Methods and workflow for spatial conservation prioritization using Zonation

Lehtomäki, Joona

2013-05-05


http://hdl.handle.net/10138/42748
https://doi.org/10.1016/j.envsoft.2013.05.001

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.
Methods and workflow for spatial conservation prioritization using Zonation

Joona Lehtomäki a,b,*, Atte Moilanen a,1

a Department of Biosciences, P.O. Box 65 (Viikinkaari 1), FI-00014 University of Helsinki, Finland
b Finnish Environment Institute, Natural Environment Centre, P.O. Box 140 (Mechelininkatu 34a), FI-00251 Helsinki, Finland

A R T I C L E   I N F O

Article history:
Received 25 January 2013
Received in revised form 3 May 2013
Accepted 5 May 2013
Available online

Keywords:
Conservation planning
Decision analysis
Decision support tool
Spatial conservation prioritization
Systematic conservation planning
Zonation software

A B S T R A C T

Spatial conservation prioritization concerns the effective allocation of conservation action. Its stages include development of an ecologically based model of conservation value, data pre-processing, spatial prioritization analysis, and interpretation of results for conservation action. Here we investigate the details of each stage for analyses done using the Zonation prioritization framework. While there is much literature about analytical methods implemented in Zonation, there is only scattered information available about what happens before and after the computational analysis. Here we fill this information gap by summarizing the pre-analysis and post-analysis stages of the Zonation framework. Concerning the entire process, we summarize the full workflow and list examples of operational best-case, worst-case, and typical scenarios for each analysis stage. We discuss resources needed in different analysis stages. We also discuss benefits, disadvantages, and risks involved in the application of spatial prioritization from the perspective of different stakeholders. Concerning pre-analysis stages, we explain the development of the ecological model and discuss the setting of priority weights and connectivity responses. We also explain practical aspects of data pre-processing and the post-processing interpretation of results for different conservation objectives. This work facilitates well-informed design and application of Zonation analyses for the purpose of spatial conservation planning. It should be useful for both scientists working on conservation related research as well as for practitioners looking for useful tools for conservation resource allocation.

© 2013 The Authors. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Conservation prioritization is about decision support for conservation planning (Ferrier and Wintle, 2009). It aims to answer questions about when, where, and how we can efficiently achieve conservation goals (Pressey et al., 2007; Wilson et al., 2007). Spatial conservation prioritization utilizes computational tools and analyses that are relevant for ecologically informed spatial allocation of conservation actions or placement of other land uses (Kukkala and Moilanen, 2012). Methods of spatial prioritization evolved starting from simple complementarity-based minimum set reserve selection algorithms that operated on relatively small data sets and presence-absence data (reviewed by Sarkar et al. (2006)). More recently, methods have become able to accommodate various cost factors and much increased ecological realism by implementing, for example, methods to deal with species-specific connectivity and uncertainty, and software implementations have become able to deal with much larger landscapes and a variety of data types (Kukkala and Moilanen, 2012).

Spatial conservation prioritisation is a form of conservation assessment (sensu Knight et al., 2006) which can be utilized as a technical phase inside the broader operational model of systematic conservation planning (SCP) that focuses on planning, implementing, and monitoring conservation (Margules and Pressey, 2000; Margules and Sarkar, 2007; Pressey and Bottrill, 2008; Kukkala and Moilanen, 2012). In this study, we concentrate on the interface between spatial conservation prioritization and implementation-oriented conservation planning, specifically in the context of the Zonation spatial planning software2 (Moilanen et al., 2005, 2009b).

* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
* Corresponding author. Department of Biosciences, P.O. Box 65 (Viikinkaari 1), FI-00014 University of Helsinki, Finland. Tel.: +358 9 191 57714.
E-mail addresses: joona.lehtomaki@helsinki.fi (J. Lehtomaki), atte.moilanen@helsinki.fi (A. Moilanen).
1 Tel.: +358 9 191 57753.

2 Zonation version 3 is freely available from http://cbig.it.helsinki.fi/software/zonation/ for MS Windows.
There are several categories of factors that can introduce their own nuances into prioritization problems (Moilanen et al., 2009c). Details of the planning problem depend on the type of conservation action considered, including protection, management, maintenance, and restoration of habitats (Pressey et al., 2007; Wilson et al., 2009). A reserve network could be planned for immediate implementation or for recurrent yearly operations (Costello and Polasky, 2004; Pressey et al., 2007). Biodiversity could be considered from the perspective of representation in a reserve network or from the perspective of landscape-wide retention, which involves potential threats and opportunities both within reserves and the surrounding landscape (Pressey et al., 2004). The level of detail included in the ecological model that explicitly or implicitly underlies the decision-making influences the difficulty of implementing a decision analysis (Possingham et al., 2000; Wilson et al., 2009).

