Received: 11 May 2018 | Accepted: 25 August 2018
DOI: 10.1111/2041-210X.13084

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Not all data are equal: Influence of data type and amount in spatial conservation prioritisation

Heini Kujala[1] | José Joaquín Lahoz-Monfort[1] | Jane Elith[1,2] | Atte Moilanen[3,4]

[1]School of Biosciences, The University of Melbourne, Melbourne, Victoria, Australia
[2]ARC Centre of Excellence for Environmental Decisions, The University of Melbourne, Melbourne, Australia
[3]Finnish Natural History Museum, University of Helsinki, Helsinki, Finland
[4]Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland

Correspondence
Heini Kujala
Email: heini.kujala@unimelb.edu.au

Funding information
Australian Research Centre Discovery Project, Grant/Award Number: DP160101003; Academy of Finland Centre of Excellence programme, Grant/Award Number: 250444

Handling Editor: Andrés Lopez-Sepulcre

Abstract

1. Decisions about land use significantly influence biodiversity globally. The field of spatial conservation prioritisation explores allocation of conservation effort, including for reserve network expansion, targeting habitat restoration, or minimising ecological impacts of development. Inevitably, the utility of such planning depends on the quantity and quality input data, including spatial information on biodiversity, threats, and cost of action. In this work we systematically develop understanding about the significance of these different data types in spatial conservation prioritisation.
2. We clarify the common ways different data types enter an analysis, develop mathematical models to understand the effects of data in spatial prioritisation, and survey literature to establish typical quantities of different types of data used. We use Jackknife analysis to derive the expected change in site values, when a single new data layer is added to a prioritisation. We validate mathematical formulae for expected impacts using simulations.
3. A survey of scientific literature reveals that typical spatial prioritisation analyses include hundreds of biodiversity feature layers (species, habitat types, ecosystem services), but the count of cost, threat or habitat condition layers is typically 0–5. Due to these differences, and the mathematical formulations commonly used to combine data types, the influence of a single cost, threat, or habitat condition data layer can be an order or two higher than the influence of a single biodiversity feature layer. In a classical cost-effectiveness formulation (benefits divided by costs, B/C) the influence of a single cost layer can even be as large as the joint influence of thousands of species distributions. We also clarify how changes in data impact site values and spatial priority rankings differently, with the latter being further influenced by data correlations, the spread of numeric values inside data layers and other data characteristics. For example, costs influence priorities significantly if cost is positively correlated with biodiversity, but the correlation is the other way around for biodiversity and habitat condition.
4. This work helps conservation practitioners to direct efforts when collating data for spatial conservation planning. It also helps decision makers understand where to focus attention when interpreting conservation plans and their uncertainties.

KEYWORDS
biodiversity, costs, data quality, habitat condition, spatial prioritisation, systematic conservation planning, threat, uncertainty

1 | INTRODUCTION

Many conservation decisions are spatially explicit, including reserve network expansions, targeting areas for habitat restoration or reintroductions, and directing surveys to find endangered or invasive species. In the past two decades, several analytical approaches and software tools have been developed to help practitioners and decision makers choose locations for action so that conservation objectives are met cost effectively (Margules & Pressey, 2000; Moilanen, Wilson, & Possingham, 2009). Following the principle of complementarity, these tools utilise mathematical formulations and optimisation methods to analyse often large amounts of spatial data and to identify sets of locations that jointly best meet case-specific conservation objectives (Kukkala & Moilanen, 2013; Wilson et al., 2007).

Spatial prioritisation results must always be interpreted in the context of the quantity and quality of data used as inputs (Lehtomäki & Moilanen, 2013). When prioritising for conservation, analyses invariably use data about distributions of biodiversity features including species, habitat types and their condition, and ecosystem services. Costs, direct or indirect, are another central component: inclusion of cost in spatial prioritisation allows identification of cost-efficient solutions (Armsworth, 2014; Naidoo et al., 2006; Wilson et al., 2007). Further important data category is threats (pressures, drivers), which are used to target, or avoid, areas with negative influences on biodiversity, depending on the objective (Joppa et al., 2016). Threats are also often intimately linked to the appropriate actions that need to be taken, and as such influence conservation costs.

Data are always, to some extent, uncertain. As biodiversity data are never complete, analyses implicitly or explicitly rely on the concept of surrogacy (Rodrigues & Brooks, 2007), with the assumption that a sufficiently large sample of biodiversity features will adequately represent biodiversity as a whole. Questions of uncertainty and surrogacy are widely acknowledged amongst conservation planners, and concerns about adequacy of data commonly arise. Statistical species distribution models (SDMs, Franklin, 2013) have become popular for extrapolating biodiversity patterns when comprehensive survey data are not available, the usual situation for most species and areas. As in all models, several sources of error introduce uncertainties into SDMs: low number of observations available for model fitting, biases or gaps in observations or predictors, and the assumptions and quality of the statistical model itself (Barry & Elith, 2006).

Conservation scientists have traditionally devoted much effort to collecting and improving biodiversity data and models underpinning spatial conservation plans. Yet, other important information such as current or future conservation costs or anthropogenic threats is also uncertain (Joppa et al., 2016). Relatively little effort has been spent improving maps of threats or costs (Armsworth, 2014), and this could be problematic if they drive spatial prioritisations strongly (Balmford, Gaston, Blyth, James, & Kapos, 2003; Bode et al., 2008; Naidoo et al., 2006). It is, therefore, relevant to question how uncertainties in different data types affect conservation decisions.

Prior work suggests that there are limits to the amount of species data needed for successful planning (Grantham et al., 2008; Kujala, Moilanen, & Gordon, 2018), but these studies do not discuss the relative roles of different data types. Some have explored the impact of different data gaps on conservation plans (e.g., Carwardine et al., 2010; Visconti et al., 2013; Wilson & , 2005) but only within the context of a specific conservation case, making it difficult to separate the influence of data uncertainty from other factors, such as conservation objectives, targets, data characteristics and correlations (Armsworth, 2014; Ferraro, 2003), and prioritisation methods. In all such analyses, it is useful to differentiate between *site value* and *the priority of a site*. Here we define site value as the expected numerical value of a location, aggregated across all data layers. Site value can be solely based on the biodiversity benefits of acting at the site, as defined by the known or predicted biodiversity present. This is sometimes referred to as the *conservation value* of a site in the literature, although terminology varies (Kukkala & Moilanen, 2013; Margules & Pressey, 2000). In spatial prioritisations, the costs and consequences of acting at a site also affect its numerical value, in which case terms *cost-effectiveness* or *return-on-investment* (ROI) are often used (Armsworth et al., 2017). Priority of a site, on the other hand, measures the relative urgency and cost-effectiveness of acting at a location in relation to other candidate sites. A *priority ranking* is a rank order interpretation of conservation priority.

This study analyses how different data types influence spatial conservation plans. We start by clarifying how they typically enter a spatial prioritisation analysis. We then show how knowledge of the mathematical structure, together with information about the typical number of data layers used for each data type, can help to estimate the relative influence a data type will have on site values and priorities. Because it is desirable that the most influential data is most accurate, this framework should help researchers and conservation practitioners make improved decisions about when and where to spend resources to improve spatial data.

2 | MATERIALS AND METHODS

We start by conceptualising how different types of data typically enter spatial prioritisations (Figure 1). We differentiate between spatial data layers that enter analysis directly versus indirect data that influence other layers. Usually spatial distributions of biodiversity features and costs enter analysis directly. Environmental factors usually enter indirectly as explanatory variables within species (or habitat) distribution models. Habitat condition is often used to modify biodiversity layers.

Accessibility of ecosystem services can be used to link ecosystem service provision (supply) to demand. Threats can enter an analysis in several ways, either directly or via impact on other layers (Figure 1, and "Threats" section, below). The need for, and information content of, data layers may also be affected by other factors. For example, threats may dictate the actions needed to achieve conservation goals, which in turn can influence information on costs (Figure 1). The number of data layers per data type is also highly variable. Biodiversity is usually described by numerous feature layers, whereas only one or a few cost layers are used (Armsworth, 2014). Differences in the position in mathematical formulation and in layer counts lead to substantial differences in the relative impacts of different data types. To illustrate these differences, we ask: If a single new data layer is added to an existing conservation analysis, how much will the conservation value of a location change depending on type of layer added?

