Quinn, J.M., Croker, G.F., Smith, B.J. & Bellingham, M.A. (2009). Integrated catchment management effects on flow, habitat, instream vegetation and macroinvertebrates in Waikato, New Zealand, hill-country streams. New Zealand Journal of Marine and Freshwater Research, 43, 775–802.
S16849	Carswell, F., Burrows, L. & Mason, N. (2009). Above-ground carbon sequestration by early-successional woody vegetation: A preliminary analysis - Science for Conservation 297. Department of Conservation, Wellington.
S16935	Zheng, T., Hamel, K.L. & Buddle, B.M. (2010). A serological survey of the prevalence of antibodies against enteroviruses in brushtail possums

Study ID	Reference
	(*Trichosurus vulpecula*) in New Zealand: Enteroviruses have a limited distribution in brushtail possums. New Zealand Veterinary Journal, 58, 23–28.
S16958	McKergow, L.A., Pritchard, M., Elliott, A.H., Duncan, M.J. & Senior, A.K. (2010). Storm fine sediment flux from catchment to estuary, Waitetuna-Raglan Harbour, New Zealand. New Zealand Journal of Marine and Freshwater Research, 44, 53–76.
S17210	Price, S., Whitehead, D., Sherlock, R., McSeveny, T. & Rogers, G. (2010). Net exchange of greenhouse gases from soils in an unimproved pasture and regenerating indigenous *Kunzea ericoides* shrubland in New Zealand. Australian Journal of Soil Research, 48, 385–394.
S17301	Reid, D.J., Quinn, J.M. & Wright-Stow, A.E. (2010). Responses of stream macroinvertebrate communities to progressive forest harvesting: Influences of harvest intensity, stream size and riparian buffers. Forest Ecology and Management, 260, 1804–1815.
S17776	Shearer, K.A. & Young, R.G. (2011). Influences of geology and land use on macroinvertebrate communities across the Motueka River catchment, New Zealand. New Zealand Journal of Marine and Freshwater Research, 45, 437–454.
S18250	Knox, C.D., Cree, A. & Seddon, P.J. (2012). Direct and Indirect Effects of Grazing by Introduced Mammals on a Native, Arboreal Gecko (*Naultinus gemmeus*). Journal of Herpetology, 46, 145–152.
S18389	Tate, K.R., Walcroft, A.S. & Pratt, C. (2012). Varying atmospheric methane concentrations affect soil methane oxidation rates and methanotroph populations in pasture, an adjacent pine forest, and a landfill. Soil Biology and Biochemistry, 52, 75–81.
S18515	Hewitt, A., Forrester, G., Fraser, S., Hedley, C.B., Lynn, I. & Payton, I. (2012). Afforestation effects on soil carbon stocks of low productivity grassland in

Study ID	Reference
	New Zealand. Soil Use and Management, 28, 508–516.
S18879	Fraser, P.M., Curtin, D., Harrison-Kirk, T., Meenken, E.D., Beare, M.H. & Tabley, F.J. et al. (2013). Winter Nitrate Leaching under Different Tillage and Winter Cover Crop Management Practices. Soil Science Society of America Journal, 77, 1391.
S19160	Wakelin, S.A., Macdonald, L.M., O'Callaghan, M., Forrester, S.T. & Condron, L.M. (2014). Soil functional resistance and stability are linked to different ecosystem properties. Austral Ecology, 39, 522–531.
S19193	Waterhouse, B.R., Adair, K.L., Boyer, S. & Wratten, S.D. (2014). Advanced mine restoration protocols facilitate early recovery of soil microbial biomass, activity and functional diversity. Basic and Applied Ecology, 15, 599–606.
S19214	Rader, R., Bartomeus, I., Tylianakis, J.M. & Laliberté, E. (2014). The winners and losers of land use intensification: Pollinator community disassembly is non-random and alters functional diversity. Diversity and Distributions, 20, 908–917.
S19282	Miller, A.L., Diez, J.M., Sullivan, J.J., Wangen, S.R., Wiser, S.K. & Meffin, R. et al. (2014). Quantifying invasion resistance: the use of recruitment functions to control for propagule pressure. Ecology, 95, 920–929.
S19506	Rahman, M.H., Holmes, A.W. & Saunders, S.J. (2014). Spatio-temporal variation in soil organic carbon under kiwifruit production systems of New Zealand. Acta Horticulturae, 1018, 279–286.
S19543	Sparling, G.P., Lewis, R., Schipper, L.A., Mudge, P. & Balks, M. (2014). Changes in soil total C and N contents at three chronosequences after conversion from plantation pine forest to dairy pasture on a New Zealand Pumice soil. Soil Research, 52, 38–45.
S19544	Mudge, P., Schipper, L.A., Baisden, W., Ghani, A. & Lewis, R. (2014). Changes

Study ID	Reference
	in soil C, N and 15N along three forest-pasture chronosequences in New Zealand. Soil Research, 52, 27–37.
S19646	Wilson, D.J., Norbury, G. & Walker, S. (2014). How does woody succession affect population densities of passerine birds in New Zealand drylands? New Zealand Journal of Ecology, 38, 257–267.
S20016	Chen, C.R., Condron, L.M., Davis, M.R. & Sherlock, R. (2003). Seasonal changes in soil phosphorus and associated microbial properties under adjacent grassland and forest in New Zealand. Forest Ecology and Management, 177, 539–557.
S20029	Davis, M.R. & Lang, M.H. (1991). Increased nutrient availability in topsoils under conifers in the South Island high country. New Zealand Journal of Forestry Science, 21, 165–179.
S20058	Kasai, M., Brierley, G.J., Page, M.J., Marutani, T. & Trustrum, N.A. (2005). Impacts of land use change on patterns of sediment flux in Weraamaia catchment, New Zealand. Catena, 64, 27–60.
S20084	Oliver, G.R., Beets, P.N., Garrett, L.G., Pearce, S.H., Kimberly, M.O. & Ford-Robertson, J. et al. (2004). Variation in soil carbon in pine plantations and implications for monitoring soil carbon stocks in relation to land-use change and forest site management in New Zealand. Forest Ecology and Management, 203, 283–295.
S20087	Parfitt, R.L., Percival, H.J., Dahlgren, R.A. & Hill, L.F. (1997). Soil and solution chemistry under pasture and radiata pine in New Zealand. Plant and Soil, 191, 279–290.
S20088	Parfitt, R.L. & Ross, D.J. (2011). Long-term effects of afforestation with *Pinus radiata* on soil carbon, nitrogen, and pH: A case study. Soil Research, 49, 494–503.
S20106	Ross, D.J., Tate, K.R., Scott, N.A., Wilde, R.H., Rodda, N.J. & Townsend, J.A.

Study ID	Reference
	(2002). Afforestation of pastures with *Pinus radiata* influences soil carbon and nitrogen pools and mineralisation and microbial properties. Australian Journal of Soil Research, 40, 1303–1318.
S20108	Rowe, D.K., Chisnall, B.L., Dean, T.L. & Richardson, J. (1999). Effects of land use on native fish communities in east coast streams of the North Island of New Zealand. New Zealand Journal of Marine and Freshwater Research, 33, 141–151.
S20128	Townsend, C.R., Downes, B.J., Peacock, K.A. & Arbuckle, C.J. (2004). Scale and the detection of land-use effects on morphology, vegetation and macroinvertebrate communities of grassland streams. Freshwater Biology, 49, 448–462.
S20129	Townsend, C.R., Thompson, R.M., McIntosh, A.R., Kilroy, C., Edwards, E. & Scarsbrook, M.R. (1998). Disturbance, resource supply, and food web architecture in streams. Ecology Letters, 1, 200–209

Appendix 6. Conversion of confidence intervals to variance and imputation of missing values

Conversion of 95% confidence intervals to variance

$$s^2 = \frac{95\%CI}{t - critical} \times \sqrt{n}$$

Where s^2 denotes the variance, 95%CI corresponds to the 95% confidence interval and n is the sample size reported by the authors. The t-critical value is the value in the t-distribution for the corresponding alpha and degrees of freedom. For all studies in our dataset where we needed to apply this conversion, the degrees of freedom where not available so we approximated the t-critical value as two. This was done because the two-sided t-distribution values for an alpha of 0.05 range between 2.57 and asymptopte at 1.96 for 5 or more degrees of freedom.

Regression model for applying Taylor's Law

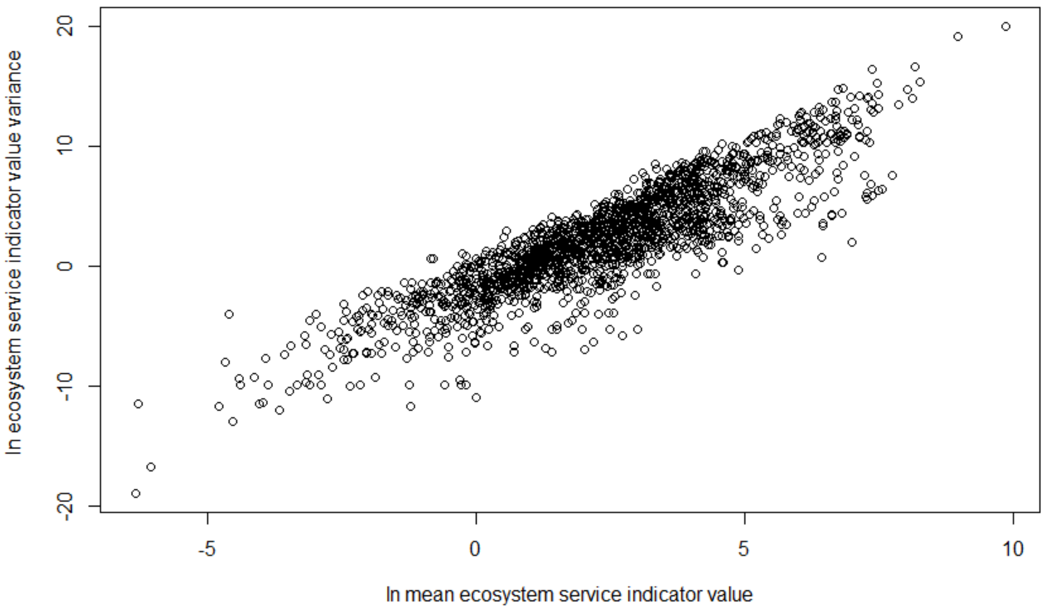


Fig. S2. *Regression (in natural logarithm space) of the mean ecosystem service indicator values for all land covers reported in our dataset into their corresponding variances.*

The equation used to impute variances from means based on regression coefficients from the relation between both in natural logarithm space is shown below. Coefficients are given to three decimal places in the equation, however their full values were used in the actual calculation of the imputed values.

$$s_{imp}^2 = e^{[-2.147 + (1.878 \times \ln(|x|)]}$$

Where s_{imp}^2 is the imputed variance and x the mean used for the imputation.

Appendix 7. Summary of log Response Ratios per land cover and ecosystem service combination.

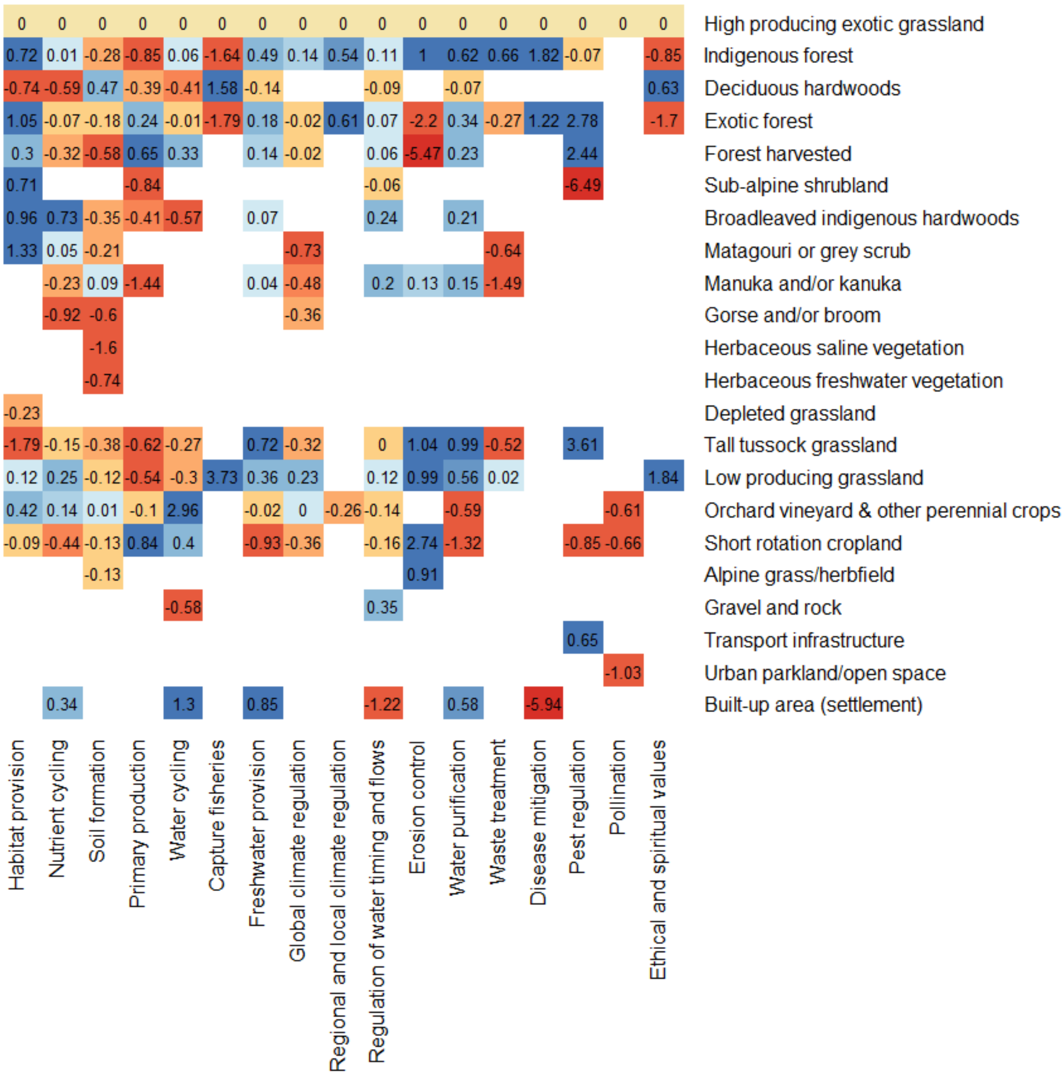


Fig. S3. *Aggregated log Response Ratios of ecosystem service provision across land covers. Values are given relative to the high producing exotic grassland reference. Red-orange tones highlight cases where the land covers perform comparatively worse than the reference in the provision of a service, while blue tones signal land covers that perform comparatively better. The darker the blue or red-orange tone, the greater the ratio separating the land cover to the reference in the provision of the corresponding service.*

Appendix 8. Classification of land covers and ecosystem services for PERMANOVA analysis

Table S4. Delimitation of categorical variables used in PERMANOVA of land cover effects across ecosystem services

Land cover	Forest Cover	Production	Type of vegetation cover
Indigenous forest	Present	No	Native
Deciduous hardwoods	Present	No	Exotic
Exotic forest	Present	Yes	Exotic
Forest harvested	Present	Yes	Exotic
Broadleaved indigenous hardwoods	Present	No	Native
Tall tussock grassland	Absent	No	Native
Low producing grassland	Absent	No	Native
High producing exotic grassland	Absent	Yes	Exotic
Orchard vineyard & other perennial crops	Absent	Yes	Exotic
Short rotation cropland	Absent	Yes	Exotic

Vegetation cover was not included in the analysis since all but one of land cover groups in the native / exotic groups overlap with those in the production / non-production groups.