Spatial conservation prioritization is usually done within a wider decision-making context in which the needs of many land users and stakeholders are acknowledged (Ferrier and Wintle, 2009). At the outset of any planning process, it is crucial that objectives (aims, goals) are explicitly set for all of the processes and criteria involved (Ferrier and Wintle, 2009; Runge et al., 2011). This also includes the explicit consideration of which decision-support tool is most suitable for the task at hand which can involve integrating large biological and socio-economic datasets as well as several software tools (Segan et al., 2011). Furthermore, setting of objectives (Opdam et al., 2002; Wilson et al., 2009), stakeholder involvement (Knight et al., 2006), policy recommendations (Sutherland et al., 2006), quality verification (Langford et al., 2011), and monitoring (Lindenmayer and Likens, 2009) are all stages that may be repeated over periods of many years.

Here we describe a workflow for running a conservation prioritization analysis with the Zonation software (see Moilanen et al., 2005, 2009b, 2011a, 2012; Moilanen and Arponen, 2011 for references). Zonation has been applied across terrestrial, riverine, marine, and urban environments (Leathwick et al., 2008; Moilanen et al., 2008; Gordon et al., 2009). It includes a set of useful analysis features, including uncertainty analysis and seven ways of dealing with connectivity (Section 2.4). In can operate on species, ecosystems (Kremen et al., 2008; Lehtomäki et al., 2009), ecosystem services (Moilanen et al., 2011a; Thomas et al., 2012), or any such biodiversity feature, and can be applied to landscapes up to tens of millions of elements (grid cells) of biodiversity feature data (Arponen et al., 2012). In addition to target-based planning (Carwardine et al., 2009), Zonation includes multiple alternative ways of aggregating conservation value across species and space (Moilanen, 2007).

As with any sophisticated tool, using Zonation requires both conceptual understanding about analysis options as well as experience and knowledge on how to establish a sensible workflow, which can be a major obstacle in the use of Zonation, due to the many analysis options available. While the analytical features available in Zonation are well documented, there is a scarcity of accessible information about what should happen before and after the computational analysis itself. Here we summarize previously scattered information about the process of implementing spatial conservation prioritization with Zonation. We explain all parts of the typical workflow, concentrating on what happens before and after the computational analysis itself. While the present work is most relevant for Zonation, much of the workflow should be relevant for any method and software that is applied for spatial prioritization.

2. Methods

2.1. Zonation: main concepts, algorithms, and outputs

Because the details of the Zonation software and its algorithms have been extensively documented elsewhere (Moilanen et al., 2005, 2009b, 2011a, 2012), we summarize only the features that facilitate understanding of the present material, including interpretation of output (Section 2.7). Zonation develops a priority ranking of the entire landscape. It starts from the assumption that protecting everything would be best for conservation. If then proceeds to iteratively rank sites, at each step removing the spatial unit (grid cell, planning unit) that leads to the smallest aggregate marginal loss in biodiversity. In this process, which is called the Zonation meta-algorithm, the least useful sites receive the lowest ranks (close to 0) and areas most valuable for biodiversity receive the highest ranks (close to 1). This ranking is not unique: which is within the top 5%, which in the top 25%, which in the top 50%, and so on. It is possible to identify any given top fraction or bottom fraction of the landscape in terms of perceived conservation value from this ranking, which can be visualized as a priority rank map with different colours indicating rank values (see inset in Fig. 1 and Moilanen et al., 2012). These curves quantify the proportion of the original occurrences remaining for each feature when successively smaller fractions of the landscape remain for conservation (it is implicitly assumed that all unprotected sites are lost from conservation). Performance curves are most often investigated as averaged across all features or across a small number of subgroups of features. It is also informative to investigate the minimum value across all features or subgroups as it will show the situation of the worst-off feature when a given fraction of the landscape remains for conservation. Individual performance curves are not always useful as there can be up to tens of thousands of features in the analysis.