We use simplified mathematical definitions of how data impact spatial priorities, focusing on the main data types of biodiversity features, costs, condition, and threats (see Table 1 for symbols used). The analysis uses scoring methods (Moilanen et al., 2009), but is also relevant for all target-based reserve selection algorithms (e.g., Marxan, Ball, Possingham, & Watts, 2009) and spatial priority ranking methods (e.g., Zonation, Moilanen et al., 2005): Whereas the sensitivity of more complex approaches to data type is further influenced by other components (e.g., targets, complementarity) not explored here, there are inevitable commonalities across all these methods in data preprocessing before analysis and in the broad structure of analysis, for example, in how benefits and costs are combined.

2.1 | Mathematical formulation of data type influence

2.1.1 | Common conservation value functions

We identify three common ways in which biodiversity information and costs are used to calculate the site value (V_i) of a location i:

$$V_i = \begin{cases} B_i, & \text{(a)} \\ B_i/C_i, & \text{(b)} \\ B_i - C_i, & \text{(c)} \end{cases} \quad (1)$$

Equation (1a) uses only biodiversity data, whereas the two latter ones use both biodiversity and cost data. Cost is most commonly used as denominator in a classic cost-efficiency analysis (Armsworth, 2014; Possingham, 2001). Alternatively, (weighted) costs can be subtracted from benefits (Equation 1c), in which case one or both data types need to be transformed, for example, via scaling or normalisation, to match their units. For simplicity, here we assume values in Equation 1c to be already transformed. Equation 1c is commonly used in currently available spatial prioritisation software, such as Marxan (Ball et al., 2009) and Zonation (Moilanen et al., 2011a),

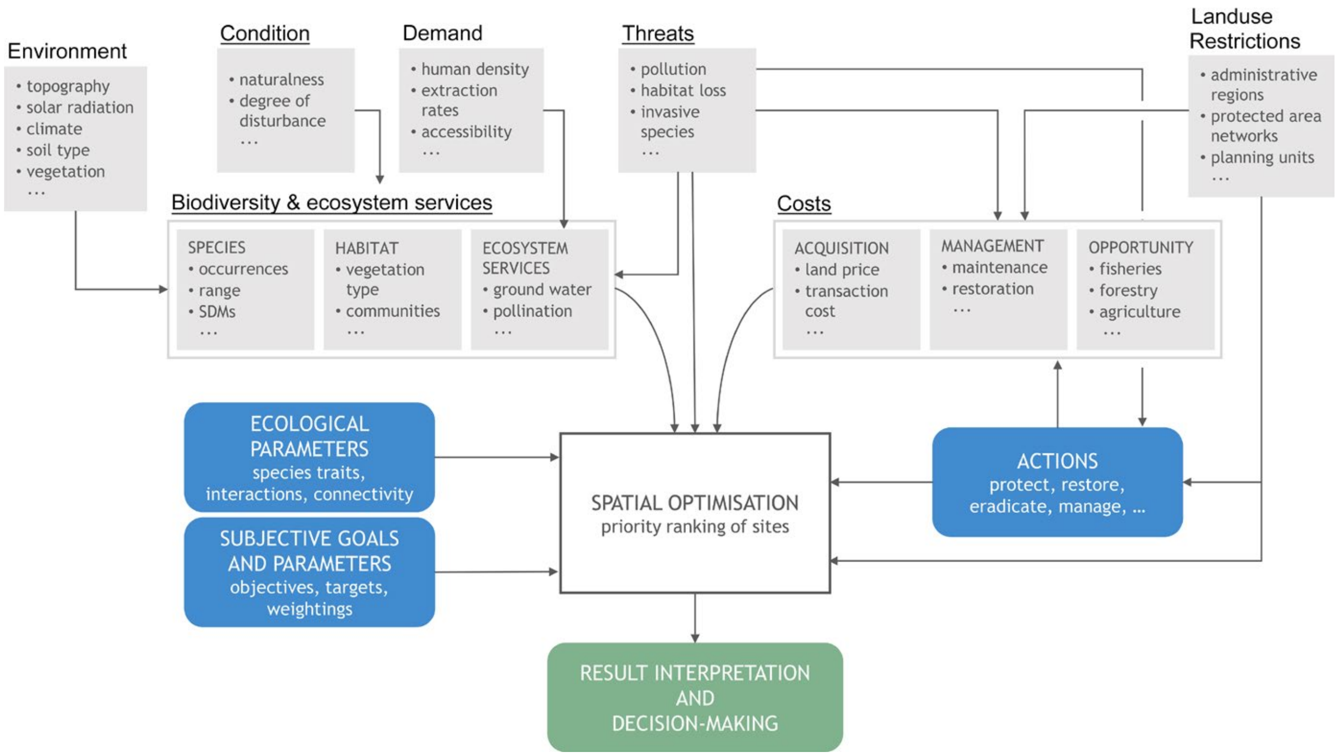


FIGURE 1 Flow of data in spatial prioritisation. Some types of data (e.g., SDMs) enter spatial prioritisation directly, whereas others have indirect impacts via other layers. Additional to spatial data (grey boxes), there are also other factors (blue boxes) that influence analysis. In all cases, the results and associated quantitative information need to be interpreted from the perspective of the objectives. Arrows indicate flow of data into and out of the analysis. We note that the connections between components in this figure are not exhaustive but illustrate typical linkages relevant for data flow. The data types assessed in this work are underlined

TABLE 1 Mathematical symbols

Symbol	Description
s_{ij}	Numerical value (e.g., occurrence level) of biodiversity feature (layer) j in location i.
c_{ik}	Numerical value of cost component (layer) k in location i.
B_i	Biodiversity benefits of location i, aggregated across biodiversity features.
C_i	Costs of including location i to a conservation plan. May be aggregated from several cost components.
n_B, n_C	Number of data layers representing biodiversity features (n_B) or costs (n_C).
$\bar{w}_B$, $\bar{w}_c$	Average of the weights given to biodiversity features (w_j^B) or cost layers (w_k^C).
h_i	Measure of habitat condition or threat that modifies the values of features at a site i. May influence all features or a subset of them (e.g., one taxon).
p	Proportion of biodiversity layers influenced by condition or threat.
q	Proportion of locations influenced by condition or threat.

although the exact formulation differs between methods. In scoring, the site value directly defines its priority (Turpie, 1995). In complementarity-based spatial prioritisation, other mechanisms are added to promote balance between biodiversity features.

2.1.2 | Common data types

As a starting point, we assume that data layers of the same type are independent, their values are identically distributed, and that, unless otherwise specified, there are no correlations between data layers of different type. We later explore deviations from this assumption.

Biodiversity features

Biodiversity features j include species, habitat types, ecosystem services, and other entities of conservation interest. Typically, biodiversity features constitute the benefits B_i of targeting a location i for conservation. Relative contributions of features to benefits may be altered using feature-specific weights w_j^B. Various methods are used to aggregate values s_{ij} across multiple features j to derive the conservation benefit at a location. A commonly used, simple approach is to calculate the weighted sum (score, index) across features:

$$B_i = \sum_{j=1}^{n_B} w_j^B s_{ij} \quad (2)$$

In this formulation weights do not necessarily sum to one. In more complicated formulations (not investigated here) benefit functions can be used to modify the treatment of features (e.g., Arponen, Heikkinen, Thomas, & Moilanen, 2005; Wu & Boggess, 1999).

We illustrate the influence of a single biodiversity data layer using a mathematical model based on the relative contribution each biodiversity feature makes to the numerical value of benefit obtained from a location. Working from Equation (2), the expected benefit from any site i can be generalised using the number of features included in the analysis (n_B), the average values of features at site i ($\bar{S}_i$) and the average weight ($\bar{w}_B$) given to the biodiversity features (detailed in Appendix S1):

$$B_i = \sum_{j=1}^{n_B} w_j^B s_{ij} \cong \bar{w}_B \bar{s}_i n_B \quad (3)$$

We further generalise this to a form representing expected average benefit across all locations i:

$$\bar{B} = E_i\left[B_i\right] = \bar{w}_B \bar{s} n_B, \quad (4)$$

where $\bar{s} = E_i[\bar{s}_i]$ is the expected average occurrence value across all features j and locations i.