Table S5. Delimitation of categorical variables used in PERMANOVA of ecosystem service provision across land covers

Ecosystem service	Scale of benefits	Biophysical domain
Habitat provision	Regional	Biotic
Nutrient cycling	Local	Edaphic
Soil formation	Local	Edaphic

Primary production	Local	Biotic
Water cycling	Global	Hydrologic
Freshwater provision	Regional	Hydrologic
Global climate regulation	Global	Edaphic
Regulation of water timing and flows	Regional	Hydrologic
Water purification	Regional	Hydrologic

Appendix 9. Overview of research effort for New Zealand

Fig. S4 shows that all of the supporting ecosystem services are represented in our database, whereas for the remaining categories our data only offer partial coverage of their services, with information on: nine out of the 11 regulating services, two of the 15 provisioning services and one of the three cultural services defined for the project. For the categories that were represented by more than one ecosystem service in our database, the number of studies per service ranged from two to 54 for the regulating services; from five to 44 for the provisioning ones and from 29 to 60 in the supporting services. A total of five studies provide evidence for the single service in the cultural category in our database.

	Habitat provision	Nutrient cycling	Soil formation	Primary production	Water cycling	Capture fisheries	Freshwater provision	Global climate regulation	Regional and local climate regulation	Regulation of water timing and flows	Erosion control	Water purification	Waste treatment	Disease mitigation	Pest regulation	Pollination	Ethical and spiritual values	Total
Indigenous forest	30	22	11	19	16	3	26	7	1	29	13	22	2	4	6	0	3	58
Deciduous hardwoods	2	1	1	1	1	1	2	0	0	1	0	1	0	0	0	0	1	2
Exotic forest	23	28	28	13	14	2	20	21	2	31	7	17	5	2	3	0	2	62
Forest harvested	6	3	2	3	4	0	5	1	0	7	3	5	0	0	2	0	0	12
Mixed exotic shrubland	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Sub-alpine shrubland	1	1	0	1	0	0	0	0	0	1	0	0	0	0	2	0	0	4
Broadleaved indigenous hardwoods	4	3	2	4	2	0	2	0	0	2	0	2	0	0	2	1	0	9
Matagouri or grey scrub	1	1	1	0	0	0	0	1	0	0	0	0	1	0	2	0	0	4
Manuka and/or kanuka	1	4	5	1	0	0	1	3	0	1	2	1	1	0	2	0	0	10
Gorse and/or broom	0	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	2
Flaxland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Herbaceous saline vegetation	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Herbaceous freshwater vegetation	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Depleted grassland	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
Tall tussock grassland	9	12	7	9	6	0	10	1	0	7	4	10	2	0	2	0	0	19
Low producing grassland	7	4	7	3	4	2	5	3	0	7	2	4	1	0	1	0	3	18
High producing exotic grassland	36	45	37	22	21	5	33	29	2	40	15	28	5	4	2	2	4	91
Orchard vineyard & other perennial crops	2	2	3	1	1	0	2	3	1	1	0	1	0	0	1	1	0	6
Short rotation cropland	6	17	17	2	2	0	4	12	0	9	4	6	0	0	1	1	0	24
Alpine grass/herbfield	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Gravel and rock	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	2
Transport infrastructure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Surface mine & dump	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
Urban parkland/open space	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	4
Built-up area (settlement)	0	1	0	0	1	0	2	0	0	1	0	1	0	1	1	0	0	3
Total	50	60	51	32	29	5	44	33	2	54	22	40	7	4	13	3	5	136

Fig. S4. ***Fig. S4.*** *Distribution of studies per ecosystem service and land cover. For most services, data are concentrated along a selection of land covers: high producing exotic grassland (with a total of 92 studies across all ecosystem services), exotic forest (64 studies in total), indigenous forest (58 studies), short-rotation cropland (24 studies) and tall tussock grassland (20 studies). In addition, eight of the 43 land cover classes in the LCDB classification were not present in our data base. These land cover classes were: "sand, gravel and rock" (i.e. the coastal strip separating land from sea), "mangrove", "fernland", "landslide", "permanent snow and ice","lake or pond", "river" and "estuarine open water". The last three units correspond to aquatic land covers which were not included in our review, whereas the remainder were simply poorly represented within the literature used for our review. Note that the total row* *and column don't match the actual sum of column and row counts because our dataset includes studies with data on multiple ecosystems and land covers. Likewise, the row and column totals do not add up to the grand total in the lower right corner which, instead, corresponds to the total number of studies in our dataset.*

Since we aggregated data from studies with multiple indicators of the same service, the matrix in Fig. S4 effectively reflects the number of data points in our spreadsheet for each ecosystem service and land cover combination. The actual number of indicators for each ecosystem service - land cover combination are shown in Fig. S5 which indicates that, overall, the number of indicators follow a similar distribution to that of the number of studies. However, for the most common ecosystem service-land cover combinations in our dataset (e.g., soil formation or nutrient cycling in both exotic forest and high producing exotic grasslands) there were as many as four to five times more indicators than studies, suggesting that studies with multiple indicators were more frequent in the land covers and ecosystem services that were also more commonly studied.

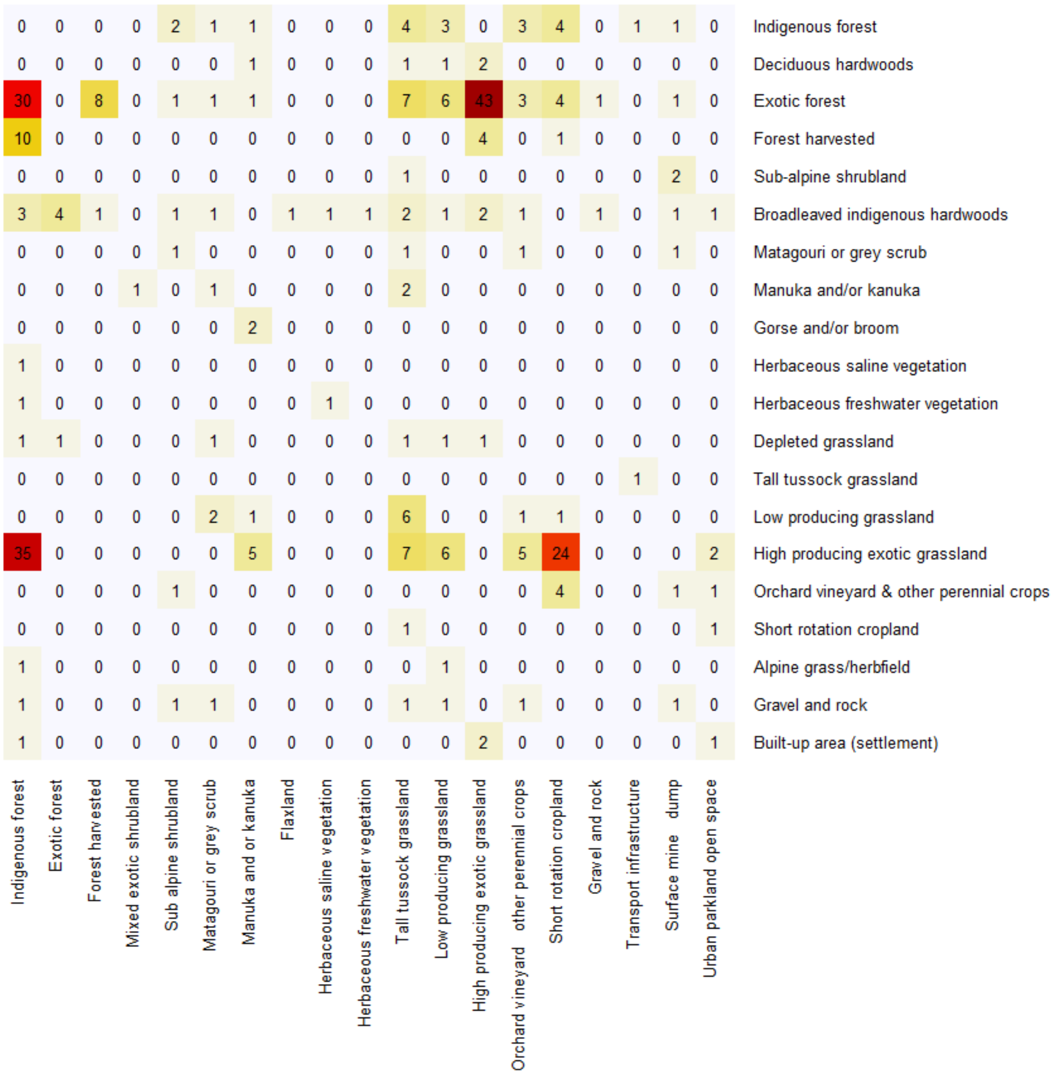


Fig. S5. *Distribution of studies per land cover comparisons. Studies contributing data on multiple ecosystem services are only counted once in each pair of land covers where they contribute data.*

Appendix 10. Evidence base and network meta-analysis for individual ecosystem services

Regulation of water timing and flows

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 36 indicators. The main aspects quantified by these indicators pertain to soil characteristics that either provide greater regulation by enhancing soil water retention or have detrimental effects in the provision of this service by promoting increased runoff. In addition, there are also some indicators on stream channel characteristics (such as its dimensions) that affect its ability to regulate water flow over time and that, to an extent, can be altered by land cover.

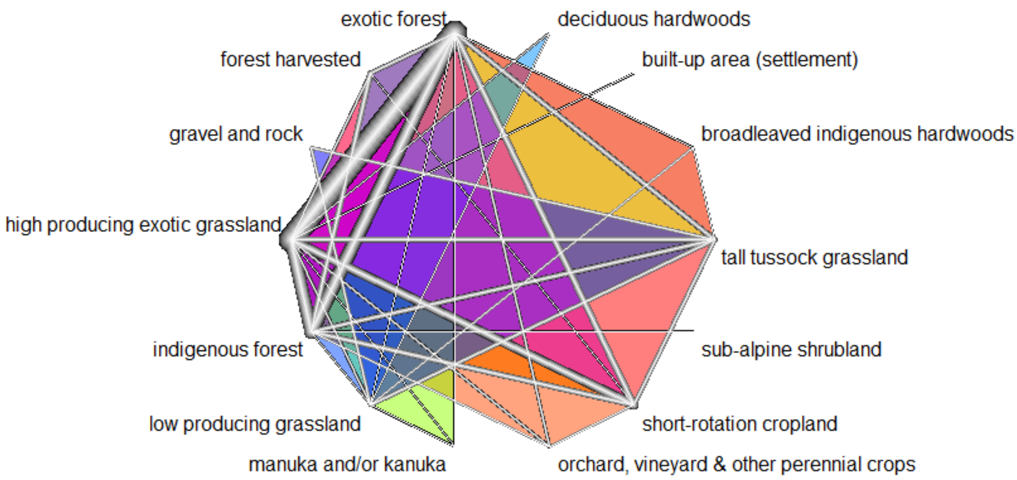


Fig. S6. *Evidence network for land cover comparisons on regulation of water timing and flows. In all evidence network graphs, lines connect pairs of land covers that are compared in one or more studies and their thickness is inversely proportional to the standard error for the comparison, with thicker lines indicating smaller standard errors and, consequently, a greater evidence for the comparison. Shaded areas indicate the presence of multi-arm studies which compare three or more land covers.*

Evidence base: This evidence network is formed by 126 pairwise comparisons of 14 land covers. Data were obtained from 54 different studies, each contributing a minimum of one and a maximum of four pairwise land cover comparisons. As indicated by the thicker lines in Fig. S6, the land covers that are most commonly compared are:

- High producing exotic grassland (64 comparisons)
- Exotic forest (58 comparisons)
- Indigenous forest (50 comparisons)
- Short-rotation cropland (19 comparisons)

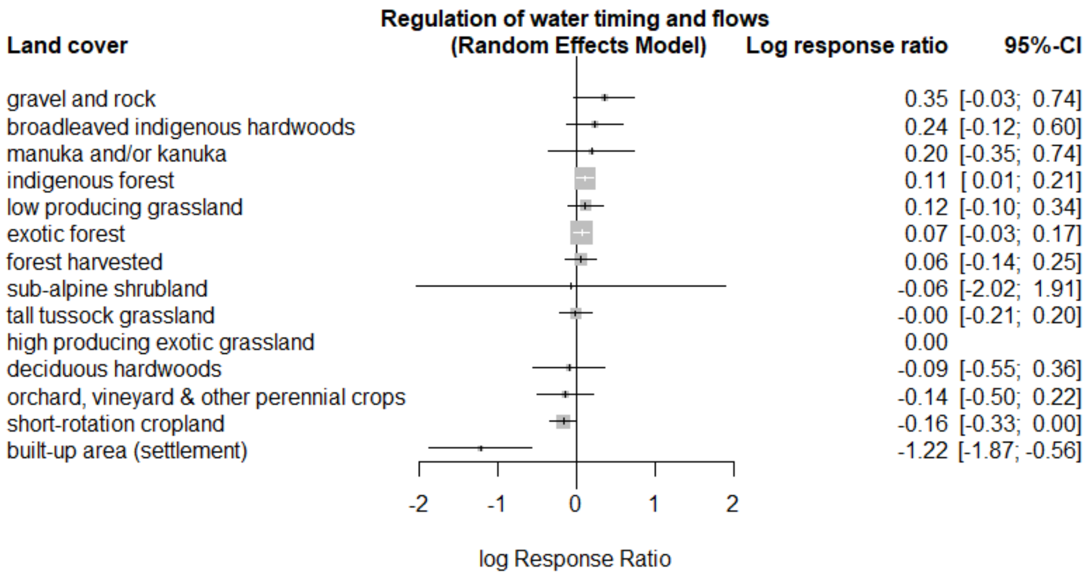


Fig. S7. *Forest plot of land cover contrasts in the provision of regulation of water timing and flows. Random effects model with high producing exotic grassland as a reference. Log Response Ratios depicted here are the network meta-analysis model estimates of the overall ratios between each land cover and high producing exotic grassland. The model accounts for the direct and indirect comparisons in the evidence network, as well as the random effects from having comparisons on the same land covers drawn from different studies. Bars indicate the 95% confidence intervals for each estimate while grey boxes reflect the relative weight of the comparison between each land cover and high producing exotic grassland in the overall* *model estimates. Comparisons that have greater weights are depicted with larger boxes. Land covers are presented in descending order of their P-Scores which are calculated from the magnitude and precision of the log Response Ratio estimates for each land cover and provide a means to rank treatment effects (i.e. land covers) according to their comparative effectiveness (6)*

Measures of heterogeneity/network inconsistency:

τ^2 = 0.165

I^2 = 53.452

Main results:

- Overall there is a gradient from native vegetation (manuka/kanuka, broadleaved hardwoods and indigenous forest) to more artificial and production-oriented land covers (high producing exotic grassland, orchard, vineyard and perennials, harvested forest).
- Cropland, exotic forest both seem to behave similarly to indigenous forest, as do broadleaved indigenous hardwoods, manuka an/kanuka, and low producing grassland.
- Built-up area stands out as the worst performing land cover in terms of on regulation of water timing and flows, which is likely explained by the presence of impervious surfaces and channel morphologies that enhance runoff.
- The high infiltration capacity in gravel and rock probably accounts for its high ranking in the provision of this service.