The main principle of the computational strategy of Zonation can be summarized as seeking to maximise retention of weighted range-size corrected feature richness (Moilanen et al., 2011a). A key to the operation of Zonation is the definition of marginal loss of biodiversity inside the Zonation meta-algorithm. For this loss there are multiple alternative definitions, which allow various concepts of conservation value, including those that emphasize species richness (the additive benefit function formulation, ABF) or rarity (core-area Zonation, CAZ) to variable degrees (Moilanen, 2007; Moilanen et al., 2011a). In fact, one of the first choices faced when initiating analysis is between ABF and CAZ. ABF produces high return on investment (Laitila and Moilanen, 2012), but may allow lower representation levels for features occurring in species-poor or expensive parts of the landscape (Moilanen, 2007). ABF is appropriate when the data is considered to be a surrogate for biodiversity broadly. CAZ aims to ensure high-quality locations for all features (Moilanen et al., 2005, 2011a), which may result in a lower return on investment because relatively higher effort must be expended on features that occur in species-poor or expensive areas. CAZ is more appropriate when an analysis focuses primarily on represent species. In this paper, we use Zonation as a common tool-based abstraction of the process.

2.2. The analysis framework

Fig. 1 summarizes the stages of a typical spatial conservation prioritization project using Zonation. Many of the stages are not Zonation-specific, and other analytical tools could be introduced into the process with small structural changes in analysis flow. It is worth noting that conservation prioritization is only one part of an operational model for conservation planning (Knight et al., 2006), and there are further successful conservation action, effective conservation implementation and management strategies are also needed.

The first step is setting conservation objectives and assessing whether the particular objectives require spatial conservation prioritization. Questions that can be addressed with Zonation are summarized in Section 2.7 (interpretation of results). The second stage is preparation of an ecologically based model of conservation value (Section 2.3) that must be informative for the objectives of the overall study. Often, the preparation of the ecological model requires the setting of weights and connectivity responses for biodiversity features (Section 2.4). Ideally, the ecological model would be developed based primarily on ecological data describing the distribution and state of biodiversity coupled with a good understanding of species' autoecology and anthropogenic factors such as conservation preferences. However, in reality the model must rely on data that is available, and the preparation of the ecological model goes hand in hand with the preparation of data. Section 2.5 summarizes factors relevant for data pre-processing.

After the objectives of the prioritization have been defined, and the ecological model and corresponding data prepared, it is possible to initiate spatial analysis. To understand how different analysis options influence results, it is important to develop the analysis in stages of increasing complexity (Section 2.6). At a more practical level, awareness of analysis options feeds back into formulation of the ecological model and into data processing—it is useless to plan for an analysis that cannot be executed. The third major stage is verification and interpretation of results (Section 2.7). To conclude, we discuss factors that may influence the full planning and analysis process, advantages and disadvantages perceived by stakeholders (Section 2.8), and resources needed by the different process stages (Section 2.9).
same time (Knight et al., 2006). When this is the case, observations from monitoring and any new information can feedback back into objective setting and data preparation for successive iterations of analysis and action.

The successful execution of each of the stages given in Fig. 1 depends on factors such as availability of data and whether results of a prioritization analysis are relevant for the planning case at hand. Table 1 outlines best-case and worst-case scenarios for the various stages of the whole spatial conservation prioritization process. Best and worst case scenarios bound extremes that rarely occur in real-life. Table 1 also includes information for a likely real-world scenario about difficulties encountered during a prioritization project. If several analysis components initially fall into the worst-case category, the whole prioritization may turn out infeasible in practice.

2.3. The ecological model

Conceptually, an ecologically based model of conservation value forms the foundation of spatial conservation prioritization. With the ecological model we refer to the entire set of data, weights or targets set to the feature data layers, and analysis options (e.g. connectivity) that are used to build an analysis that produces output relevant for the planning case at hand. Table 1 outlines best-case and worst-case scenarios for the various stages of the whole spatial conservation prioritization process. Best and worst case scenarios bound extremes that rarely occur in real-life. Table 1 also includes information for a likely real-world scenario about difficulties encountered during a prioritization project. If several analysis components initially fall into the worst-case category, the whole prioritization may turn out infeasible in practice.

2.3. The ecological model

Conceptually, an ecologically based model of conservation value forms the foundation of spatial conservation prioritization. With the ecological model we refer to the entire set of data, weights or targets set to the feature data layers, and analysis options (e.g. connectivity) that are used to build an analysis that produces output relevant for the planning case at hand. Thus the ecological model encompasses both the model that is used to produce the input data layers as well as analysis details. The complexity of the model depends on the availability of data and the overall objectives set in the first stage. The model can be relatively simple if, for example, the objective is to increase the population size of a focal species. Alternatively, an objective that involves satisfying species-level representation targets for a given set of species or aiming for a balanced representation across all habitat types that occur in the planning region while accounting for habitat condition and pair-wise similarity between habitats will require a more complex ecological model. It is worth mentioning that in most common cases of spatial conservation prioritization with Zonation, the ecological model implicitly assumes a static landscape (although it is possible to mimic temporal dynamics as described in context of climate change in Section 2.7).