Costs

The most commonly considered cost is acquisition, that is, the price of land/water area targeted for conservation. Other costs include transaction, management, and opportunity costs (Armsworth, 2014). Both single and multiple cost data layers may be used in spatial prioritisation and costs can be weighted just like biodiversity features, for example, when exploring different ways of accounting for different opportunity costs (Moilanen, Leathwick, & Quinn, 2011b). As with biodiversity (Equation 4), the aggregate cost value can be generalised to:

$$\bar{C} = E_i\left[C_i\right] = \bar{w}_C \bar{c} n_C \quad (5)$$

where n_C is the number of cost layers, $\bar{c} = E_i\left[\bar{c}_i\right]$ is average cost across all cost layers k and sites i, and $\bar{w}_c$ is the average weight given to a cost layer. Frequently, there would be only one cost layer, which simplifies this expression.

Habitat condition

Condition describes the degree of intactness (naturalness) of a habitat at a location. A common way to account for condition is to use site-specific condition (h_i) to multiply the site-specific values of either all features (if all features are affected similarly by condition) or some features (if features are affected differently) (Moilanen et al., 2011b). Multiplier values <1 are used to indicate reduced condition relative to a reference state, such as the natural state (Kujala, Whitehead, Morris, & Wintle, 2015). Condition multipliers h_i differ from feature-specific weights w_j^B in that the former vary spatially and the latter are per-feature multipliers. In some formulations, threat or accessibility of ecosystem services can be treated structurally in the same way as condition is treated here.

Using p to note the proportion of features affected by condition, we can express the expected benefit over all locations based on biodiversity and condition data as:

$$\overline{B^h} = \bar{B}\left(1-p\right) + \bar{B}p\bar{h} \tag{6}$$

where $\bar{B}$ is benefit ignoring condition, $\bar{h} = E_i[h_i]$ is the average condition multiplier across the landscape, and $\bar{B}p\bar{h}$ and $\bar{B}(1-p)$ are the aggregated benefits across features affected and not affected, respectively, by condition; this assumes species impacted by condition have similar properties to those not impacted. If habitat condition only affects a fraction q of the locations i, parameter p can be replaced by q in the equation above (see Equation 6b in Appendix S1). Here lower condition leads to reduction in benefits, however, in some cases it may be desirable to target degraded locations for, for example, restoration actions, in which case Equation 6 needs to be further modified (e.g., Moilanen et al., 2011b).

Threats

Threats typically describe the presence, intensity, and/or frequency of any (manageable or nonmanageable) pressure that threatens biodiversity, including habitat loss, pollution, invasive species, or overexploitation of stocks (Joppa et al., 2016). In spatial conservation planning, threats are typically used to guide conservation actions either towards (e.g., invasive species eradication) or away from threatened areas (e.g., when threat is nonmanageable). Several options exist to include threats in spatial prioritisation, the most suitable depending on objectives of analysis. First, threat layers can be included as features with positive weights (Equation 1a and 3), in which case priority is increased at locations with threat. Second, threat layers can be given negative weights, in which case they operate effectively as opportunity costs, modelling increased cost of threat management (Equation 1c). Finally, threats can also be used structurally the same way as a condition layer (Equation 6). If so, the interpretation is that habitat conditions (or species occurrence levels) are expected to become reduced in areas with an unmanageable threat. The impact of a single threat layer depends on which of these options is chosen. There also are examples where threat values have been inverted before use in analysis, which changes the direction and scaling of the effect in the above formulations (Whitehead et al., 2014).

2.1.3 | Relative influence of different data types

We explore the relative influence of a data layer to site value through a Jackknife analysis, that is, by adding the layer to an analysis and comparing the conservation values with and without the additional layer. We consider three scenarios:

(1) the 1st instance of adding a data layer to the analysis (e.g., with and without a single cost data layer);
(2) the nth instance of adding a data layer (e.g., with 67 and 68 biodiversity features);
(3) the values within a single data layer change on average by a fraction x, mimicking the typical situation where information underlying a data layer changes.

To illustrate scenario (3), we calculate the relative impact for each data type when x is 0.1 or 0.5, respectively.

We mark the value at site i at the starting situation as V_i^0, using any of the objective functions in Equation 1 to calculate site value. We then use V_i^1 to describe the altered site value after adding or changing a data layer. Then, the relative change in value at site i is:

$$\Delta V_i = \frac{V_i^1 - V_i^0}{V_i^0} = \frac{V_i^1}{V_i^0} - 1 \tag{7}$$

and the expected (average) change over all sites is:

$$\overline{\Delta V} = \mathrm{E}_i\left[\Delta V_i\right] = \mathrm{E}_i\left[\frac{V_i^1}{V_i^0}\right] - 1 \tag{8}$$

Not all combinations of data type and value function (Equation 1a–c) are common (e.g., adding the first biodiversity data layer), hence we restrict our analyses to representative cases. Table 2 gives equations for the relative influence a layer of a specific data type is likely to have on the site value of a location. The derivation of these equations is shown in Appendix S1.

2.1.4 | Site values vs priority ranks

Conservation decisions are often made based on rankings of site value, simply by selecting the most highly valued areas (i.e., scoring approach). Although newly added data may change the values of sites, this need not influence rankings. Earlier studies suggest that the influence of added data on conservation rankings of locations depends on both its correlation with, and the internal variation in values compared to, existing data (Figure 2). In classic cost-effectiveness analyses (Equation 1b), cost data becomes a dominant driver of priority ranks when costs are positively correlated with biodiversity values and have larger internal variation (Ferraro, 2003). In contrast, condition most affects priority ranks when negatively correlated with biodiversity occurrence (Equation 6). With biodiversity, largest changes in ranks are expected with large negative correlation between previous and new data and few features in the analysis (Figure 2).

2.2 | Simulation analysis

We verify the correctness of the above mathematical formulations using a simulated multifeature prioritisation (detailed in Appendix S2 and S3). We first generated sets of hypothetical species, with values ranging from 1 to 10, into a 20 × 20 grid using unconditional Gaussian simulation (R package "GSTAT" v.1.1-3). We summed the species layers to produce a baseline prioritisation (Equation 1a and 2), giving equal weights to all species. Next, we simulated three additional data layers with values between 1 and 10, but with a predefined

TABLE 2 Expected change in site value when data layers of different type are either added or changed. Columns give three scenarios of data change. Rows are for representative cases of data type and site value functions (Equation 1a–c). For each combination, the table gives the expected change in site value in location i as $\Delta V_i = \frac{V_i^1}{V_i^0} - 1$ (Equation 7), where V_i^1 refers to the value after data change. The expected change in value over the landscape would be $\overline{\Delta V_i} = E_i\left[\Delta V_i\right] = E_i\left[\frac{V_i^1}{V_i^0}\right] - 1$ (not shown)