Nutrient cycling

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 161 indicators, most of which focus on the cycling and flow of nutrients within the soil system and characteristics of the soil environment that promote or hinder nutrient cycling. The latter were taken as negative indicators for the provision of nutrient cycling, as were the indicators on nutrient loss from the soil system. In addition, the data also include indicators on how land cover conditions plant uptake and the processing of nutrients both in the soil and freshwater systems. A large number of the indicators pertain to nitrogen and phosphorus however there is also information on other nutrients including: calcium, carbon, chlorine, copper, magnesium, potassium, sodium, sulfur and zinc (we have followed the Millennium Ecosystem Assessment (7) in their delimitation of the nutrients for this service as those relevant for plant growth).

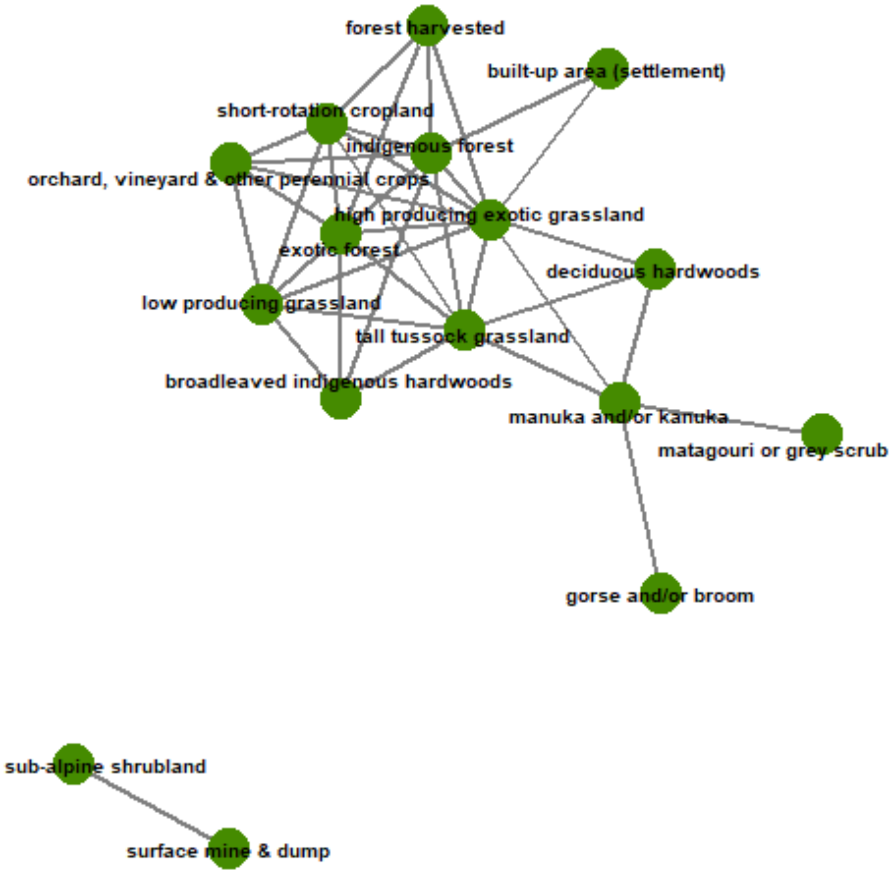


Fig. S8. *Land-cover comparison networks for nutrient cycling.*

Evidence base: The evidence base for this service is split into two networks of land cover comparisons depicted in Fig. S8. The smaller of these networks holds the comparison between sub-alpine shrubland and surface mine & dump, for which there is only evidence from a single study and, consequently, are not connected to any of the land covers in the larger network. In the smaller network the single study evidence defines a log Response Ratio of approximately 1.435 in favor of the sub-alpine shrubland over the surface mine & dump (the standard error of this estimate is approximately 0.054). In what follows we focus exclusively on the evidence base and network meta-analysis for the larger network of land covers in this service.

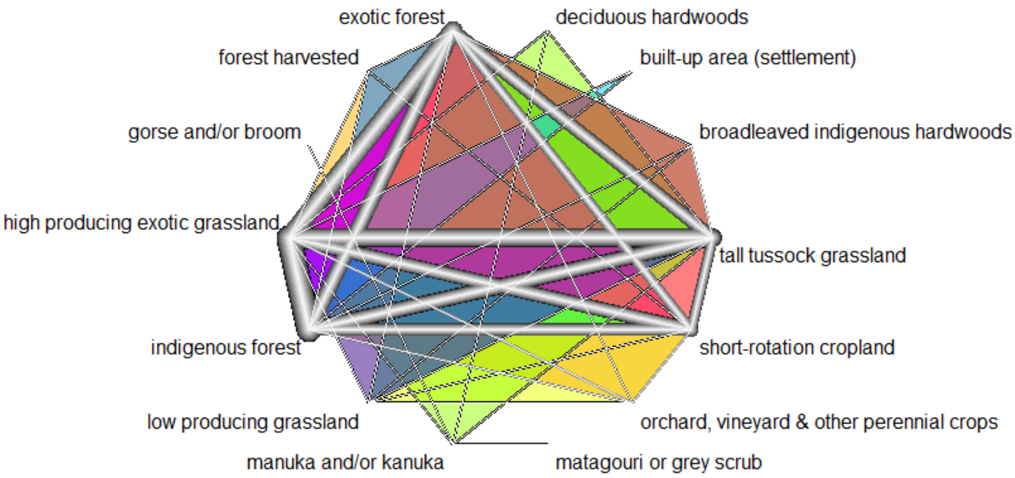


Fig. S9. *Evidence network for land cover comparisons on nutrient cycling.*

This evidence network is formed by 123 pairwise comparisons of 14 land covers. Data were obtained from 59 different studies, each contributing a minimum of one and a maximum of four pairwise land cover comparisons. As indicated by the thicker lines in Fig. S9, the land covers that are most commonly compared are:

- High producing exotic grassland (67 comparisons)

- Exotic forest (47 comparisons)
- Indigenous forest (39 comparisons)
- Short-rotation cropland (30 comparisons)

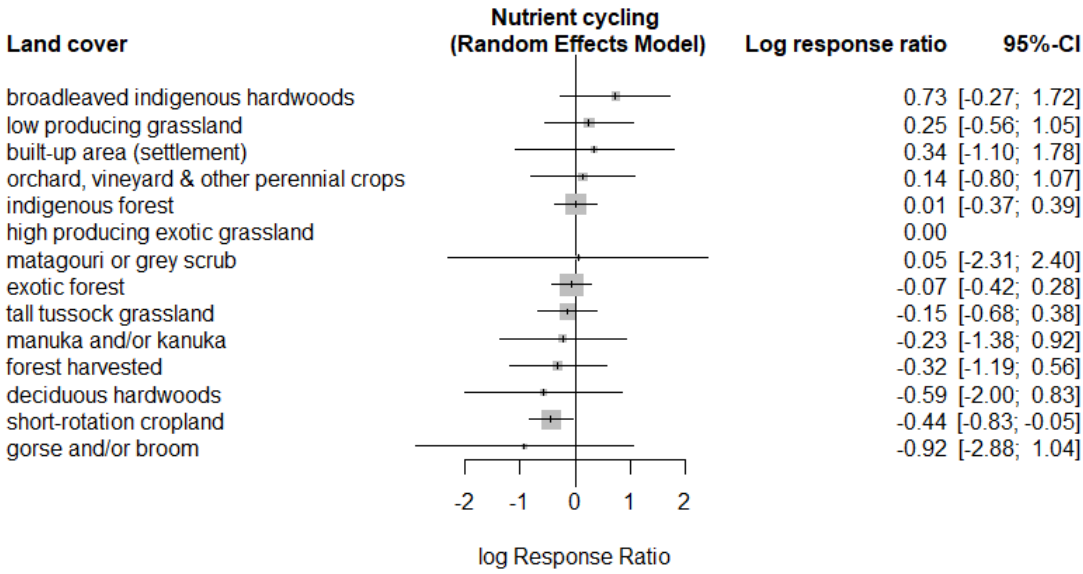


Fig. S10. *Forest plot of land cover contrasts in the provision of nutrient cycling.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 0.802$

$I^2 = 96.271$

Main results:

- With the exception of short-rotation cropland, the confidence intervals for most land covers overlap the high producing exotic grassland reference. Moreover, exotic forests and orchards, vineyards & other perennial crops also exhibit very small effect estimates, suggesting they may share similar nutrient cycling dynamics to those found in high producing exotic grasslands with artificial nutrient enrichment inducing more dynamic processing rates in the soil system (8). On the contrary, short-rotation croplands and other land covers dominated by exotic species but lacking the artificial enrichment (forest harvested, deciduous hardwoods and gorse and/broom) perform worse in the provision of this service than the reference cover.

- The wider confidence intervals in the forest plot shown in Fig. S10 correspond to the land covers that had the least number of direct comparisons within the evidence network, while the land covers with narrower intervals are the ones that were informed by the greatest number of comparisons.

Habitat provision

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 80 indicators which, for the most part, expressed aspects relating to the availability of resources and/or conditions favorable to wildlife within a land cover, habitat occupation or use by native fauna and the health of native animal species within a given land cover.

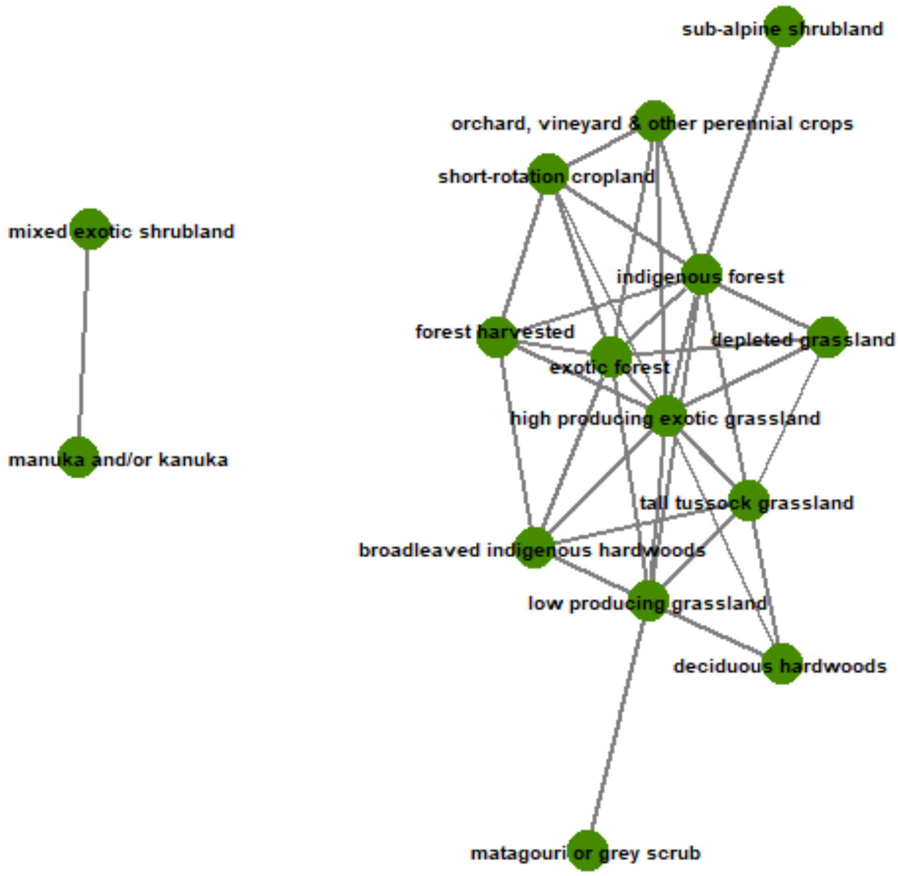


Fig. S11. *Land-cover comparison networks for habitat provision.*

Evidence base: As shown in Fig. S11, there are two networks connecting the land cover comparisons for this service. The smaller of these networks encompasses the mixed exotic shrublands and manuka/kanuka (which are compared only in one study for this service), while the remaining land covers are connected in the larger network. Evidence for the smaller network suggests manuka and/or kanuka is marginally better than mixed exotic shrubland in the provision of habitat (with a log response ratio of approximately 0.025 for manuka and/or kanuka over the mixed exotic shrubland, and a standard error of approximately 0.079 in this estimate). Below we present the evidence base and network meta-analysis for the larger network of land covers in this service.