Ideally, the ecological model would be based on distributions and expected persistence of all species occurring in the region. In reality such data is never available, and the conceptual formulation of the ecological model is constrained by the availability of data. Frequently some serviceable (spatial) data, such as observed or modelled species distributions, already exist. The question then becomes how to best utilize this information in an analysis. Here, surrogacy relationships between feature groups can be relevant, especially if available data is taxonomically biased (Kremen et al., 2008).

Overall, development of the ecological model can be an iterative process in which data is investigated and either rejected or incorporated into the analysis procedure, or further information is collected to fill particular gaps. The ecological model could include a range of different components describing biodiversity features, such as species, habitats, environments, various ecosystem services, carbon sequestration, ecosystem processes, genes, etc. These features are entered as individual entities into an analysis. Features can be given differential weights and possibly connectivity responses (Section 2.4); the development of features across time could also be accounted for (Thomson et al., 2009; Carroll et al., 2010). Frequently, one would also include additional ecologically relevant considerations such as connectivity (Section 2.4), habitat condition (Leathwick et al., 2010; Moilanen et al., 2011b), uncertainty of inputs (Moilanen et al., 2006), the effects of conservation actions (Moilanen et al., 2011a), or the ecological similarities between nominally different habitat types (Lehtomäki et al., 2009).

In addition to ecological factors, one could also be concerned about costs, needs of alternative land uses, and priorities that vary between administrative regions (Moilanen and Arponen, 2011; Moilanen et al., 2011a). A basic Zonation analysis is done using grid cells as selection units, but, if relevant, analysis could also be based on planning units defined via land ownership or hydrological catchment division (Moilanen et al., 2008; Leathwick et al., 2010). The analysis structure, features, weights of features, and additional considerations should be developed by the joint effort of a team of experts with relevant knowledge of ecology, available data, and socio-political constraints. At this stage, understanding of the analysis features available in Zonation is also highly useful. The latest Zonation V3.1 manual (Moilanen et al., 2012) is a useful source of information about available Zonation features and how to implement different types of analyses.

2.4. Weights and connectivity settings

Priority weights assigned to features influence the balance between features in the prioritization solution. In Zonation, weights for features always need to be...
decided. The simplest starting point is to treat everything as equal and give all features the same weight ($w_f = 1.0$). However, there are good reasons to assign features variable weights; for example, one might wish to assign endemic species elevated priorities.

Next we summarize a two-stage process for assigning weights. We assume a common case where input data can be divided into different groups, such as layers for birds, insects, habitat types, alternative land uses, connectivity layers corresponding to habitats, etc. Weights are first allocated as relative measures within each group, and in the second stage weights are balanced between groups. Note that a small range size is not a reason for assigning an elevated weight since the aggregation of conservation value inside Zonation applies successive range-size renormalization for all features (Moilanen et al., 2005; Moilanen et al., 2011a), implying that features with a small range receive elevated priority already. Also, the construction of the algorithm adds emphasis in narrow range features.

Factors that influence priority weights include anything that could influence target-setting in systematic conservation planning (Carwardine et al., 2009), and it should be noted that weight-setting is not an exact science as subjective valuation is involved. While weight-setting can be made arbitrarily complex, it should be noted that the initial stage of the Zonation algorithm is such that a sensible and efficient balance between features is obtained even with the use of default weights (Moilanen et al., 2005).

### Table 1

<table>
<thead>
<tr>
<th>Component</th>
<th>Best-case</th>
<th>Typical case</th>
<th>Worst-case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objectives</td>
<td>Clear, quantitative, and measurable.</td>
<td>Defined, but for different purposes. Only partly quantitative and measurable.</td>
<td>Unspecific and poorly measurable.</td>
</tr>
<tr>
<td>Preparation of the ecological model</td>
<td>Ecology well understood, representative data.</td>
<td>Ecological background partly understood, some data is available.</td>
<td>Source of conservation value poorly understood.</td>
</tr>
<tr>
<td>Individual data sets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Representativeness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability Format</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data as a whole</td>
<td>Relevant and fully adequate for the purpose.</td>
<td>Relevant, but partly inadequate for the most relevant analysis. Deficiencies must be accounted for in interpretation of results.</td>
<td></td>
</tr>
<tr>
<td>Spatial prioritization</td>
<td>Human and computational resources adequate.</td>
<td>Some human resources available, but will require collaboration. Some computational resources available, but may restrict analysis capability.</td>
<td>Competent analysts unavailable. Lack of computing resources reduces quality of analysis possible.</td>
</tr>
<tr>
<td>Post-processing</td>
<td>No technical problems: quickly completed.</td>
<td>Time consuming and slow the first time, but faster the next time done.</td>
<td>Lack of understanding of options: delays due to technical difficulties.</td>
</tr>
<tr>
<td>Verification</td>
<td>No technical problems.</td>
<td>Mostly good and technically correct results, but part of analyses will have to be redone.</td>
<td>Analysis failed due to data or analysis setup: needs to be redone, possibly multiple times.</td>
</tr>
<tr>
<td>Recommendations</td>
<td>Corresponds to objectives.</td>
<td>Corresponds to most objectives.</td>
<td>Poorly reflect objectives.</td>
</tr>
<tr>
<td>On-the-ground verification</td>
<td>Confirms the conservation relevance.</td>
<td>Not done.</td>
<td>Expected conservation value not found.</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Resources available, confirms conservation success.</td>
<td>Insufficient resources available and usually done for other purposes.</td>
<td>Resources unavailable and/or recommendations proven inadequate.</td>
</tr>
</tbody>
</table>