Type of added/changed layer	V_i^1	1st layer added	nth layer added[a]	Fraction x change in one layer[a,b,c]
Biodiversity				
Equation (1a): $V_i = B_i$ **Equation (4):** $\bar{B} = \bar{w}_B \bar{s}_i n_B$			$\Delta V_i = \frac{\bar{w}_B \bar{s}_i (n_B+1)}{\bar{w}_B \bar{s}_i n_B} - 1 = \frac{1}{n_B}$	$\Delta V_i = \frac{\bar{w}_B \bar{s}_i (n_B-1) + \bar{w}_B \bar{s}_i (1+x)}{\bar{w}_B \bar{s}_i n_B} - 1 = \frac{\bar{w}_B \bar{s}_i (n_B+x)}{\bar{w}_B \bar{s}_i n_B} - 1 = \frac{x}{n_B}$
Equation (1b): $V_i = B_i / C_i$			$\Delta V_i = \frac{\bar{w}_B \bar{s}_i (n_B+1)}{C_i} / \frac{\bar{w}_B \bar{s}_i n_B}{C_i} - 1$ $= \frac{\bar{w}_B \bar{s}_i (n_B+1)}{C_i} \times \frac{C_i}{\bar{w}_B \bar{s}_i n_B} - 1 = \frac{1}{n_B}$	$\Delta V_i = \frac{\bar{w}_B \bar{s}_i (n_B-1) + \bar{w}_B \bar{s}_i (1+x)}{C_i} \times \frac{C_i}{\bar{w}_B \bar{s}_i n_B} - 1 = \frac{x}{n_B}$
Cost				
Equation (1b): $V_i = B_i / C_i$		$\Delta V_i = \frac{B_i}{\bar{w}_C c_i} \times \frac{1}{B_i} - 1$ $= \frac{1}{\bar{w}_C c_i} - 1$ Or $\left(\frac{1}{c_i} - 1\right)$ when no weight used[d]	$\Delta V_i = \frac{B_i}{\bar{w}_C \bar{c}_i (n_C+1)} \times \frac{\bar{w}_C \bar{c}_i n_C}{B_i} - 1$ $= \frac{n_C}{n_C+1} - 1 = -\frac{1}{n_C+1}$	$\Delta V_i = \frac{B_i}{\bar{w}_C \bar{c}_i (n_C-1) + \bar{w}_C \bar{c}_i (1+x)} \times \frac{\bar{w}_C \bar{c}_i n_C}{B_i} - 1$ $= \frac{B_i}{\bar{w}_C \bar{c}_i (n_C+x)} \times \frac{\bar{w}_C \bar{c}_i n_C}{B_i} - 1 = -\frac{x}{n_C+x}$
Equation (1c): $V_i = B_i - C_i$		$\Delta V_i = \frac{B_i - \bar{w}_C c_i}{B_i} - 1 = -\frac{\bar{w}_C c_i}{\bar{w}_B \bar{s}_i n_B}$	$\Delta V_i = \frac{B_i - \bar{w}_C \bar{c}_i (n_C+1)}{B_i - \bar{w}_C \bar{c}_i n_C} - 1$ $= \frac{1}{n_C - \frac{B_i}{\bar{w}_C \bar{c}_i}} = \frac{1}{n_C - \frac{\bar{w}_B \bar{s}_i n_B}{\bar{w}_C \bar{c}_i}}$ Approaches 1 when n_c=1 and w_c is large.	$\Delta V_i = \frac{B_i - \bar{w}_C \bar{c}_i (n_C-1) - \bar{w}_C \bar{c}_i (1+x)}{B_i - \bar{w}_C \bar{c}_i n_C} - 1$ $= \frac{B_i - \bar{w}_C \bar{c}_i (n_C+x)}{B_i - \bar{w}_C \bar{c}_i n_C} - 1 = \frac{x}{n_C - \frac{B_i}{\bar{w}_C \bar{c}_i}} = \frac{x}{n_C - \frac{\bar{w}_B \bar{s}_i n_B}{\bar{w}_C \bar{c}_i}}$
Condition (and structural analogues)				
Equation (1a): $V_i = B_i$ **Equation (6):** $B^h = B_i(1-p) + B_i p h_i$		$\Delta V_i = \frac{B_i(1-p) + B_i p h_i}{B_i} - 1 = p(h_i - 1)$	As on the left.[e]	Here $E\left[h_i^1\right] = (1+x)\,E\left[h_i^0\right]$, denoted $(1+x)\,h_i$ $\Delta V_i = \frac{B_i(1-p) + B_i p h_i (1+x)}{B_i(1-p) + B_i p h_i} - 1 = \frac{p h_i x}{1-p+p h_i} = \frac{x}{1 - \frac{1}{h_i} + \frac{1}{p h_i}}$

TABLE 2 (Continued)

Type of added/changed layer	V_i^1	1st layer added	*n*th layer added[a]	Fraction x change in one layer[a,b,c]
Equation (1b): $V_i = B_i/C_i$		$\Delta V_i = \frac{B_i(1-p)+B_i p h_i}{C_i} \times \frac{C_i}{B_i} - 1 = p(h_i - 1)$	As on the left.[e]	As above as C_i cancels out.
Equation (1c): $V_i = B_i - C_i$		$\Delta V_i = \frac{B_i(1-p)+B_i p h_i - C_i}{B_i - C_i} - 1 = \frac{p(h_i-1)}{1-\frac{C_i}{B_i}}$	As on the left.[e]	$\Delta V_i = \frac{B_i(1-p)+B_i p h_i(1+x) - C_i}{B_i(1-p)+B_i p h_i - C_i} - 1 = \frac{B_i p h_i x}{B_i(1-p)+B_i p h_i - C_i} = \frac{x}{1-\frac{1}{h_i}+\frac{1}{p h_i}-\frac{C_i}{B_i p h_i}}$

[a]Calculations assume the same expectations for n and $(n+1)$ layers. [b]The fraction change x should be interpreted as an average change in pixel values drawn from normal distribution, $X = N(0, \sigma)$. [c]A fraction x change in one data layer also implies an average change in fraction x for each location *i*. [d]Typically, no explicit weight would be used for the cost component in a B/C analyses because the weighting would not change the ranking of sites. [e]We assume s_{ij} values already include any previously added condition layers.

strong negative (Spearman's correlation coefficient, $\rho = -0.9$), neutral ($\rho \approx 0$) or strong positive ($\rho = 0.9$) spatial correlation with the baseline priority map. We then re-prioritised the original data by including each new data layer either as an equally weighted biodiversity feature (Equation 1a), cost (Equation 1b) or condition layer (Equation 6). We compared the expected average change in the site values ($E[\Delta V_i]$) (Equation 8) to observed changes ($O[\Delta V_i]$) and calculated the Spearman's rank correlation between the baseline and the new priority ranks of sites. Figure 3 illustrates the analysis steps with a smaller 5 × 5 grid.

In these simulations, adding the new data layer as a species represents the "add *n*th layer" scenario. The cost and condition examples represent the scenario in which the new type of layer enters analysis for the first time. Supporting Information (Appendix S2 and S3) includes further simulations of the mathematical expressions above, including selected variants of cost-effectiveness formulation (Equation 1b,c), variable numbers of biodiversity layers, and different scenarios of addition or change in a layer of specific type. All simulations were repeated 1,000 times for each added data layer to calculate a range of potential outcomes. All data layers were produced and simulations run using R (v.3.3.1).

2.3 | Use of data types in literature

To clarify the differences in the number of layers typically used for each data type, we draw on examples from scientific literature. Kullberg and Moilanen (2014) reported that only 39 out of 207 conservation prioritisation articles published between 2010 and 2012 used information on cost, condition or threats together with biodiversity data. These studies cover a variety of prioritisation approaches, from simple biodiversity scores (equivalent to Equation 1) to complex complementarity-based optimisations. We re-visit these 39 articles to record the number of data layers used for each data type, thereby providing an *a-priori* expectation of typical data layer counts.