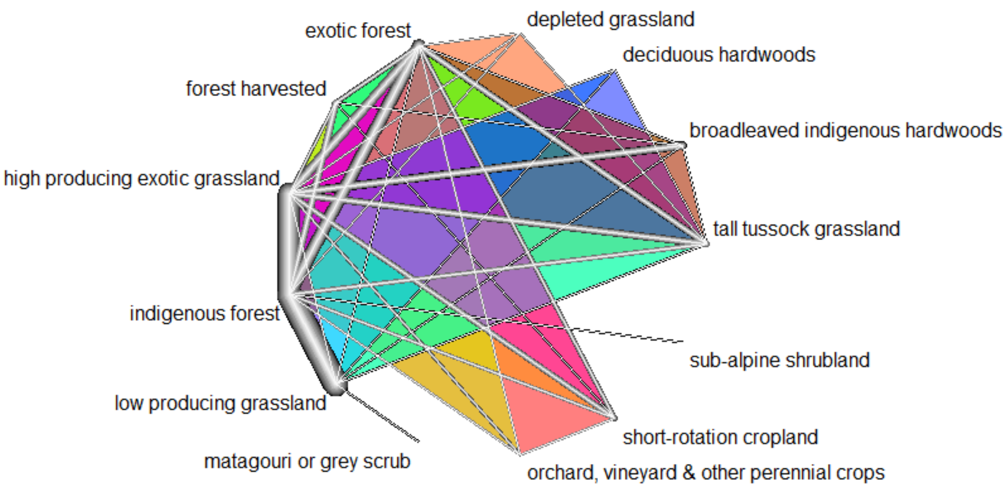


Fig. S12. *Evidence network for land cover comparisons on habitat provision.*

This evidence network is formed by 122 pairwise comparisons of 13 land covers. Data were obtained from 49 different studies, each contributing a minimum of one and a maximum of four pairwise land cover comparisons. As indicated by the thicker lines in Fig. S12, the land covers that are most commonly compared are:

- High producing exotic grassland (61 comparisons)
- Indigenous forest (51 comparisons)
- Exotic forest (49 comparisons)
- Tall tussock grassland (17 comparisons)

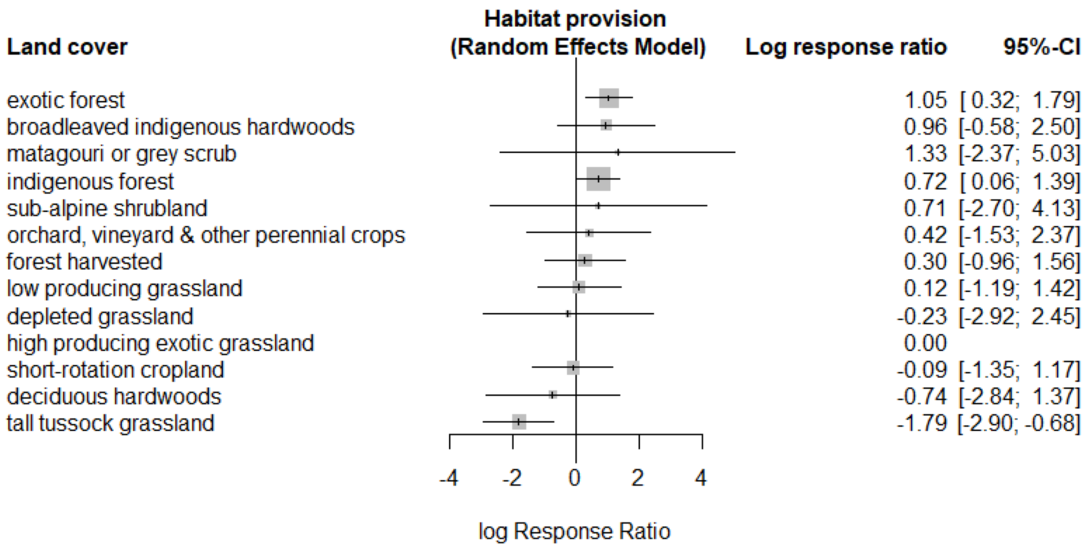


Fig. S13. *Forest plot of land cover contrasts in habitat provision. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 1.699

I^2 = 99.552

Main results:

- Exotic and native forests are both significantly better than high producing exotic grassland in providing habitat and, although non-significant, the exotic forest ranks slightly higher than the native one in delivering this service.
- Tall tussock grasslands rank poorly and are significantly worse than both exotic and native forests and high producing exotic grasslands in the provision of habitat.
- All croplands and grasslands (low, high and depleted) perform similarly in the provision of this service and, overall, rank below the forest and native shrublands.

Soil formation

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 111 different indicators that cover aspects such as: soil aggradation and degradation processes (the latter having a negative effect on soil formation), soil structure and stability, the availability of nutrients and favorable conditions for plant growth in the soil.

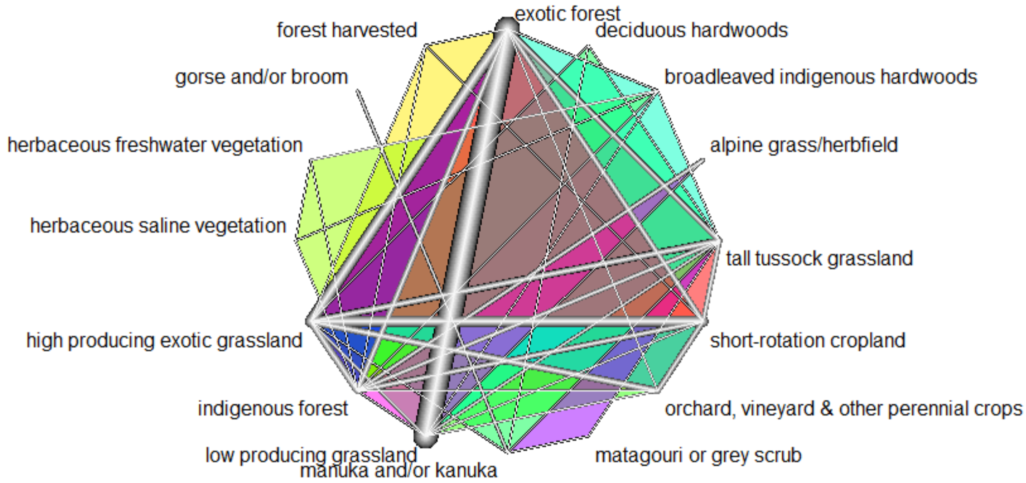


Fig. S14. *Evidence network for land cover comparisons on soil formation.*

Evidence base: This evidence network is formed by 110 pairwise comparisons of 16 land covers. Data were obtained from 51 different studies, each contributing a minimum of one and a maximum of four pairwise land cover comparisons. As indicated by the thicker lines in Fig. S14, the land covers that are most commonly compared are:

- High producing exotic grassland (54 comparisons)
- Exotic forest (42 comparisons)

- Short-rotation cropland (30 comparisons)
- Indigenous forest (25 comparisons)

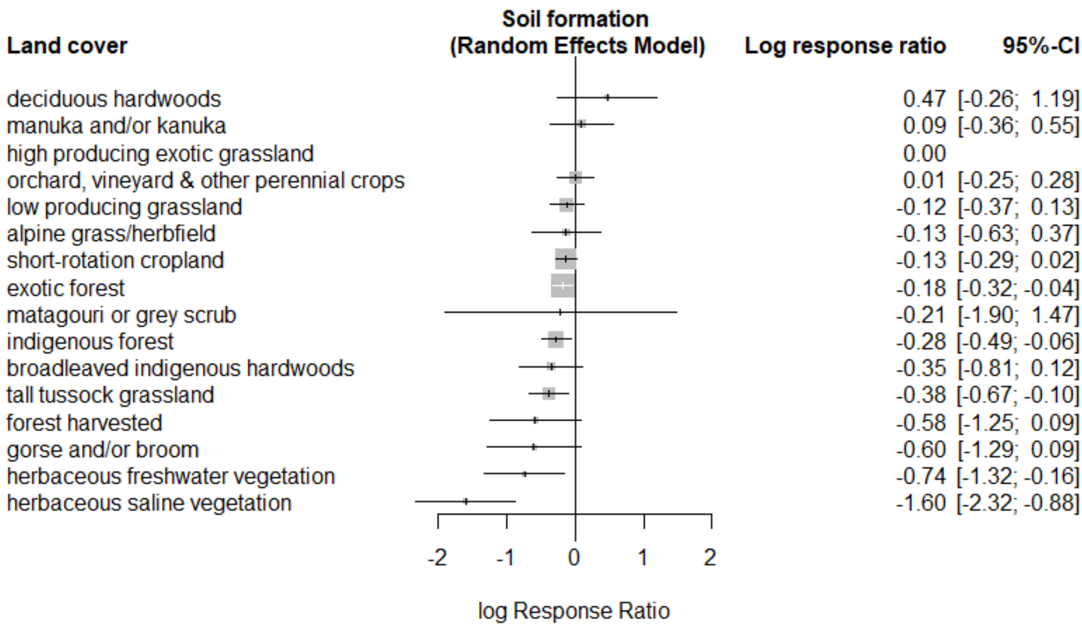


Fig. S15. *Forest plot of land cover contrasts in the provision of soil formation. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 0.241$

$I^2 = 67.163$

Main results:

- No land cover is significantly better than the high producing exotic grassland in promoting soil formation however, the ratio estimate for deciduous hardwoods does place them a bit above the reference in the provision of this service.

- High producing exotic grasslands rank well in this service, this is likely a result of both the high artificial nutrient inputs (that result in a greater nutrient availability for plants and, by our accounting, increased soil formation) and the fact that this land cover tends to be found in areas where the soils are well developed and very favorable to plant growth. The small (and unsignificant) differences between high producing exotic grassland and both cropland covers (short rotation and the orchard, vineyard & other perennial crops) could be explained by similar factors.

- Except for manuka and/or kanuka, all native land covers rank similarly in the provision of this service. Indigenous forests and tall tussock grasslands are significantly worse than the reference land cover, while broadleaved indigenous hardwoods and matagouri also do worse than the reference but not significantly so.
- Gorse and/or broom and forest harvested also rank below the reference land cover in delivering this service, but their wide confidence intervals make these differences statistically non-significant. Wide confidence intervals also apply to herbaceous freshwater and saline vegetation which, nevertheless, still perform significantly worse than the reference and many of the other land covers. This makes sense given how they are prone to influxes of water that prevent soil forming processes.

Freshwater provision

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 71 indicators all of which expressed a measure of land cover effects on the quantity or quality of freshwater provided by streams. Indicators on the quality of water draw mostly on measures from concentrations of nutrients commonly linked to euthrophication (namely nitrogen and phosphorus), sediments and fecal contamination.

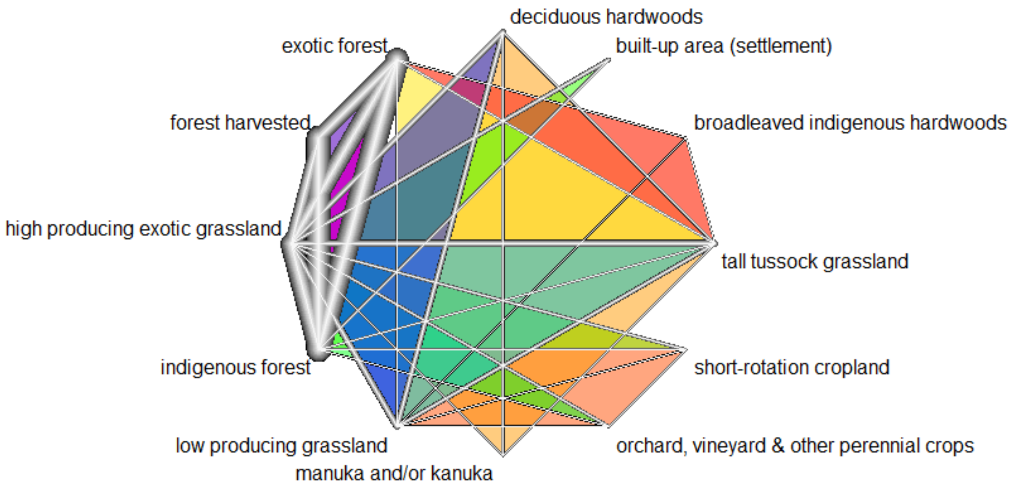


Fig. S16. *Evidence network for land cover comparisons on freshwater provision.*

Evidence base: This evidence network is formed by 96 pairwise comparisons of 12 land covers. Data were obtained from 44 different studies, each contributing a minimum of one and a maximum of three pairwise land cover comparisons. As indicated by the thicker lines in Fig. S16, the land covers that are most commonly compared are:

- High producing exotic grassland (52 comparisons)
- Indigenous forest (40 comparisons)
- Exotic forest (35 comparisons)
- Tall tussock grassland (16 comparisons)

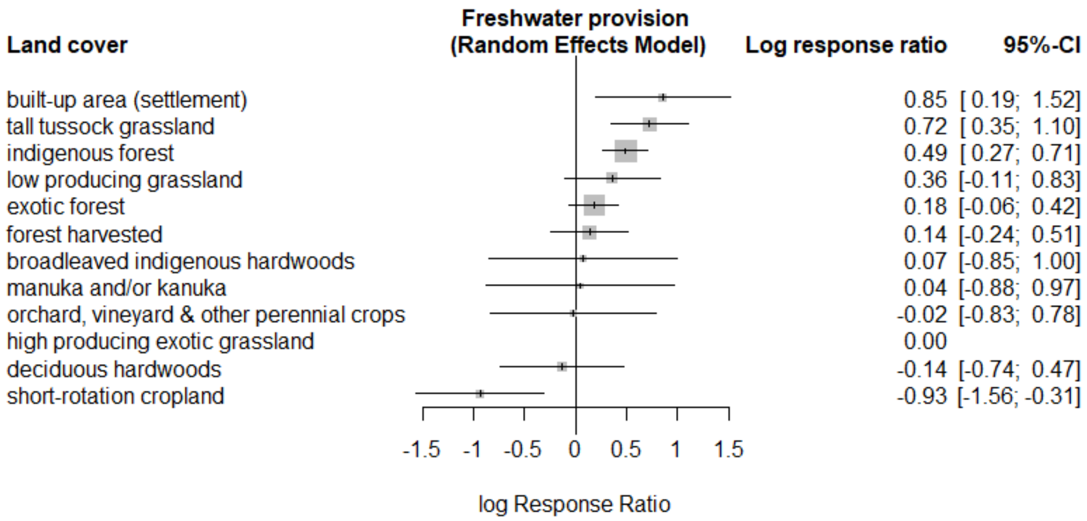


Fig. S17. *Forest plot of land cover contrasts on freshwater provision. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 0.423$

$I^2 = 81.514$

Main results:

- It is striking to find built-up area as the highest ranking land cover in providing this service. However the log response estimate is bounded by wide confidence intervals. For this specific land cover we only had information on how it compares to high producing exotic grassland in terms of specific stream flow, where it exceeds the grassland by 4 times. So the effect we see here could be an artifact of this single data point.
- Tussock grasslands and indigenous forests both perform significantly better than the high producing exotic grassland in providing freshwater. While low producing grasslands also tended to do better than the reference but not significantly so.

- Exotic and harvested forests have very similar rankings, performing slightly better than the reference but not in a significant way. Likewise, the deciduous hardwoods, tend to do worse than the reference in providing freshwater but the difference is not significant.
- Short-rotation cropland performs poorly in the provision of this service and has significant differences not only with the reference land cover, but also with some of the ones that rank high in its provision (built-up area, tall tussock and low producing grasslands, exotic, harvested and indigenous forests).
- For the remainder of the land covers, the confidence intervals are wide and intersect those of all the other land covers.

Water purification

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 78 indicators that provide a measure of either: the filtering of pollutants and excess nutrients from the water (either in the soil or in aquatic systems, for processes that are affected by land cover), or the accumulation of pollutants and toxic substances in freshwater.