At this stage weights have been allocated to all features, but there may be a second stage in which weights are balanced between feature groups, assuming several exist. For example, consider a data set consisting of categorical distribution data for 30 habitat types and distribution data for 100 birds and 200 insect species. Assuming that all weights are equal ($w_f = 1.0$ for all), data groups would have aggregate weights of 30 for habitats, 100 for birds, and 200 for insects, leading to insects having the greatest influence on analysis outcomes. A more flexible way to set up the weighting is to assign relative weights to each group. One could consider the habitat type group as the most fundamental and assign this group an aggregate weight of 30. Birds could be the second most important, but less so than habitats, and thus one could assign birds an aggregate weight of 20. Data for insects could be viewed as somewhat supplementary and perhaps unreliable, and one could give them an aggregate weight of 10, thereby implying that the weight of one insect should in fact be 1/20 of that of one habitat type—a major departure from having everything equal. Next, the weights of individual layers are rescaled so that they sum up to the aggregate weights assigned to groups. More formally, let feature $j$ belong to group $b$, and let $M_b$ and $N_b$ be the numbers of multiplicative and additive weighting factors corresponding to feature $j$, and let $N_i$ and $N_b$ be the numbers of multiplicative and additive weighting factors, respectively. The initial relative weight of feature $j$ becomes:

$$ r_f^j = \frac{N_b}{\sum_i N_i} \prod_{i=1}^{M_b} \sum_{j=1}^{N_b} A_{ij} $$

A spatial analysis using Zonation means accounting for different responses the analysis features may have on given spatial scales. The most common responses to be considered are connectivity and fragmentation: selection units (grid cells, planning units) have a particular location in space and in the analysis the attributes of a selection unit are influenced by attributes of the surrounding selection units at
specified scales. Specification of connectivity and/or fragmentation responses is another task that is almost always encountered when Zonation analyses are set up. Without these, Zonation analyses are only implicitly spatial: selection units have a location in the spatial domain, but the ranking of a selection unit is only influenced by what is actually present at that particular location.

In Zonation analyses it is a common strategy to enter distributions of species or habitats into the analysis as layers representing local habitat quality. Then, to account for connectivity, the local habitat quality layers can be entered into the same analysis—a second time with connectivity transformations applied to them (note that this option does not do this automatically and it is up to the user to adjust the input files accordingly). The analysis now includes both local habitat quality and feature-specific connectivity, two fundamental components of spatial ecology. These components usually share the same overall weighting scheme, but a multiplier can be used to scale the relative priority given to connectivity layers as compared to local habitat quality. Commonly, this connectivity multiplier would get a value of 0.5, reflecting a view that habitat quality is more fundamental than connectivity, as without habitat quality there is no connectivity. Note that the connectivity multiplier is independent from the actual method used to perform the connectivity transformations (Table 2). The overall weighting scheme has to be developed case by case to correspond to the particulars of the planning need and objectives. However, in most typical cases, only a couple of factors would be considered in the weight setting. Here, multiple potentially relevant factors have been listed for completeness; we do not imply that weight-setting should be as complicated as described above.

Table 2 summarizes methods for incorporating connectivity or accounting for species’ sensitivity to fragmentation presently available in Zonation. Except for the BQP and NQP, these methods can in principle be combined in the same analysis, but this will complicate interpretation of the analysis outcome. Note that an analysis option called “edge removal” is enabled by default in Zonation; this will add a “spatial” component to the analysis even without any explicit connectivity options used. As the name suggests, with this option Zonation removes cells only from the edges of the remaining analysis area, thereby speeding up the cell removal process. This effect is minor, but it will introduce some spatial patterns into the priority rank maps. Note also, that a given biodiversity feature may have connectivity responses on multiple spatial scales, which can be accounted for by entering different connectivity components into the analysis. For example, Rayfield et al. (2009) had different connectivity scales for home-range scale habitat use and for juvenile dispersal of the American marten.