3 | RESULTS

3.1 | The relative influence of data types and number of data layers

When the benefits of acting in a location are defined by the (weighted) summed values of biodiversity features present (Equation 1a), the impact of adding or changing any single biodiversity data layer depends on the total number of biodiversity features (Table 2, first row, "*n*th layer added"; Supporting Information Figure S1). Our literature survey indicates that the number of biodiversity features used in conservation analyses varies greatly, ranging from just a few to thousands of features, with a mean of 741 (Table 3). Considering Equations 1–3, it becomes evident that when the biodiversity value of a location is based on many (e.g., >100) features, adding one more feature is likely to have negligible impact on the site value

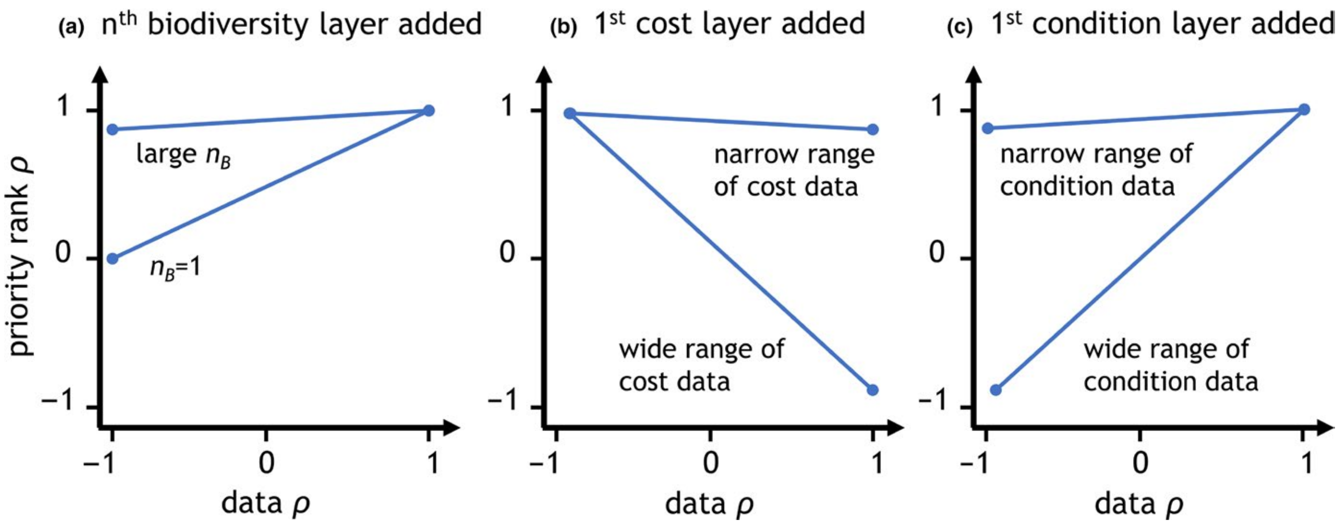


FIGURE 2 A schematic of how spatial correlation between old and new priority ranks (*Y*-axis) links to spatial correlation between new and old data (*X*-axis), when added data represents (a) biodiversity feature, (b) costs, or (c) condition. The direction of this relationship varies with data type. With costs and condition data, influence on rankings also depends on the relative range of values compared to biodiversity values (b and c). The effect of additional biodiversity data further depends on how many features are already included in the analysis (n_B)

(Supporting Information Figure S1), as the addition of one layer induces <1% change in information input (assuming the new layer is not given a high relative weight). Similarly, a change (update) of one layer has an even smaller impact than the addition (or removal) of one layer (Table 3).

The relative impact of new cost data also depends on the number of cost layers included, which usually is between zero and a few (Kullberg & Moilanen, 2014; Table 3). Therefore, n_C is typically orders of magnitude smaller than n_B, and the relative impact of new cost data to aggregate cost (Equation 5) is likely to be much larger than the impact of new biodiversity data to aggregate benefits (Equation 3; Table 2; Supporting Information Figure S2). This is particularly so for the cost-efficiency formulation *B*/*C*, in which biodiversity value, aggregated across potentially thousands of biodiversity features, is divided by a single or an aggregate of just a few cost values. Interestingly, the effects of the cost layer are independent from the biodiversity layers in the *B*/*C* formulation (Table 2, third row; Supporting Information Figure S2), implying that the impact of a single cost layer can equal the influence of all biodiversity layers put together.

Information on condition and threat are also typically based on few layers (Table 3). In the reviewed literature, analyses accounting for habitat condition and/or threat most often used only a single data layer to describe these aspects. The main difference between condition and cost data is that condition may have varying impact on spatial priorities depending on the fraction of features it influences (Equation 6; Supporting Information Figure S3). If a single condition layer is applied to all features, the impact is larger than when only some features are multiplied or when multiple condition layers are used for separate species groups, in which case some impacts may be counter balanced. The impact of a condition layer is also reduced if it impacts only a small part of the landscape.

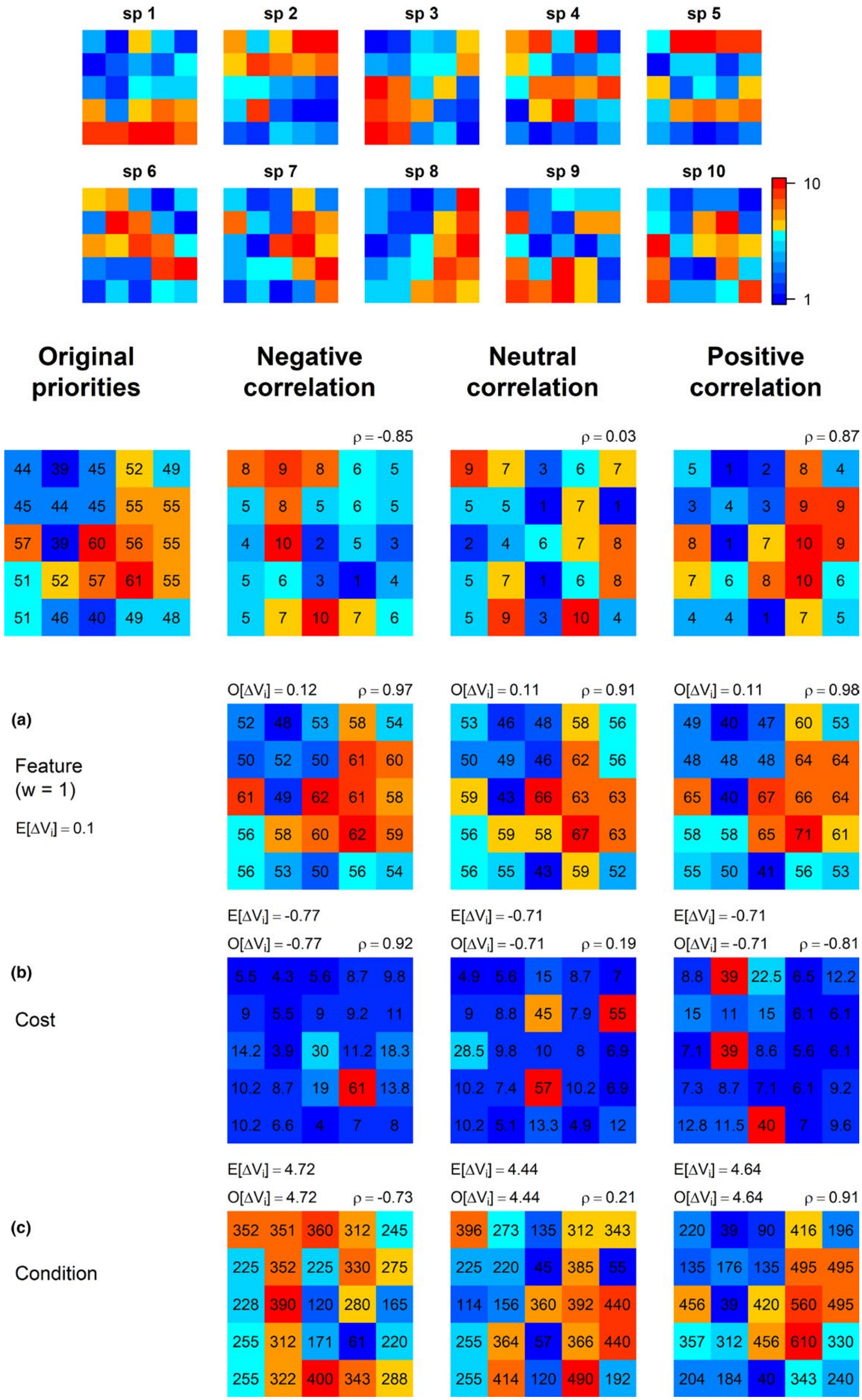
Threats may be accounted for in several ways, similar to the use of feature layers, costs or condition. The impact of a threat will consequently depend on the fractions of layers and landscape impacted. As the typical number of threat layers is small (Table 3), the expectation is that a single threat layer may have relatively high impact, especially if the layer is used structurally like cost or condition.