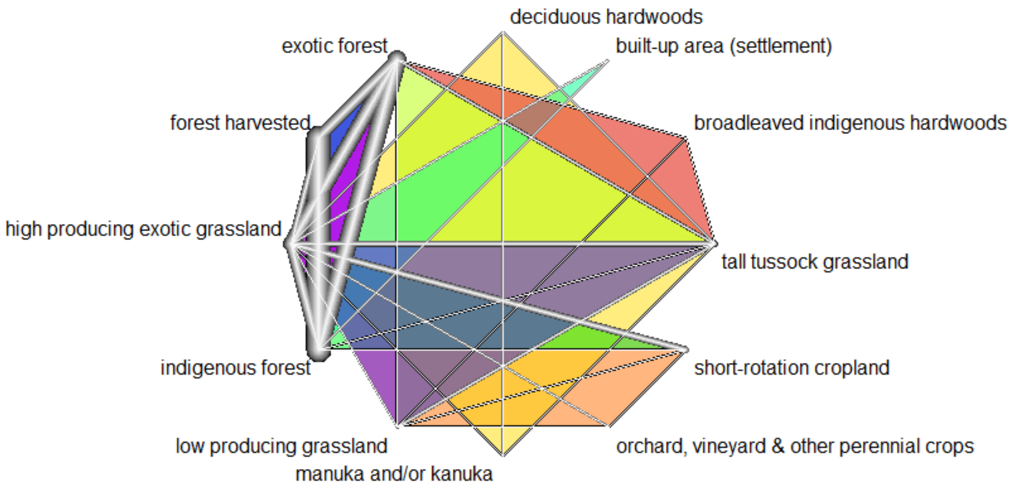


Fig. S18. *Evidence network for land cover comparisons on water purification.*

Evidence base: This evidence network is formed by 80 pairwise comparisons of 12 land covers. Data were obtained from 40 different studies, each contributing a minimum of one and a maximum of three pairwise land cover comparisons. As indicated by the thicker lines in Fig. S18, the land covers that are most commonly compared are:

- High producing exotic grassland (41 comparisons)
- Indigenous forest (32 comparisons)
- Exotic forest (28 comparisons)
- Tall tussock grassland (16 comparisons)

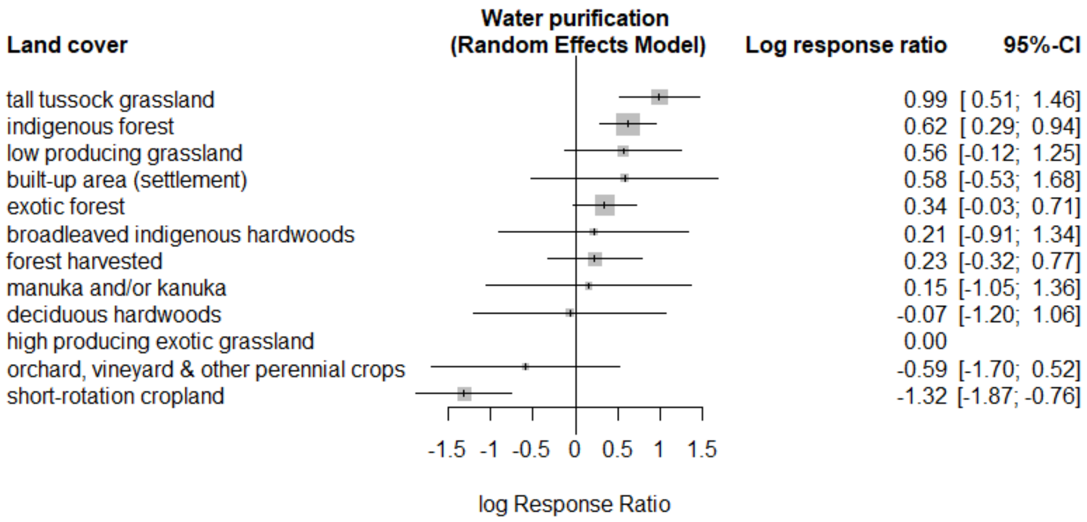


Fig. S19. *Forest plot of land cover contrasts in the provision of water purification. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 0.588$

$I^2 = 90.579$

Main results:

- Both tall tussock grasslands and indigenous forests stand out as land covers that do significantly better than the reference in the provision of this service, and overall, rank above all other land covers. Exotic and harvested forests also tend to perform slightly better the reference land cover in this service, although not significantly so.
- Croplands & high producing grasslands rank poorly, with short rotation croplands performing significantly worse than many of the land covers, including the high producing grasslands.
- Broadleaved indigenous hardwoods, manuka and/or kanuka, deciduous hardwoods and orchard, vineyard & other perennial crops all have wide confidence intervals and, as a result, are not significantly different from the other land covers.

- The relatively high ranking estimate for built areas is quite surprising here. However, the estimate is bounded by large confidence intervals and supported by direct comparisons with only two other land covers (high producing exotic grassland and indigenous forest), both of which stem from a single study.

Global climate regulation

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 29 indicators most of which are based on measures of greenhouse gas emission and sequestration processes in the soil. The majority of the indicators focus on carbon dioxide however, the data also include a few on methane and nitrous oxide.

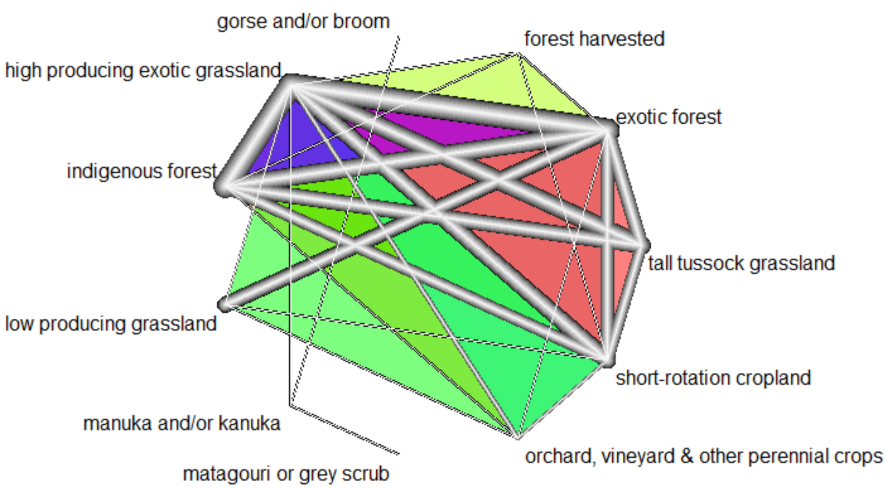


Fig. S20. *Evidence network for land cover comparisons on global climate regulation.*

Evidence base: This evidence network is formed by 75 pairwise comparisons of 11 land covers. Data were obtained from 33 different studies, each contributing a minimum of one and a maximum of four pairwise land cover comparisons. As indicated by the thicker lines in Fig. S20, the land covers that are most commonly compared are:

- High producing exotic grassland (45 comparisons)
- Exotic forest (34 comparisons)
- Short-rotation cropland (25 comparisons)
- Indigenous forest (19 comparisons)

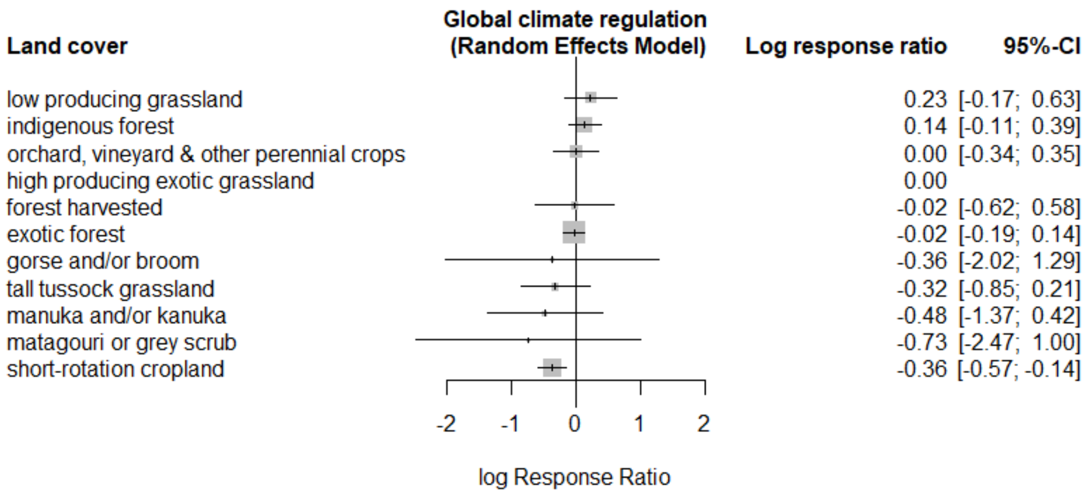


Fig. S21. *Forest plot of land cover contrasts in the provision of global climate regulation. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 0.335$

$I^2 = 94.585$

Main results:

- There are no significant differences between the high producing exotic grassland reference and all other land covers, except for short-rotation croplands which performs significantly worse than the reference and indigenous forest in delivering this service. The fact that short-rotation cropland does worse than the high producing grassland, and that the latter is not significantly different from the native land covers could, in part, be explained by the selection of indicators available for this service (which focus mainly on processes at the soil level instead of the entire land system which would include the effects of livestock).

Primary production

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 23 indicators. The larger proportion of these indicators express primary productivity as the amount of biomass within a given land cover, however there are also some indicators on the effects land covers have over primary productivity in streams (e.g.,. by providing more or less shade cover).

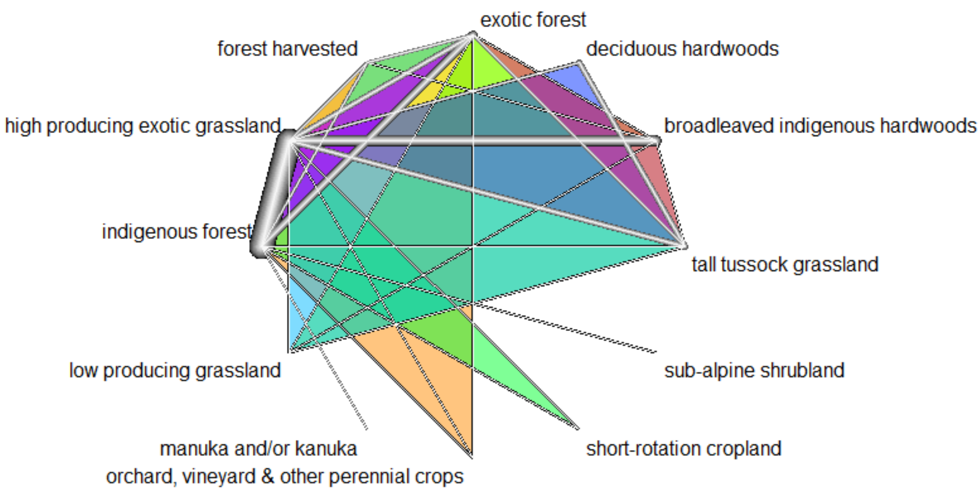


Fig. S22. *Evidence network for land cover comparisons on primary production.*

Evidence base: This evidence network is formed by 66 pairwise comparisons of 12 land covers. Data were obtained from 32 different studies, each contributing a minimum of one and a maximum of three pairwise land cover comparisons. As indicated by the thicker lines in Fig. S22, the land covers that are most commonly compared are:

- High producing exotic grassland (33 comparisons)
- Indigenous forest (27 comparisons)
- Exotic forest (25 comparisons)
- Tall tussock grassland (14 comparisons)

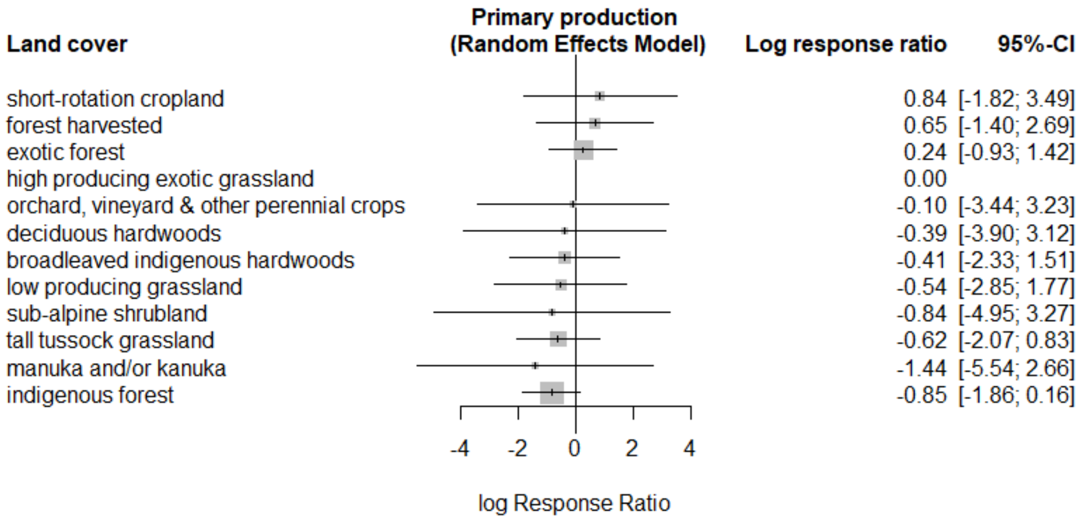


Fig. S23. *Forest plot of land cover contrasts in the provision of primary production. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 2.021$

$I^2 = 99.602$

Main results:

- All of the land cover effect estimates for this service have wide confidence intervals that overlap each other and intersect the reference mark for the high producing exotic grassland. However, the production land covers (exotic and harvested forests, croplands, deciduous hardwoods and high producing exotic grasslands) tend to be better at delivering this service than the native ones (broadleaved indigenous hardwoods, tall tussock grassland, sub-alpine shrubland, manuka and/or kanuka and indigenous forest). This is likely due to the high biomass turnover found in production systems and reflected in the measures of biomass accumulation we have for this service. Under the LCDB definition, forest harvested includes areas where native and exotic forests have been cleared and within those, areas were the forest has been replanted and is up to 5 years old. The fast growth rate of young forests could thus account for the high rank of this land cover in the provision of primary production.

Water cycling

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 16 indicators, all of which quantify the amount of water flowing through the different components of the terrestrial segment of the cycle. For this service we defined a positive relationship between all measures of water flow and the provision of the service, such that the greater the flow at a given land cover, the larger the contribution of that land cover to the provision of the service. This follows the MEA definition of water cycling as a supporting service that benefits all living organisms by allowing the movement of water through ecosystems (9).