### 2.5. Pre-processing of data

Data pre-processing can easily be the most resource-hungry phase of a conservation prioritization project (Table 3), except for the on-the-ground implementation phase (Gibbons et al., 2011). Fig. 2 summarizes stages of data pre-processing. If the data have already been collated, pre-processed, validated, and converted into a suitable format, then the spatial prioritization analysis can commence very rapidly. This, however, is rarely the case, and extensive pre-processing is frequently required. Data pre-processing starts with the acquisition of the relevant primary data (stage 1, Fig. 2). While data can originate from direct observations, the generally poor availability of extensive and systematic observational data often makes it necessary to rely on surrogate data sets (Elith and Leathwick, 2009). Distributions of habitats may be derived from remotely sensed data. Remotely sensed environmental variables, such as temperature and rainfall, can be used as such or as explanatory variables in species distribution modelling (SDM) to relate species distributional data to a set of geographic and/or environmental predictors (Elith and Leathwick, 2009).

Regardless of the source of the data, it needs to be stored in an orderly fashion to facilitate maintenance and access to it (stage 2, Fig. 2). Data are normally held in a database which frequently is just a collection of plain files organized into folders. The database does not have to be a local one, as high-quality data sets are increasingly becoming available online (Jetti et al., 2012).

Next, the desired subset of the data is selected and extracted from the whole database. This subset enters the actual pre-processing stage (stage 3, Fig. 2), which may include a large variety of geospatial processing steps. Up until this stage the data can be in either vector or raster format, but Zonation requires that all input data are formatted as raster grids. A typical pre-processing task is therefore conversion from vector to raster format. All the raster maps need to have the same spatial extent and resolution (i.e. all the raster maps are perfectly aligned) which needs to be taken into account while doing the conversion.

When converting from vector to raster format, it is better to use as fine a resolution as is relevant for the planning problem at hand and permitted by the raw data, because prioritzation results are sensitive to the resolution used—low resolution degrades the utility of the analysis (Arponen et al., 2012). Due to computational limitations, high-resolution raster maps often need to be aggregated to a lower resolution, which may not be a trivial task. When aggregating cells into a lower resolution, special attention needs to be paid to which function is used to assign the value to the lower resolution aggregate cell. Zonation can operate with various types of input features and the selection of the function may depend on the exact nature of the data being aggregated. According to our experience, for the most common data types (probability of occurrence, coverage of habitat type) summing the higher resolution cells that will comprise the lower resolution cell is appropriate. For binary presence-absence data (either species or habitats) summing the number of higher resolution cells equals to the number of occurrences within the aggregate cell. While aggregation loses the exact spatial location information of the high-resolution cells, information about the quality of the high-resolution cells is retained within the

### Table 2

Summary of connectivity methods available in Zonation version 3. Speed is as compared to a run that does not apply connectivity; speed differences will depend on the particulars of the data.

<table>
<thead>
<tr>
<th>Method</th>
<th>Speed</th>
<th>Feature-specific</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning units</td>
<td>Faster, depending on planning unit size</td>
<td>No – affects all features</td>
<td>Not an actual connectivity method, but employing planning units will introduce aggregation of rank priorities (Moilanen et al., 2008; Leathwick et al., 2010). The speedups resulting from planning units will depend on the size of the planning units.</td>
</tr>
<tr>
<td>Edge removal</td>
<td>10× faster</td>
<td>No – structural</td>
<td>Restricts removal of cells during ranking to the edge of remaining areas, thus promoting structural connectivity and speeding up computations significantly (Moilanen et al., 2005).</td>
</tr>
<tr>
<td>Boundary Length Penalty – BLP</td>
<td>Often 2× faster</td>
<td>No – structural</td>
<td>Boundary Length Penalty; penalizes edge length to area ratio resulting on more compact shapes. Speeds computations in combination with edge removal due to reduced edge length of remaining area (Moilanen and Wintle, 2007).</td>
</tr>
<tr>
<td>Distribution smoothing</td>
<td>~2× slower</td>
<td>Yes</td>
<td>Converts a habitat (suitability) map into a corresponding metapopulation-type connectivity map, assuming a radially symmetric negative-exponential dispersal kernel (Moilanen, 2005; Moilanen et al., 2005). Implemented during data input and pre-processing, and slows computations primarily due to the duplications of analysis features (habitat distributions &amp; connectivity transformed maps).</td>
</tr>
<tr>
<td>Boundary quality penalty, BQP</td>
<td>10–1000× slower</td>
<td>Yes</td>
<td>Feature-specific response to habitat loss in a specific neighbourhood around focal cell (Moilanen and Wintle, 2007), or, e.g. how sensitive species are to fragmentation. Slows down computations significantly because a change to a focal cell influences not only the cell itself but also its neighborhood.</td>
</tr>
<tr>
<td>Neighbourhood quality penalty, NQP</td>
<td>~1–10× slower</td>
<td>Yes</td>
<td>Generalization of the BQP technique to a riverine system, where features have directed connectivity both upriver and downriver (Moilanen et al., 2008). Not as slow as the BQP because water catchments are used as planning units. Relatively fast if planning units are large and slow if planning units are small.</td>
</tr>
<tr>
<td>Connectivity interaction</td>
<td>1–3× slower</td>
<td>Yes, pair-wise between 2 features</td>
<td>Connectivity method primarily intended to model connectivity when there are multiple partially similar environment types (such as forest types) that help each other’s connectivity to varying degrees (Lehtomäki et al., 2009). Typically doubles the number of features in analysis thus slowing down computations.</td>
</tr>
</tbody>
</table>
| Matrix connectivity          | 2–3× slower    | Yes, between many features | }
lower resolution aggregate cell. As an exception, if one wishes to retain information about species richness indices or other similar aggregate quantities, then selecting the maximum of such quantities within the aggregate cell may be a good option. Usually however, it is desirable to work on original species distribution information rather than on aggregate data layers.