Simulation results in Supplementary material (Supporting Information Figures S1–S3) confirm that our mathematical formulae in Table 2 closely approximate changes observed in simulation experiments. This confirms both our mathematical derivation (Appendix S1) and the estimates of relative expected change in Table 2, above.

3.2 | Expected change in site value and spatial correlations between data types

Figure 3 illustrates how site value and priority are influenced by the addition of a single new layer of a specific data type. When an equally weighted biodiversity layer is added to a pool of 10 layers,

FIGURE 3 Illustration of the influence of an additional data layer on site values and priorities. The two top rows show distributions of ten simulated species. Row three shows a baseline prioritisation result (sum across the ten simulated distributions), and three additional data layers of random values with strongly positive, neutral or strongly negative correlation with the baseline values. The rows (a)–(c) show updated spatial priorities when the additional data layers have been added as (a) an equally weighted additional biodiversity feature, (b) cost (Equation 1b), or, (c) condition (Equation 6, p = 1). The numbers above each re-prioritisation give the expected ($E[\Delta V_i]$) and observed ($O[\Delta V_i]$) average change in site value between the two prioritisations, and the Spearman's rank correlation coefficient (ρ) between the original and new priority ranks. Note that except for the first two rows, colours are not comparable between plots but cell values are shown



sp 1
sp 2
sp 3
sp 4
sp 5
sp 6
sp 7
sp 8
sp 9
sp 10
10
1
Original priorities
Negative correlation
Neutral correlation
Positive correlation
44 39 45 52 49
45 44 45 55 55
57 39 60 56 55
51 52 57 61 55
51 46 40 49 48
ρ = -0.85
8 9 8 6 5
5 8 5 6 5
4 10 2 5 3
5 6 3 1 4
5 7 10 7 6
ρ = 0.03
9 7 3 6 7
5 5 1 7 1
2 4 6 7 8
5 7 1 6 8
5 9 3 10 4
ρ = 0.87
5 1 2 8 4
3 4 3 9 9
8 1 7 10 9
7 6 8 10 6
4 4 1 7 5
(a)
Feature (w = 1)
E[ΔVi] = 0.1
O[ΔVi] = 0.12 ρ = 0.97
52 48 53 58 54
50 52 50 61 60
61 49 62 61 58
56 58 60 62 59
56 53 50 56 54
O[ΔVi] = 0.11 ρ = 0.91
53 46 48 58 56
50 49 46 62 56
59 43 66 63 63
56 59 58 67 63
56 55 43 59 52
O[ΔVi] = 0.11 ρ = 0.98
49 40 47 60 53
48 48 48 64 64
65 40 67 66 64
58 58 65 71 61
55 50 41 56 53
(b)
Cost
E[ΔVi] = -0.77
O[ΔVi] = -0.77 ρ = 0.92
5.5 4.3 5.6 8.7 9.8
9 5.5 9 9.2 11
14.2 3.9 30 11.2 18.3
10.2 8.7 19 61 13.8
10.2 6.6 4 7 8
E[ΔVi] = -0.71
O[ΔVi] = -0.71 ρ = 0.19
4.9 5.6 15 8.7 7
9 8.8 45 7.9 55
28.5 9.8 10 8 6.9
10.2 7.4 57 10.2 6.9
10.2 5.1 13.3 4.9 12
E[ΔVi] = -0.71
O[ΔVi] = -0.71 ρ = -0.81
8.8 39 22.5 6.5 12.2
15 11 15 6.1 6.1
7.1 39 8.6 5.6 6.1
7.3 8.7 7.1 6.1 9.2
12.8 11.5 40 7 9.6
(c)
Condition
E[ΔVi] = 4.72
O[ΔVi] = 4.72 ρ = -0.73
352 351 360 312 245
225 352 225 330 275
228 390 120 280 165
255 312 171 61 220
255 322 400 343 288
E[ΔVi] = 4.44
O[ΔVi] = 4.44 ρ = 0.21
396 273 135 312 343
225 220 45 385 55
114 156 360 392 440
255 364 57 366 440
255 414 120 490 192
E[ΔVi] = 4.64
O[ΔVi] = 4.64 ρ = 0.91
220 39 90 416 196
135 176 135 495 495
456 39 420 560 495
357 312 456 610 330
204 184 40 343 240

TABLE 3 Numbers of layers used for each data type across 39 articles reviewed, and the relative influence of a change in a single data layer on the site value (V_i). The table gives the average, mode (median for biodiversity) and maximum number of layers used, and the proportion of studies that used only a single layer of that data type (1-only). The studies (n = count) using cost and threat are subdivided according to specific formulations. Relative impact is illustrated by calculating the proportional change in V_i using Equation (1) and (7), and the mean number of layers for each data type and formulation. We used an average value $p = 0.9$ as per literature and assumed an average condition value of 0.5 for h_i (with a typical range of 0–1). Relative impact shown only for data types that do not require information on layer values

	n	Typical number of data layers used				Change (%) in V_i when one data layer changes by	
Data type		Mean	Mode	Maximum	1-only (%)	10%	50%
Biodiversity	39	741.1	97[a]	6,078	0.0	0.01	0.08
Cost	23	2.1	1	5	39.1		
B/C (Equation 1b)	11	1.9	1	5	45.5	−5.0	−20.8
B-C (Equation 1c)	12	2.4	2	5	25.0	–	–
Condition	11	4.0	1	8	36.4	7.4	37.2
Threat	19	2.5	1	7	36.8		
B*T (Equation 6)	5	1.2	1	2	80.0	7.4	37.2
B-T (Equation 1c)	5	3.8	4	7	20.0	–	–
Postprocessing[b]	9	2.5	4	4	11.1		

[a]Median. [b]Threat was not used in the priority ranking itself but instead overlapped with priorities in a separate postprocessing analysis.

the expected average change in site value is $E\left[\Delta V_i\right] = \frac{1}{n_B} = \frac{1}{10} = 0.1$. (Table 2), closely matching the change observed ($O[\Delta V_i] = 0.11$–0.12). The approximately 10% change in site values produces a smaller change in priority rankings ($\rho = 0.91$–0.98). Variation in change is explained by spatial correlation between the new layer and original priorities: site values are on average altered most by new species that have a negative correlation (Figure 2; Supporting Information Figure S1).

As anticipated, adding a first cost or condition layer results in notably larger changes both in values and priority rankings. In our example, dividing the benefits with the first cost layer changes the site values on average by 70–80% and alters the priority rankings from relatively little ($\rho = 0.92$) to nearly opposite patterns ($\rho = -0.81$) in comparison to the original priorities. Our results support earlier notions (Ferraro, 2003; Naidoo et al., 2006) that when costs and biodiversity benefits are positively correlated, including costs strongly impacts priorities (Figure 3; Supporting Information Figure S4). Inclusion of the first condition layer caused a major (440%–470%) change in site values. Opposite to costs, the rankings of B and the product h^*B are almost the same when the spatial correlation between benefits and condition is positive ($\rho = 0.91$, Figure 3; Supporting Information Figure S4), and vice versa for negative correlation ($\rho = -0.72$).

In addition to the spatial correlation between data types, our simulations confirm that the influence of a data layer on priority ranks is further dependent on the relative range of values within the layer. Notably, high changes in ranks due to adding cost or condition only occur when the internal variation of cost or condition values reaches a certain level in relation to biodiversity values (Supporting Information Figure S4).

4 | DISCUSSION

There are many factors that shape conservation priority patterns. This work clarifies the major differences in influences of different data types on site value and priority ranks, with implications for data collection and analysis interpretation.

Several observations stand out. First, when scoring is used to rank candidate locations across many biodiversity features, concerns about the accuracy of a single feature layer are likely to be of less significance, given the minor effect one layer has on aggregate value. Second, costs, threats and habitat condition will typically have impacts orders of magnitude higher than those of individual biodiversity feature layers. In a classic cost-effectiveness analysis (*B/C*) the effect of a single cost layer can be equal to the joint effects of all biodiversity feature layers together. Cost and threat layers both typically have very low layer counts (Table 3), and given their high influence (Figure 3), more attention should be directed to their production and verification. Third, there are alternatives to how costs and threats enter analysis (see Methods) and it is important to understand the interpretations of these alternatives. Fourth, the impact of fractionally changing one biodiversity feature layer is on average smaller than the effect of adding one completely new layer. Fifth, effects of spatial correlation depend on data type.