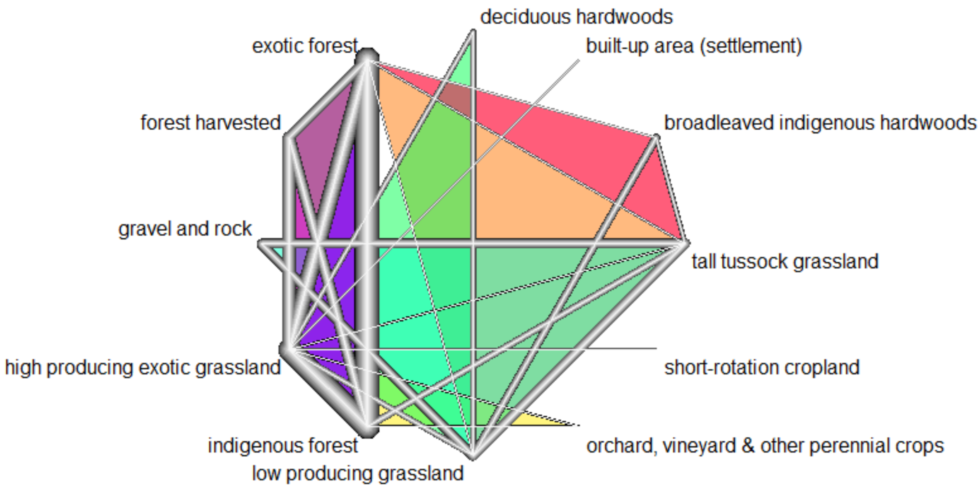


Fig. S24. *Evidence network for land cover comparisons on water cycling.*

Evidence base: This evidence network is formed by 61 pairwise comparisons of 12 land covers. Data were obtained from 29 different studies, each contributing a minimum of one and a maximum of three pairwise land cover comparisons. As indicated by the thicker lines in Fig. S24, the land covers that are most commonly compared are:

- High producing exotic grassland (31 comparisons)
- Exotic forest (26 comparisons)
- Indigenous forest (25 comparisons)
- Tall tussock grassland (10 comparisons)

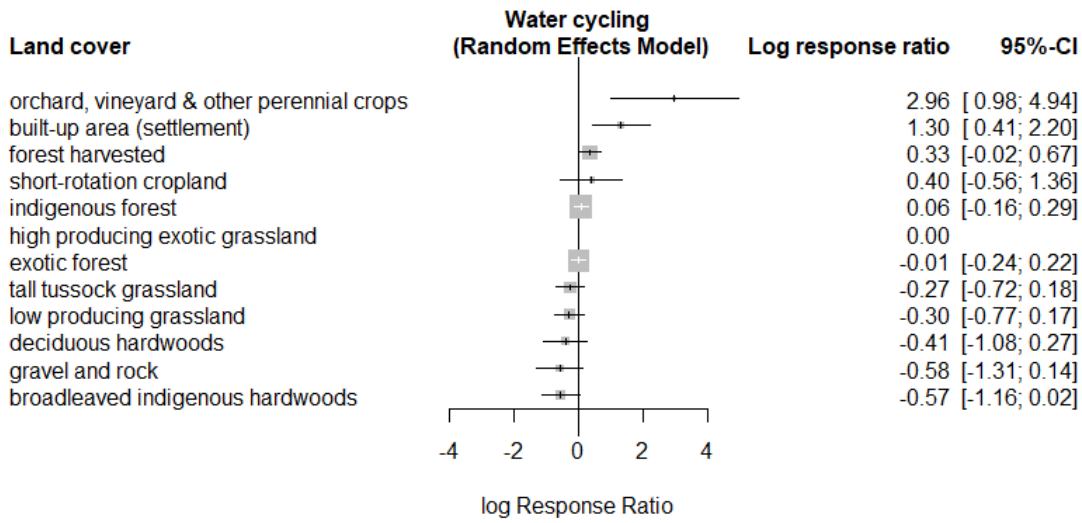


Fig. S25. *Forest plot of land cover contrasts in the provision of water cycling. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 0.332$

$I^2 = 77.458$

Main results:

- Built-up area and orchard, vineyard & other perennial crops stand out as land covers that do very well in providing this service. Their effect estimates stand out as significantly better than the high producing exotic grassland reference and most of the other land covers (except harvested forest and short-rotation cropland). For built-up (settlement) areas this could be an effect of increased flow speeds due to the presence of impervious surfaces.

- Systems that have slower water cycles (native and exotic forests, broadleaved indigenous hardwoods, tall tussock and low producing grasslands) tend to rank worse in this service than they do in the regulation of water timing and flows, suggesting a trade-off between water cycling and the regulation of flows. However, gavel and rock, which would be a system with faster cycling rates, also follows this trend. This could be driven by the direct evidence we have for gravel and rock under water cycling which involves water yield comparisons to tall tussock and low producing grasslands, and in which gravel and rock always performs worse (given its high infiltration capacity). In contrast, all other land covers with faster cycling rates (orchard, vineyard and other perennials, high producing exotic grassland and short-rotation cropland) are better at water cycling than they are at the timing and regulation of flows.

Erosion control

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 25 indicators on the magnitude of soil loss and sediment export to waterways as well as soil and stream channel characteristics that provide increased resistance to erosion.

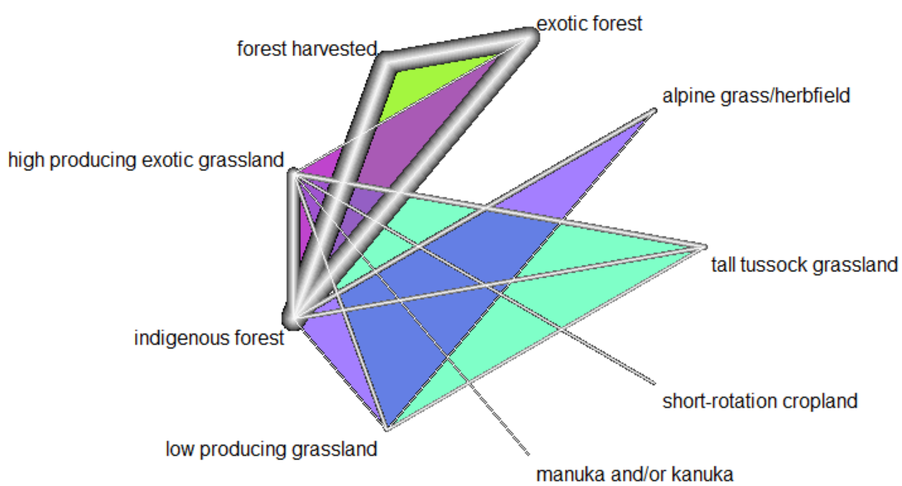


Fig. S26. *Evidence network for land cover comparisons on erosion control.*

Evidence base: This evidence network is formed by 36 pairwise comparisons of nine land covers. Data were obtained from 22 different studies, each contributing one to two pairwise land cover comparisons. As indicated by the thicker lines in Fig. S26, the land covers that are most commonly compared are:

- High producing exotic grassland (19 comparisons)
- Indigenous forest (19 comparisons)
- Exotic forest (12 comparisons)
- Forest harvested (five comparisons)

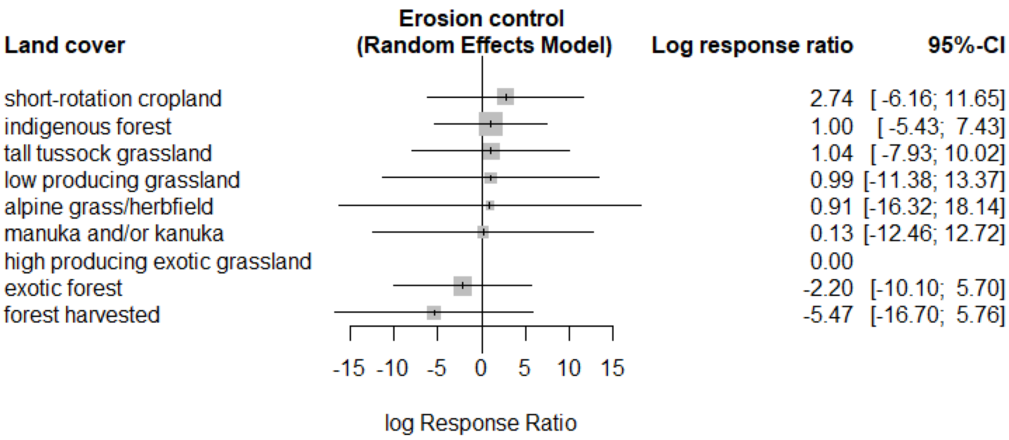


Fig. S27. *Forest plot of land cover contrasts in the provision of erosion control. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 9.08

I^2 = 99.967

Main results:

- Effect estimates are all bounded by wide confidence intervals, which yield no significant differences between any of the land covers. This could be due to the effect of environmental variables such as slope and parent material of the soil which introduce additional variation within each land cover.

- Except for short-rotation cropland, the production land covers tend to perform poorly in the provision of this service. In contrast, those with native vegetation covers (in the form of grass, forest or shrublands) have higher ranking estimates for their control over erosive processes.

Pest regulation

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 44 indicators, most of which focus on the abundance of invasive species in different land covers. However, there are also some indicators that quantify habitat occupation and use by invader species, and their response to biological controls.

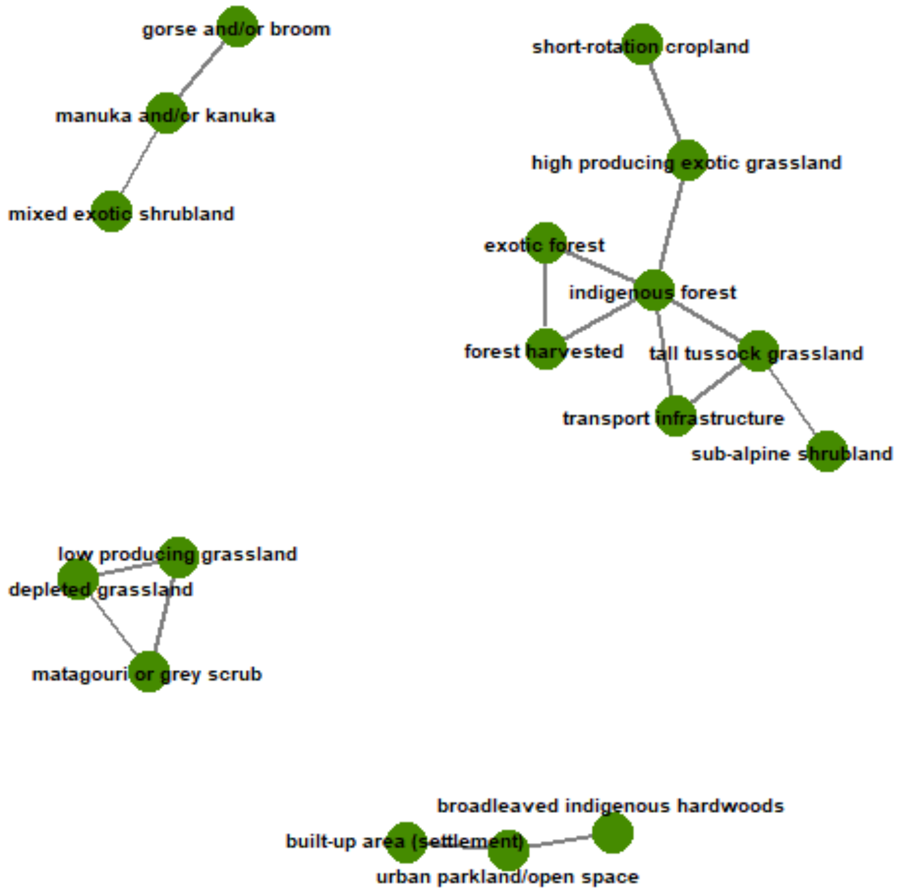


Fig. S28. *Land-cover comparison networks for pest regulation.*

Evidence base: As shown in Fig. S28, there are four networks connecting the land cover comparisons for this service: three smaller ones and a larger one. The evidence for the smaller networks is summarized in **Table S6**, whereas the remainder of this section describes the data and analysis for the larger network of land covers in this service.

Table S6. Reported response ratios for evidence subnetworks in pest regulation. Ratios are based on the natural logarithm of the quotient of land cover 1 over land cover 2

Sub-network	Land Cover 1	Land Cover 2	Log Response Ratio	Standard Error	Study ID
1	broadleaved indigenous hardwoods	urban parkland/open space	-0.045	0.206	S14626
1	built-up area (settlement)	urban parkland/open space	8.760	0.126	S16437
2	gorse and/or broom	manuka and/or kanuka	-0.572	0.304	S14266
2	manuka and/or kanuka	mixed exotic shrubland	0.117	0.065	S18250
3	depleted grassland	low producing grassland	0.105	0.593	S13628
3	depleted grassland	matagouri or grey scrub	0.063	0.550	S13628
3	low producing grassland	matagouri or grey scrub	-0.042	0.424	S13628

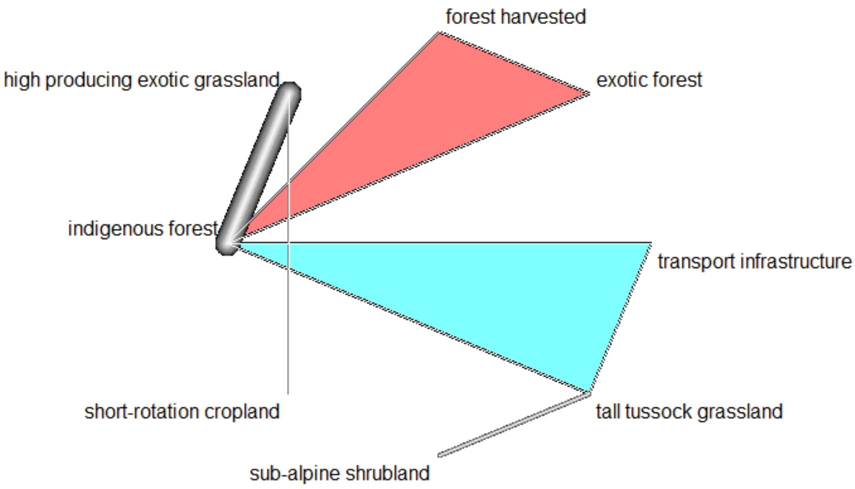


Fig. S29. *Evidence network for land cover comparisons on pest regulation.*

The largest evidence network for pest regulation is formed by 11 pairwise comparisons of eight land covers. Data were obtained from seven different studies, each contributing a minimum of one and a maximum of two pairwise land cover comparisons. As indicated by the thicker lines in Fig. S29, the land covers that are most commonly compared are:

- Indigenous forest (seven comparisons)
- Tall tussock grassland (three comparisons)
- Forest harvested (three comparisons)
- Exotic forest (three comparisons)

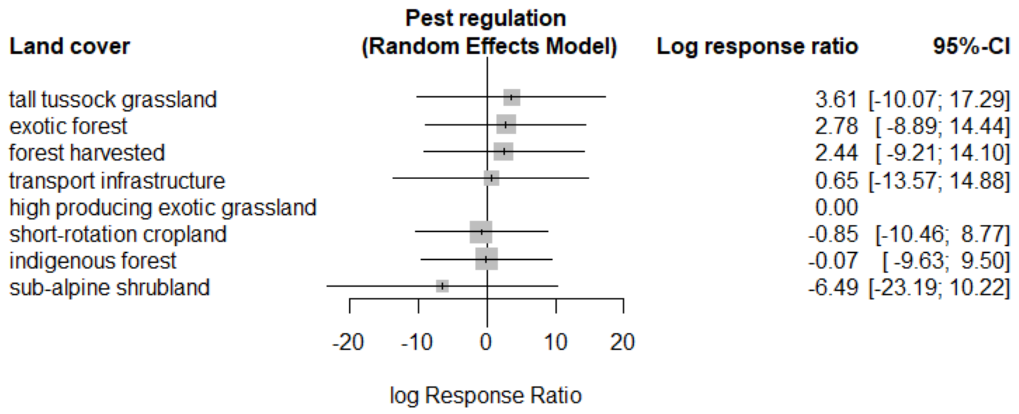


Fig. S30. *Forest plot of land cover contrasts in the provision of pest regulation. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

$\tau^2 = 4.88$

$I^2 = 96.012$

Main results:

- Effect estimates are all bounded by wide confidence intervals which yield no significant differences between the ratio estimates for any of the land covers. This could be explained by the fact that the evidence network has many comparisons converging around indigenous forest (Fig. S29) and that five out of the seven studies in this service provide evidence only on a single pair of land covers.
- In relation solely to the log Response Ratio estimates, tall tussock grasslands and exotic forests (harvested and unharvested) have the highest estimates, ranking above the high producing grassland reference in the provision of this service. In contrast, indigenous forests rank similarly to short-rotation croplands with small differences with the reference while the remaining native land cover, sub-alpine shrubland, performs worse than all other land covers in this service.

Waste treatment

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 21 indicators most of which provide a measure of the concentration and export of toxic compounds in the soil.