One should also consider the treatment of missing data: a low resolution cell should only become missing data if all its component cells are missing data. Note that the use of sum instead of mean of high-resolution cell values is relevant in particular when the number of GIS software or hardware errors is relatively high, or when the proportion of cells with data within the aggregate cell varies, as commonly is the case. If there are no high-resolution cells with missing data, then taking the sum and mean produce an equivalent outcome from the perspective of Zonation.

Good data management and documentation are essential to promote transparency and repeatability in spatial conservation prioritization. Different parts of data collection and analysis may be carried out by different participants (researchers, organizations, private companies), which underlines the importance of producing and maintaining coherent metadata. At a minimum, data should be validated and the necessary metadata descriptions should be recorded at the interface between two stages (Fig. 2).

### Table 3

Typical amounts of human resources and time demanded by different stages of the spatial conservation prioritization process. Human resources are given as high (>10 people), medium (3–10 people), and low (1–2 people). Uncertainty indicates the relative degree of potential for time-consuming surprises. One main point here is that if data is not ready, data acquisition and preparation will likely dominate the time budget.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Time (approximate % of the total)</th>
<th>Human resources</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation of the ecological model</td>
<td>30%</td>
<td>&lt;15% H</td>
<td>M</td>
</tr>
<tr>
<td>Acquisition and preparation of data</td>
<td>15%</td>
<td>80% L/H</td>
<td>M/H</td>
</tr>
<tr>
<td>Analysis with Zonation</td>
<td>20%</td>
<td>10% L</td>
<td>L</td>
</tr>
<tr>
<td>Interpretation of results</td>
<td>15%</td>
<td>&lt;5% M</td>
<td>L</td>
</tr>
<tr>
<td>Communication</td>
<td>20%</td>
<td>~5% L</td>
<td>M</td>
</tr>
</tbody>
</table>

### 2.6. The concept of balance and staged development of analysis

Computational analysis with Zonation has been extensively described elsewhere (Moilanen et al., 2005, 2009b; Moilanen and Arponen, 2011; see also Section 2.3. Here we present a different interpretation concentrated around the concept of balance. Trade-offs cannot be avoided in conservation. The most basic trade-off is between focal species (or other biodiversity features): in the world of limited budgets, using more money on one species implies less money for others. Protecting a lot of one environment implies that less of another can be protected. The balance between species can be influenced by targets or priority weights set to them. However, there are several other balances that may need to be attended to, including (i) the balance between features; (ii) the balance between habitat area, habitat quality, and connectivity for each feature (Hodgson et al., 2011); (iii) the balance between currently present features and their projected occurrences in the future; (iv) the balance between conservation benefits and costs; (v) the balance between conservation and alternative land uses; (vi) the balance between different administrative regions; and (vii) the balance of costs incurred by different stakeholders. Balancing between all these different factors is not a trivial task and complicates the development of a sensible analysis setup, but the default settings and the working principles of Zonation have been designed to deal with trade-offs in an a-priori sensible manner.