Our results are aligned with previous studies showing how costs may in some situations override biodiversity in spatial prioritisations (Armsworth, 2014; Armsworth et al., 2017; Balmford et al., 2003; Bode et al., 2008; Ferraro, 2003; Leathwick et al., 2008; Naidoo et al., 2006). Our work provides a mathematical explanation and shows how the influence of cost layers significantly depends on the combination of costs and benefits in the objective function

(Equation 1). Highest impacts of costs are expected in the B/C cost-efficiency formulation. If, on the other hand, there are many cost layers in a B–C structure, the relative influence of each cost layer will be smaller. Finally, higher internal variation in a layer leads to greater impacts (Supporting Information Figure S4) (Ferraro, 2003; Naidoo & Ricketts, 2006).

It is one thing to point out that cost, threat, and condition layers should be of high quality and another to produce those layers. All these data types are dependent on human behaviour and preferences, which is a major source of uncertainty. Acquisition and opportunity costs and demand for ecosystem services can change together with land use and economic shifts (Armsworth, 2014; Arponen, Cabeza, Eklund, Kujala, & Lehtomäki, 2010). Despite the important role of threats in conservation resource allocation, a recent study found that reasonable quality global data were available for 14 threats only (Joppa et al., 2016). Armsworth (2014) summarises shortfalls in available cost data, and argues that current cost proxies do not correctly reflect true conservation spending. Linkage to human influence implies that cost, threat and condition layers may have a very short period of validity compared to, for example, geophysical data. The relatively low (*c.* 20%) usage of cost and threat information in spatial prioritisation studies identified by Kullberg and Moilanen (2014) may reflect poor availability of data or possibly also concerns about data quality. An operationally critical question is, that if cost, condition or threat data are highly uncertain, should they be included in or excluded from the analysis? Our results suggest these data types strongly influence the location of priority sites and are therefore critical components of spatial prioritisation. They also imply that if these data are sufficiently uncertain, they will bias results in a semirandom and counter-productive manner. From our results we cannot determine when uncertainty in data warrants their inclusion/exclusion and hence this needs to be assessed case-specifically. When faced with data uncertainty, one option is to replicate analyses both with and without costs/threats/condition so that their effects can be separated from those of biodiversity distribution. Value of Information analyses (Yokota & Thompson, 2004) can provide further insight on how data improvements are likely to affect the results, with respect to the effort of additional data collection. Although there currently exists no definition of adequately good data, it could be argued that improvements to data are not necessary if they will not change the decision at hand (Runge, Converse, & Lyons, 2011; Yokota & Thompson, 2004). Equally important is to develop more standardised data collection procedures for cost, condition, and threat information and to improve data sharing.

Compared to costs and threats, distributions of many species and habitats are primarily influenced by abiotic and biotic factors and only secondarily by human activities. Habitat mappings can be improved through more relevant predictor variables, and more species observations—particularly in poorly surveyed areas—can be collected to improve species distribution data. If a new species is given an extremely high weight, or if multiple data layers are changed simultaneously, for example, through updates in environmental layers used to produce species distribution models, the expected change in site values increases. Habitat condition information will have limited impact on spatial priorities if it already influences many SDMs directly or indirectly. On the other hand, binary (presence-absence) habitat distributions can change significantly when modified by additional habitat condition information. Hence, the quality of condition data is important when habitat categories (ecosystem, community, etc.) or species range maps are used in spatial prioritisation. Furthermore, administrative borders and land use restrictions can impact spatial solutions significantly. These data (land ownership, governmental borders, protected areas, etc.) were not considered here, as they usually are of comparatively higher quality and higher certainty.

Additional to data, priority rankings are determined by the relative differences between cells, which do not necessarily change when cell values themselves change. Conservation priorities are influenced by factors such as the objectives and preferences of the decision maker, species-specific targets, interdependencies between locations (e.g., complementarity, connectivity, resource flow), and the algorithms used in prioritisation. Thus, the interaction between methods, assumptions, and data need to be accounted for when interpreting analyses. Here we used a simple scoring algorithm that aggregates benefit across features. Methodologically, complementarity-based spatial prioritisation employs more complex algorithms to balance the solution between features (Ball et al., 2009; Cabeza, 2003; Margules & Pressey, 2000; Moilanen et al., 2005). This has the consequence that the solution becomes more sensitive to additional biodiversity data (Kujala, Moilanen, et al., 2017). However, if additional data is effectively the same as the existing data, little change in priorities should be expected (Di Minin & Moilanen, 2012; Kujala, Moilanen, et al., 2017). While understanding complex interactions between prioritisation options is beyond the scope of this study, present results are directly relevant for scoring methods and cost-effectiveness analysis—both common in conservation decision making.

We have clarified the sensitivity of spatial conservation planning to uncertainties in different types of data, which helps direct attention during data collection, analysis preparation and decision making. Researchers and conservation managers should aim to improve spatial data and modelling in the most relevant and cost-effective manner possible.

ACKNOWLEDGEMENTS

All authors were supported by Australian Research Centre Discovery Project grant DP160101003. A.M. was also supported by the Academy of Finland Centre of Excellence programme 2012–2017, grant 250444. We thank M. McCarthy and three anonymous reviewers for insightful comments and suggestions on the previous versions of this work.

AUTHOR'S CONTRIBUTIONS

All authors contributed to the formulation of the research question; H.K. and A.M. designed methodology with support from J.J.L.-M.

and J.E.; H.K. built and conducted the simulations and reviewed literature; H.K. and A.M. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Data produced in this work has been made available via figshare online repository at https://doi.org/10.26188/5b85dbbab4127 (Kujala, Lahoz-Monfort, Elith, & Moilanen, 2018).

ORCID

Heini Kujala http://orcid.org/0000-0001-9772-3202

José Joaquín Lahoz-Monfort http://orcid.org/0000-0002-0845-7035

Jane Elith http://orcid.org/0000-0002-8706-0326

REFERENCES

Armsworth, P. R. (2014). Inclusion of costs in conservation planning depends on limited datasets and hopeful assumptions. *Annals of the New York Academy of Sciences*, *1322*, 61–76. https://doi.org/10.1111/nyas.12455

Armsworth, P. R., Jackson, H. B., Cho, S.-H., Clark, M., Fargione, J. E., Iacona, G. D., … Sutton, N. A. (2017). Factoring economic costs into conservation planning may not improve agreement over priorities for protection. *Nature Communications*, *8*, 2253. https://doi.org/10.1038/s41467-017-02399-y

Arponen, A., Cabeza, M., Eklund, J., Kujala, H., & Lehtomäki, J. (2010). Costs of integrating economics and conservation planning. *Conservation Biology*, *24*, 1198–1204. https://doi.org/10.1111/j.1523-1739.2010.01539.x

Arponen, A., Heikkinen, R. K., Thomas, C. D., & Moilanen, A. (2005). The value of biodiversity in reserve selection: Representation, species weighting, and benefit functions. *Conservation Biology*, *19*, 2009–2014. https://doi.org/10.1111/j.1523-1739.2005.00218.x

Ball, I. R., Possingham, H. P., & Watts, M. (2009). Marxan and relatives: Software for spatial conservation prioritisation. In A. Moilanen, K. A. Wilson, & H. P. Possingham (Eds.), *Spatial conservation prioritization: Quantitative methods and computational tools* (pp. 185–195). Oxford, UK: Oxford University Press.