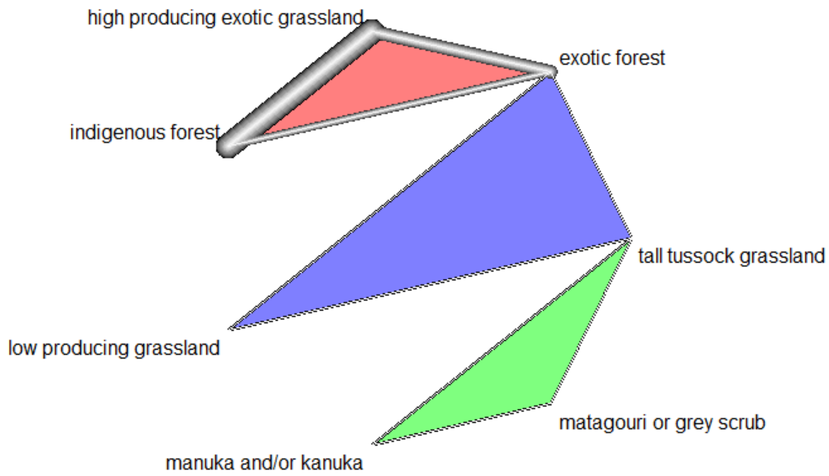


Fig. S31. *Evidence network for land cover comparisons on waste treatment.*

Evidence base: This evidence network is formed by 13 pairwise comparisons of seven land covers. Data were obtained from seven different studies, each contributing one to two pairwise land cover comparisons. As indicated by the thicker lines in Fig. S31, the land covers that are most commonly compared are:

- Exotic forest (seven comparisons)
- High producing exotic grassland (six comparisons)
- Tall tussock grassland (four comparisons)
- Indigenous forest (three comparisons)

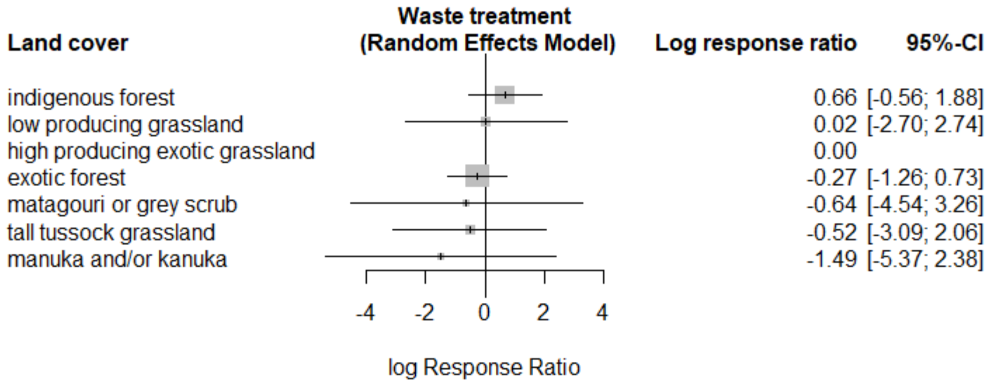


Fig. S32. *Forest plot of land cover contrasts in the provision of waste treatment. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 0.911

I^2 = 94.681

Main results:

- No significant differences can be found between the land covers in this service since the confidence intervals all overlap each other and extend over both sides of the baseline reference. Furthermore, with the exception of manuka and/or kanuka, the estimated log Response Ratio between every land covers and the high producing exotic grassland reference is always between one and minus one.
- Indigenous forest and grassland land covers have rank the highest in the provision of this service, while the exotic forest and the remaining native land covers (tall tussocks, manuka and/or kanuka, matagouri or grey scrub) all rank poorly an perform comparatively worse than the reference land cover.

Capture fisheries

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 22 indicators on the abundance, biomass, size and growth of freshwater fish.

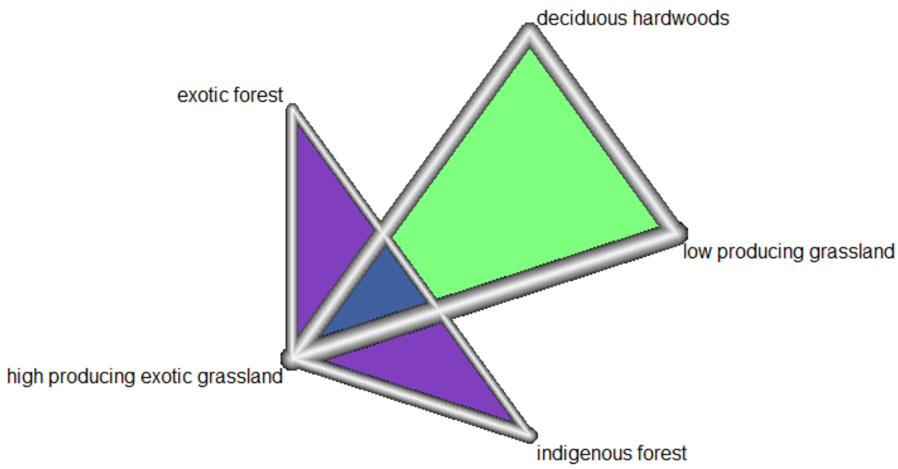


Fig. S33. *Evidence network for land cover comparisons on capture fisheries.*

Evidence base: This evidence network is formed by 11 pairwise comparisons of five land covers. Data were obtained from five different studies, each contributing one to two pairwise land cover comparisons. The land covers that are most commonly compared are:

- High producing exotic grassland (eight comparisons)
- Indigenous forest (five comparisons)
- Exotic forest (four comparisons)
- Low producing grassland (three comparisons)

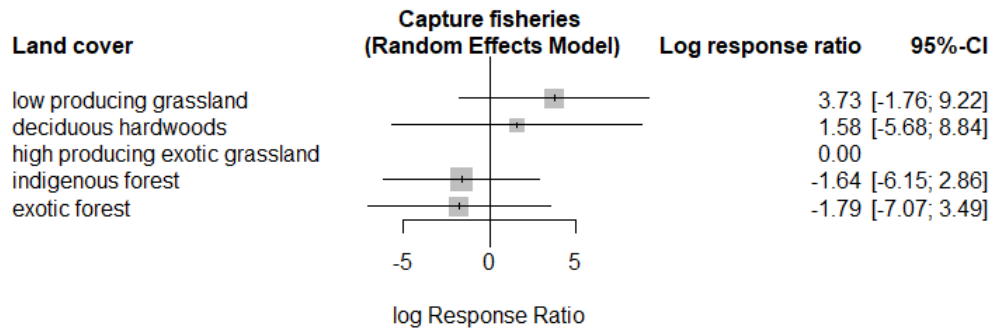


Fig. S34. *Forest plot of land cover contrasts in the provision of capture fisheries. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 3.955

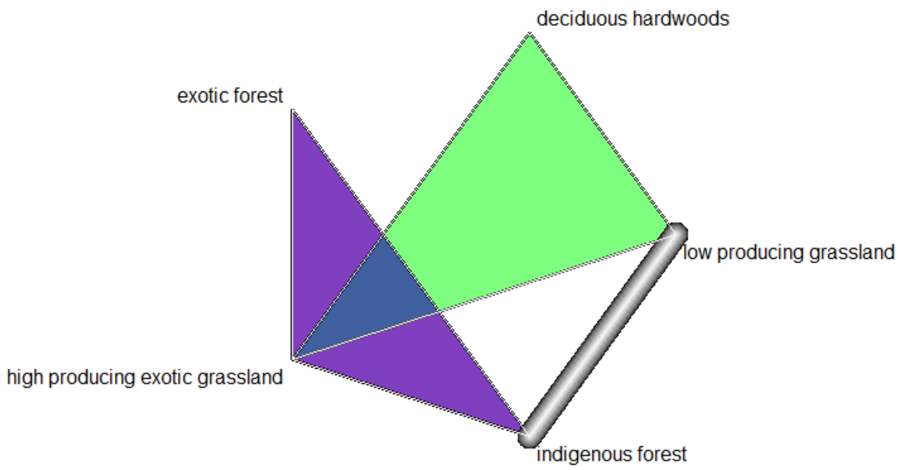
I^2 = 99.399

Main results:

- The wide confidence intervals for all land covers preclude any significant differences between any of the land covers. However, in terms of the ratio estimates, both indigenous forests tend to deliver less of the capture fisheries service than the low and high producing grasslands and the deciduous hardwoods.

Ethical and spiritual values

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from 14 indicators, all of which express the abundance, biomass, and growth of culturally valuable fauna namely, bats an eels. Information for the latter includes most of the indicators for the capture fisheries service.



992

993 ***Fig. S35.*** *Evidence network for land cover comparisons on ethical and spiritual values.*

994

995 **Evidence base:** This evidence network is formed by 11 pairwise comparisons of five land
996 covers. Data were obtained from five different studies, each contributing one to two
997 pairwise land cover comparisons. The land covers that are most commonly compared are:

998 • High producing exotic grassland (seven comparisons)
999 • Indigenous forest (five comparisons)
1000 • Low producing grassland (four comparisons)
1001 • Exotic forest (four comparisons)

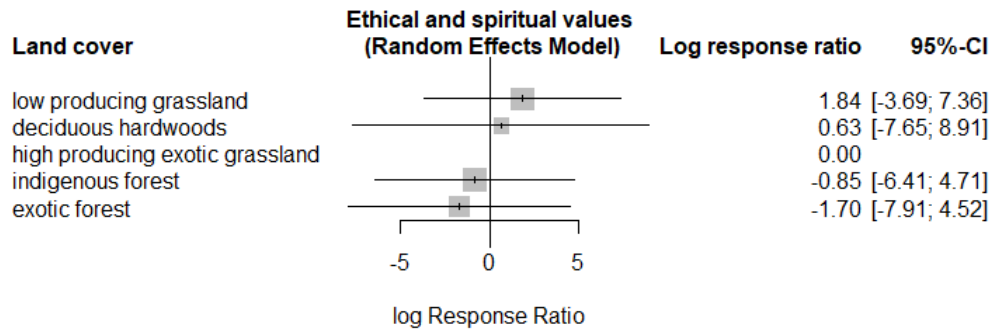


Fig. S36. *Forest plot of land cover contrasts in the provision of ethical and spiritual values. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 4.594

I^2 = 99.58

Main results:

- Land covers in this service appear to share the same trends as those for the capture fisheries service with wide confidence intervals and the estimated ratios for exotic and native forests at the negative end of the spectrum and being smaller than those for the grasslands and deciduous hardwoods. This is an effect of the large overlap between the indicators that support the evidence on this service and that of capture fisheries.

Disease mitigation

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from nine indicators on the abundance of mosquitoes and their predators, and on the presence of fecal coliforms, *Esceherichia coli* or entreocoliforms in streams and freshwater sources.

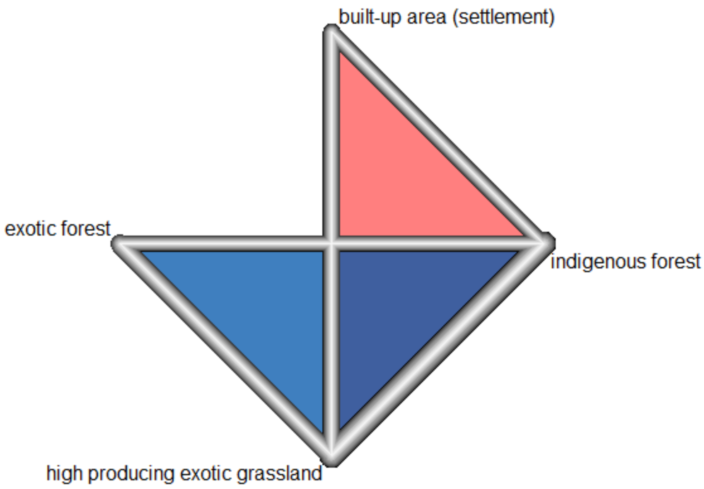


Fig. S37. *Evidence network for land cover comparisons on disease mitigation.*

Evidence base: This evidence network is formed by ten pairwise comparisons of four land covers. Data were obtained from four different studies, each contributing one to two pairwise land cover comparisons. The number of comparisons per land cover is as follows:

- Indigenous forest (seven comparisons)
- High producing exotic grassland (seven comparisons)
- Exotic forest (four comparisons)
- Built-up area (settlement) (two comparisons)

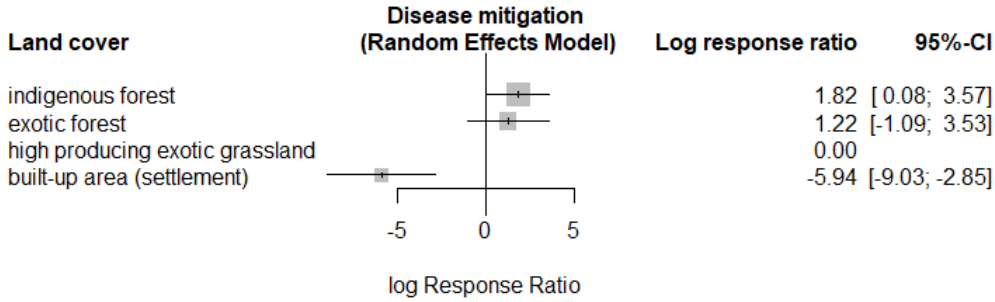


Fig. S38. *Forest plot of land cover contrasts in the provision of disease mitigation. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 1.729

I^2 = 95.505

Main results:

- Built-up area stand out as the land cover that performs significantly worse than any of the others in delivering this service.
- Production land covers (exotic forest and high producing exotic grassland) perform slightly worse than the indigenous forest with respect to disease mitigation. The difference with the indigenous forest is significant for the high producing exotic grasslands but not for the exotic forest.

Pollination

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from four indicators that quantify the potential for pollination in a land cover based on either the abundance or body size of pollinators or on the flower visitation rates and duration.

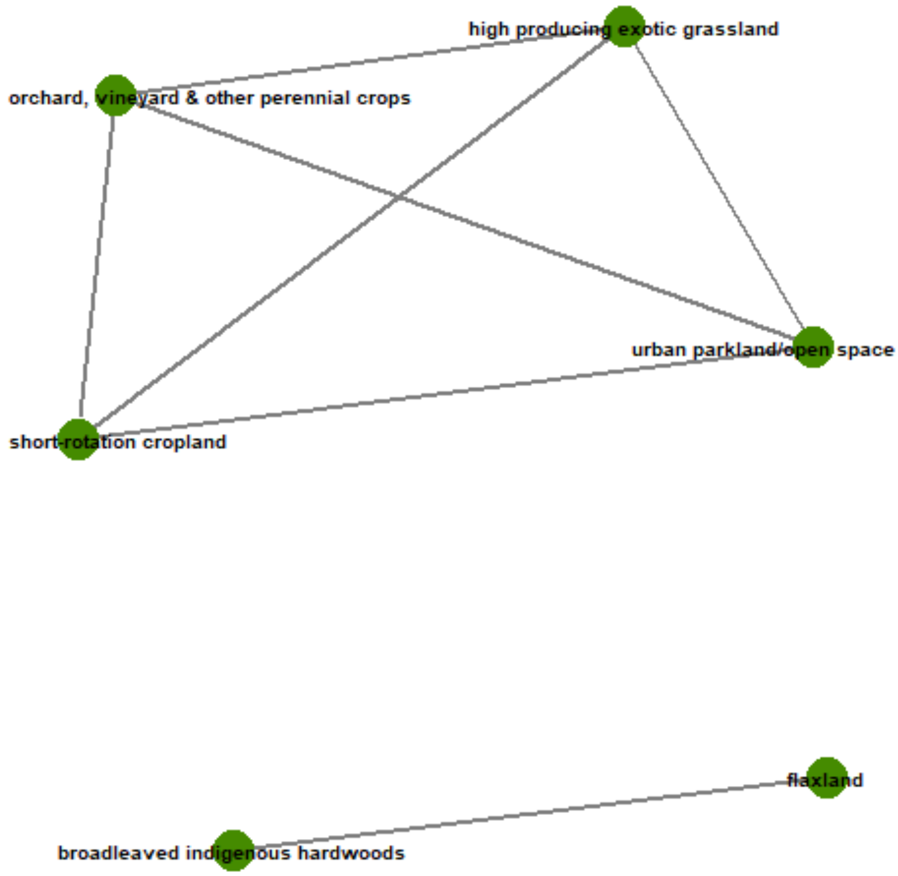


Fig. S39. *Land-cover comparison networks for pollination.*

Evidence base: As shown in Fig. S39, the evidence for this service is split into two networks: a smaller one connecting broadleaved indigenous hardwoods and flaxland and a larger one with the four remaining land covers. For the land covers in the smaller sub-network, the evidence available comes from a single study in which the log Response Ratio of broadleaved indigenous hardwoods to flaxland is approximately -0.823 and has a standard error of 0.625. Details for the larger network are presented below.

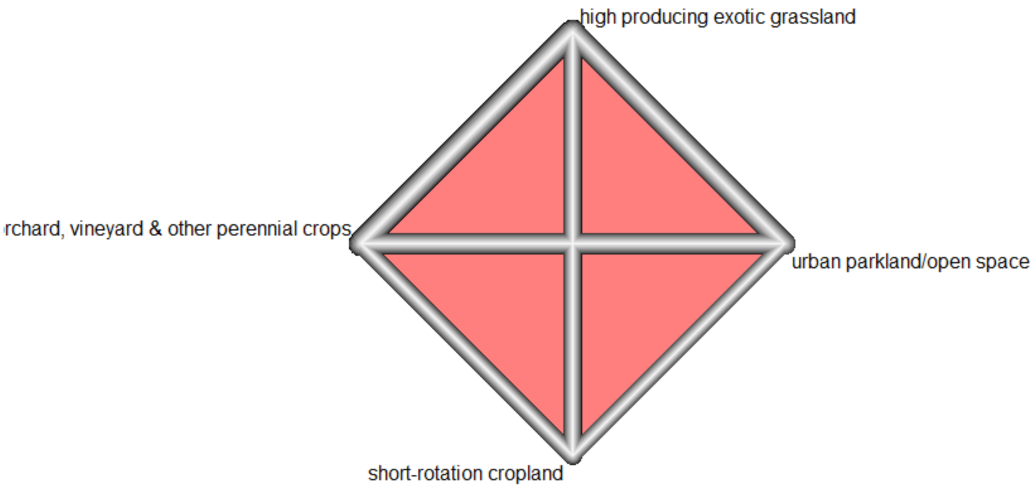


Fig. S40. *Evidence network for land cover comparisons on pollination.*

This evidence network is formed by seven pairwise comparisons of four land covers. Data were obtained from two different studies, each contributing a minimum of one and a maximum of three pairwise land cover comparisons. As indicated by the thicker lines in Fig. S44, the land covers that are most commonly compared are:

- Urban parkland/open space (four comparisons)
- High producing exotic grassland (four comparisons)
- Short-rotation cropland (three comparisons)
- Orchard, vineyard & other perennial crops (three comparisons)

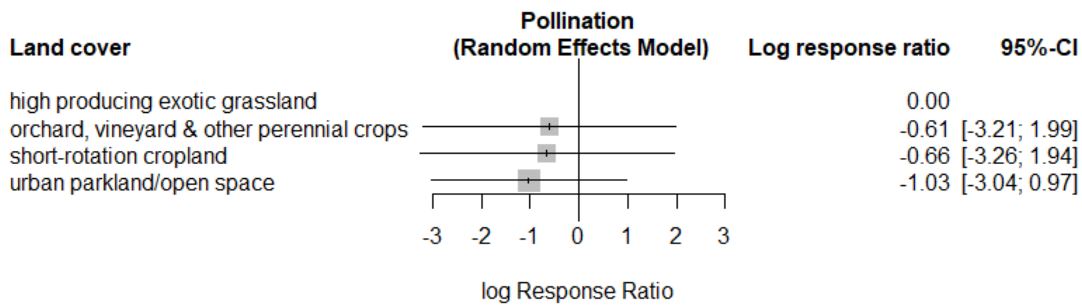


Fig. S41. *Forest plot of land cover contrasts in the provision of pollination. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = 1.412

I^2 = 95.049

Main results:

- There are no significant differences between the different land covers in the provision of this service, since all confidence intervals for the ratio estimates overlap each other and extend across the baseline reference. In part this could be a result of having only 2 studies informing this analysis.
- The ratio estimates suggest that the urban parkland/open spaces and, to a lesser extent, the croplands perform worse than the high producing exotic grasslands in delivering pollination services.

Regional & local climate regulation

Type of indicators for this ecosystem service: Comparisons for this ecosystem service were drawn from three indicators that quantify the regulation of temperatures either in stream water or near the land surface (as expressed by evapotranspiration).

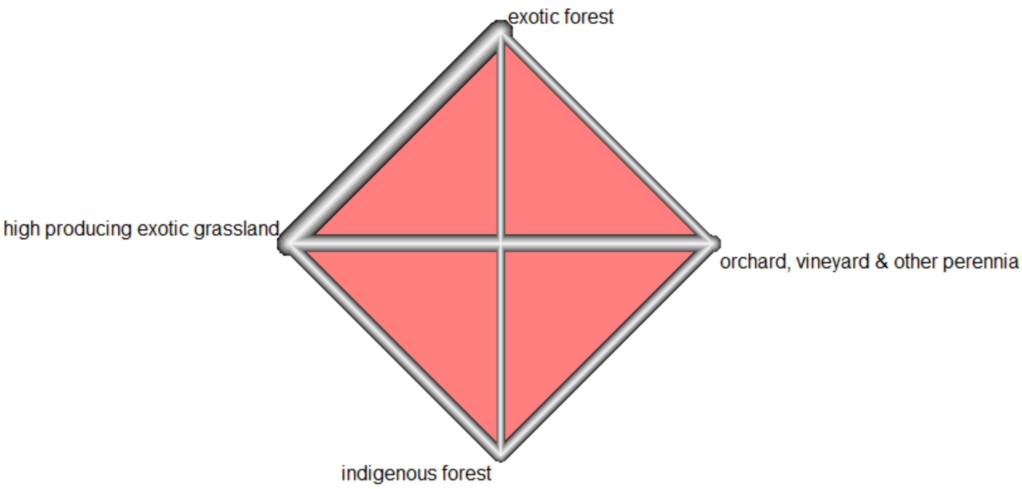


***Fig. S42**. Evidence network for land cover comparisons on regional and local climate regulation.*

Evidence base: This evidence network is formed by seven pairwise comparisons of four land covers. Data were obtained from two different studies, each contributing a minimum of one and a maximum of three pairwise land cover comparisons. The number of comparisons per land cover is as follows:

- High producing exotic grassland (four comparisons)
- Exotic forest (four comparisons)
- Orchard, vineyard & other perennial crops (three comparisons)
- Indigenous forest (three comparisons)

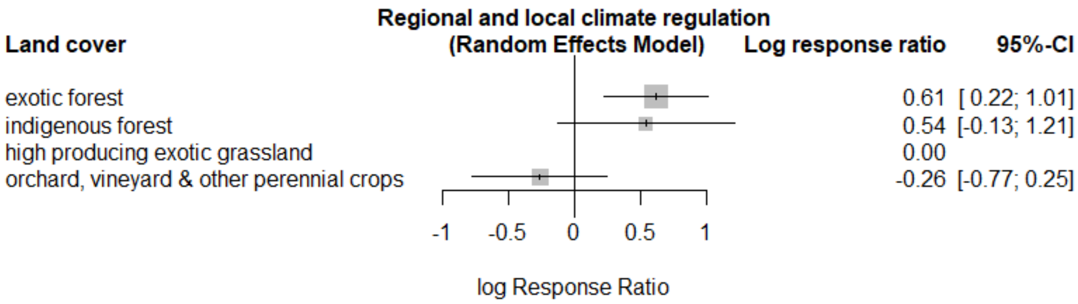


Fig. S43. *Forest plot of land cover contrasts in the provision of regional and local climate regulation. Random effects model with high producing exotic grassland as a reference.*

Measures of heterogeneity/network inconsistency:

τ^2 = Not available

I^2 = Not available

Main results:

- Besides exotic forests (which are significantly better than the reference land cover), all the land covers have overlapping confidence intervals and, consequently, are not significantly different from each other in the provision of this service. This is probably due to the limited number of studies with evidence on this service.
- The ratio estimates suggest that forested land covers tend to deliver greater climate regulation at the local and regional level than the high producing exotic grassland and the orchard land covers.

A note on confidence intervals and the size of the evidence base

The forest plots above show that, primary production, erosion control, pest regulation, waste treatment, capture fisheries, ethical & spiritual values, pollination and regional & local climate regulation all present wide, overlapping confidence intervals for all or most of their estimates. This suggests that differences in the provision of all of these services across land covers were not significant. For some of these services, this is due to smaller evidence bases, as in the case of pollination and regional & local climate regulation where the network meta-analysis was informed by only 7 - 8 comparisons taken from 2 different studies. For services with a slightly larger evidence base (e.g., capture fisheries, ethical and spiritual values, pest regulation and waste treatment) there is an asymmetry in the number of comparisons available for the different land covers, since one or two pairs of land covers harness most of the comparisons and leave limited evidence for the other pairs of comparisons (note the large differences in link weights in the corresponding evidence networks presented above). With over 60 pairwise comparisons across all land covers, the evidence base for primary production has a similar problem since most of the comparisons involve high producing exotic grasslands, indigenous and exotic forests. A similar trend can be observed with some (but not all) pairs of land covers for other well-informed services, such as regulation of water timing and flows and soil formation. For waste treatment, the low sample size results from having few comparisons (13 in total) spread over a large number of land covers (7 in all), which results in an evidence network formed by several, poorly informed links.

Appendix 11. Detailed results from PERMANOVA analyses

Table S7. Detailed output of the permutational analysis of variance (PERMANOVA, adonis) on the effects of land cover characteristics on land cover provision of multiple ecosystem services. Model specified with production as the first variable term.

Variable	Degrees of freedom	*F*	Partial R^2	*p* - value
Production	1	2.927	0.259	**0.00995**
Forest cover	1	1.226	0.109	0.289
Interaction (Production : Forest cover)	1	1.141	0.101	0.338
Residuals	6		0.531	
Total	9		1.000	

Table S8. Detailed output of the permutational analysis of variance (PERMANOVA, adonis) on the effects of land cover characteristics on land cover provision of multiple ecosystem services. Model specified with forest cover as the first variable term.

Variable	Degrees of freedom	F	Partial R^2	p - value
Forest cover	1	1.226	0.109	0.333
Production	1	2.927	0.259	**0.0199**
Interaction (Production : Forest cover)	1	1.141	0.101	0.328
Residuals	6		0.531	
Total	9		1.000	

Table S9. Comparisons of group mean dispersions for variables on land cover characteristics. Separate tests were conducted for each variable (permdisp: betadisper and permutest with 999 permutations).

Variable	$F_{(1,8)}$	p- value
Production	0.1508	0.683
Forest cover	1.1379	0.348

Table S10. Detailed output of the PERMANOVA (adonis) on the effects of ecosystem service characteristics on ecosystem service provision across multiple land covers. Model specified with biophysical domain as the first variable term.

Variable	Degrees of freedom	F	Partial R^2	p - value
Biophysical domain	2	2.253	0.312	0.065
Scale of benefits	2	2.973	0.411	**0.045**
Residuals	4		0.277	
Total	8		1.000	

Table S11. Detailed output of the PERMANOVA (adonis) on the effects of ecosystem service characteristics on ecosystem service provision across multiple land covers. Model specified with scale of benefits as the first variable term.

Variable	Degrees of freedom	F	Partial R^2	p - value
Scale of benefits	2	2.337	0.323	0.055
Biophysical domain	2	2.888	0.400	**0.035**
Residuals	4		0.277	
Total	8		1.000	

Table S12. Comparisons of group mean dispersions for variables on ecosystem service characteristics. Separate tests were conducted for each variable (permdisp: betadisper and permutest with 999 permutations).

Variable	$F_{(1,8)}$	p - value
Scale of benefits	0.486	0.688
Biophysical domain	0.823	0.623

Captions for datasets S1 to S2

SI Dataset 1: Quantitative indicators used to quantify provision of each ecosystem service. For each indicator the units used by different studies are given. Indicators that lack units are reported as index or ratio in the units column. If the indicator is a variable that was logged, units of the variable before applying the logarithm are generally given. “No. Studies” describes the number of studies reporting each of the indicators in our dataset.

SI Dataset 2: Final log Response Ratios on ecosystem service provision for pairwise comparison of land covers in each study used in our analysis. Within each study, log Response Ratios of multiple indicators of provision for the same ecosystem service have been aggregated to a single value per service for each pairwise land cover comparison of land covers. Column heading abbreviations: Ecosystem service (ES), Study ID (S.ID), Land Cover (LC), Log Response Ratio (LRR), Variance (Var) and Standard Errror (SE).

References

1. MacLeod CJ, Moller H (2006) Intensification and diversification of New Zealand agriculture since 1960: An evaluation of current indicators of land use change. *Agriculture, Ecosystems and Environment* 115(1-4):201–218.

2. Wallace BC, Small K, Brodley CE, Lau J, Trikalinos TA (2012) Deploying an interactive machine learning system in an evidence-based practice center. *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics - IHI '12*:819.

3. Wan X, Wang W, Liu J, Tong T (2014) Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology* 14(1):135.

4. Taylor LR (1961) Aggregation, variance and the mean. *Nature* 189(4766):732–735.

5. Rosenberg MS, Rothstein HR, Gurevitch J (2013) Effect sizes: conventional choices and calculations. *Handbook of Meta-Analysis in Ecology and Evolution* (Princeton University Press, Princeton), pp 61–71.

6. Rücker G, Schwarzer G (2015) Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Medical Research Methodology* 15(1):58.

7. Lavelle P, et al. (2005) Nutrient Cycling. *Ecosystems and Human Well-Being: Current State and Trends Asessment*, pp 333–353.

8. Graaff M-A de, Groenigen K-J van, Six J, Hungate B, Kessel C van (2006) Interactions between plant growth and soil nutrient cycling under elevated CO_2 : a meta-analysis. *Global Change Biology* 12(11):2077–2091.

9. MEA (2005) *Ecosystems and human well-being: Synthesis* (Island Press, Washington) Available at: http://www.who.int/entity/globalchange/ecosystems/ecosys.pdf.