Balmford, A., Gaston, K. J., Blyth, S., James, A., & Kapos, V. (2003). Global variation in terrestrial conservation costs, conservation benefits, and unmet conservation needs. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 1046–1050. https://doi.org/10.1073/pnas.0236945100

Barry, S., & Elith, J. (2006). Error and uncertainty in habitat models. *Journal of Applied Ecology*, *43*, 413–423. https://doi.org/10.1111/j.1365-2664.2006.01136.x

Bode, M., Wilson, K. A., Brooks, T. M., Turner, W. R., Mittermeier, R. A., McBride, M. F., … Possingham, H. P. (2008). Cost-effective global conservation spending is robust to taxonomic group. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6498–6501. https://doi.org/10.1073/pnas.0710705105

Cabeza, M. (2003). Habitat loss and connectivity of reserve networks in probability approaches to reserve design. *Ecology Letters*, *6*, 665–672. https://doi.org/10.1046/j.1461-0248.2003.00475.x

Carwardine, J., Wilson, K. A., Hajkowicz, S. A., Smith, R. J., Klein, C. J., Watts, M., & Possingham, H. P. (2010). Conservation planning when costs are uncertain. *Conservation Biology*, *24*, 1529–1537. https://doi.org/10.1111/j.1523-1739.2010.01535.x

Di Minin, E., & Moilanen, A. (2012). Empirical evidence for reduced protection levels across biodiversity features from target-based conservation planning. *Biological Conservation*, *153*, 187–191. https://doi.org/10.1016/j.biocon.2012.04.015

Ferraro, P. J. (2003). Assigning priority to environmental policy interventions in a heterogeneous world. *Journal of Policy Analysis and Management*, *22*, 27–43. https://doi.org/10.1002/(ISSN)1520-6688

Franklin, J. (2013). Species distribution models in conservation biogeography: Developments and challenges. *Diversity and Distributions*, *19*, 1217–1223. https://doi.org/10.1111/ddi.12125

Grantham, H. S., Moilanen, A., Wilson, K. A., Pressey, R. L., Rebelo, T. G., & Possingham, H. P. (2008). Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters*, *1*, 190–198. https://doi.org/10.1111/j.1755-263X.2008.00029.x

Joppa, L. N., O'Connor, B., Visconti, P., Smith, C., Geldmann, J., Hoffmann, M., … Burgess, N. D. (2016). Filling in biodiversity threat gaps. *Science*, *352*, 416–418. https://doi.org/10.1126/science.aaf3565

Kujala, H., Lahoz-Monfort, J. J., Elith, J., & Moilanen, A. (2018). Not all data are equal: Influence of data type and amount in spatial conservation prioritisation - Data. figshare, https://doi.org/10.26188/5b85dbbab4127

Kujala, H., Moilanen, A., & Gordon, A. (2018). Spatial characteristics of species distributions as drivers in conservation prioritization. *Methods in Ecology and Evolution*, *9*, 1121–1132. https://doi.org/10.1111/2041-210X.12939

Kujala, H., Whitehead, A. L., Morris, W. K., & Wintle, B. A. (2015). Towards strategic offsetting of biodiversity loss using spatial prioritization concepts and tools: A case study on mining impacts in Australia. *Biological Conservation*, *192*, 513–521. https://doi.org/10.1016/j.biocon.2015.08.017

Kukkala, A. S., & Moilanen, A. (2013). Core concepts of spatial prioritisation in systematic conservation planning. *Biological reviews of the Cambridge Philosophical Society*, *88*, 443–464. https://doi.org/10.1111/brv.12008

Kullberg, P., & Moilanen, A. (2014). How do recent spatial biodiversity analyses support the convention on biological diversity in the expansion of the global conservation area network? *Natureza e Conservação*, *12*, 3–10. https://doi.org/10.4322/natcon.2014.002

Leathwick, J., Moilanen, A., Francis, M., Elith, J., Taylor, P., Julian, K., … Duffy, C. (2008). Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, *1*, 91–102. https://doi.org/10.1111/j.1755-263X.2008.00012.x

Lehtomäki, J., & Moilanen, A. (2013). Methods and workflow for spatial conservation prioritization using Zonation. *Environmental Modelling & Software*, *47*, 128–137. https://doi.org/10.1016/j.envsoft.2013.05.001

Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, *405*, 243–253. https://doi.org/10.1038/35012251

Moilanen, A., Anderson, B. J., Eigenbrod, F., Heinemeyer, A., Roy, D. B., Gillings, S., … Thomas, C. D. (2011a). Balancing alternative land uses in conservation prioritization. *Ecological Applications*, *21*, 1419–1426. https://doi.org/10.1890/10-1865.1

Moilanen, A., Franco, A. M. A., Early, R. I., Fox, R., Wintle, B., & Thomas, C. D. (2005). Prioritizing multiple-use landscapes for conservation: Methods for large multi-species planning problems. *Proceedings of the Royal Society B: Biological Sciences*, *272*, 1885–1891. https://doi.org/10.1098/rspb.2005.3164

Moilanen, A., Leathwick, J. R., & Quinn, J. M. (2011b). Spatial prioritization of conservation management. *Conservation Letters*, *4*, 383–393. https://doi.org/10.1111/j.1755-263X.2011.00190.x

Moilanen, A., Wilson, K. A., & Possingham, H. (2009). *Spatial conservation prioritization: Quantitative methods and computational tools*. Oxford: Oxford University Press.

Naidoo, R., Balmford, A., Ferraro, P. J., Polasky, S., Ricketts, T. H., & Rouget, M. (2006). Integrating economic costs into conservation planning. *Trends in Ecology & Evolution*, *21*, 681–687. https://doi.org/10.1016/j.tree.2006.10.003

Naidoo, R., & Ricketts, T. H. (2006). Mapping the economic costs and benefits of conservation. *PLoS Biology*, *4*, 2153–2164.

Possingham, H. P. (2001). *The business of biodiversity: Applying decision theory principles to nature conservation*. Fitzroy, Vic.: Australian Conservation Foundation.

Rodrigues, A. S. L., & Brooks, T. M. (2007). Shortcuts for biodiversity conservation planning: The effectiveness of surrogates. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 713–737. https://doi.org/10.1146/annurev.ecolsys.38.091206.095737

Runge, M. C., Converse, S. J., & Lyons, J. E. (2011). Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. *Biological Conservation*, *144*, 1214–1223. https://doi.org/10.1016/j.biocon.2010.12.020

Turpie, J. K. (1995). Prioritizing South African estuaries for conservation: A practical example using waterbirds. *Biological Conservation*, *74*, 175–185. https://doi.org/10.1016/0006-3207(95)00028-3

Visconti, P., Di Marco, M., Álvarez-Romero, J. G., Januchowski-Hartley, S. R., Pressey, R. L., Weeks, R., & Rondinini, C. (2013). Effects of errors and gaps in spatial data sets on assessment of conservation progress. *Conservation Biology*, *27*, 1000–1010.

Whitehead, A. L., Kujala, H., Ives, C. D., Gordon, A., Lentini, P. E., Wintle, B. A., ... Raymond, C. M. (2014). Integrating biological and social values when prioritizing places for biodiversity conservation. *Conservation Biology*, *28*, 992–1003. https://doi.org/10.1111/cobi.12257

Wilson, K. A., Westphal, M.I., Possingham, H.P., & Elith, J. (2005). Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation*, *122*, 99–112. https://doi.org/10.1016/j.biocon.2004.07.004

Wilson, K. A., Underwood, E. C., Morrison, S. A., Klausmeyer, K. R., Murdoch, W. W., Reyers, B., ... Possingham, H. P. (2007). Conserving biodiversity efficiently: What to do, where, and when. *PLoS Biology*, *5*, 1850–1861.

Wu, J., & Boggess, W. G. (1999). The optimal allocation of conservation funds. *Journal of Environmental Economics and Management*, *38*, 302–321. https://doi.org/10.1006/jeem.1999.1091

Yokota, F., & Thompson, K. M. (2004). Value of information literature analysis: A review of applications in health risk management. *Medical Decision Making*, *24*, 287–298. https://doi.org/10.1177/0272989X04263157

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Kujala H, Lahoz-Monfort JJ, Elith J, Moilanen A. Not all data are equal: Influence of data type and amount in spatial conservation prioritisation. *Methods Ecol Evol*. 2018;9:2249–2261. https://doi.org/10.1111/2041-210X.13084