As a practical matter in the development of complex prioritization analyses, we emphasize that it is helpful to develop these analyses step by step. The simplest starting point is an analysis that only includes distributions of features with everything weighted equally and no connectivity effects in use. Next, one can stepwise bring in feature weights, competing land uses or costs, simple connectivity considerations, uncertainty of feature distributions, more complicated connectivity considerations, and so on. The order might vary depending on the particulars of the analysis. There are at least two reasons why these analyses should be developed in stages. First, comparison between successive analysis stages allows one to verify that the change in outcome is sensible. In our experience, unusual changes in results may indicate errors in data preparation (modelling, GIS processing, etc.). Erroreous inputs are less easy to detect when multiple new analysis components are introduced simultaneously, obscuring the individual effect of each component. Second, a comparison between successive analysis stages can be easily informative in itself. Consider, for example, an analysis with and without connectivity. Areas that rise in priority after connectivity is brought in are areas of less than ideal local habitat quality but which are needed for the connectivity of the reserve network.

### 2.7. Interpretation of results

In this section we summarize how the Zonation output can be utilized for different analytical and practical purposes and provide published examples for each case. In the following discussion we assume that the landscape is ranked evenly so that all planning units (cells) fall in the priority interval \(0 \leq x \leq 1\). In other words, each cell has a priority value between 0 and 1. From here on we will simply refer to these priority ranks as priorities. As related information, part (V) of the Zonation user manual includes ~25 recipes and sample setups for common planning problems of varying complexity (Moilanen et al., 2012).

(i) Identification of the best areas for conservation (Kremen et al., 2008; Bekessy et al., 2012; Taberlet et al., 2012). Assume we are interested in allocating a fraction \(x\) of the landscape for conservation. Areas with priorities in the interval \(1-x, 1\) indicate the top areas for conservation. The value of \(x\) depends greatly on the data, analysis setup, and objectives, but typically varies between 0.02 and 0.2. These areas contain a balanced representation of all features (habitat type and/or species distributions) included in the analysis.

Further processing using GIS software can reveal, for example, where the highest-priorities of a particular feature are located. This can be achieved by overlaying the priority rank map with a distribution map of the feature of interest. Landscape identification analysis can be used for identification of management landscapes that are spatially connected and have similar biodiversity in separate patches (Moilanen et al., 2005, 2012).
(ii) Identification of ecologically less important areas for alternative land uses (Moinian et al., 2011a). Assume we are interested in allocating activities that can have adverse effects on the environment away from the ecologically most valuable areas (Moinian, 2012). For this purpose, it is possible to identify areas of the lowest conservation priority, identified by priorities [0, b], where b is the lowest fraction of interest. Standard Zonation outputs and the performance curves give a measure of what fraction of the features’ distributions is included in this lowest-priority fraction of the landscape. Note that an apparently low-priority area might nevertheless be ecologically important due to factors for which data was not available. In other words, quality and breadth of data are important when trying to identify areas for alternative land uses, which of course applies also to defining areas for conservation.

(iii) Planning of the expansions of reserve networks (Procotor et al., 2011). This analysis must be done using a hierarchical structure of prioritization, which is specified using a mask that identifies the locations of existing reserves. These areas are then held back and assigned ranks after all non-protected areas have been ranked (Kremen et al., 2008; Lehtomäki et al., 2009). This way it is guaranteed that the highest priorities are located in the present conservation areas. Assuming present conservation areas cover fraction c of the landscape, and that an expansion of size e is sought, the areas of interest correspond to priority ranks in interval [1-c, 1-c].

(iv) Evaluation of existing or proposed reserve areas is done using the replacement cost technique (Cabeza and Moilanen, 2006; Moinian et al., 2009a) in which two different solutions are compared. The first solution is called the ideal unconstrained (Zonation) solution, which is obtained by performing the standard analysis process. The second solution is calculated with a hierarchy enforced (as in analysis iii, above) in which the highest priorities are constrained into existing proposed conservation areas that should be evaluated (e.g. Leathwick et al., 2008). Comparison between the performance curves of the ideal and constrained solution reveals how much is lost in terms of feature occurrences due to the constraints.

(v) Target based planning. (Margules and Sarkar, 2007; Carwardine et al., 2009) which is widely implemented in other SCP software, can also be implemented in Zonation (Carvalho et al., 2010). While target-based planning is not the primary analysis mode in spatial conservation prioritization, it is possible to apply the analysis so that the targets are met with a minimally small (or minimum cost) fraction of the landscape (Moinian, 2007). From this result, a threshold x is identified so that the top fraction [1-x, 1] just barely satisfies targets for all features.

(vi) Replacement cost analysis. Item (iv), above, is a special case of replacement cost calculations. In a more general form, we are interested in the contrast between an ideal free solution and another solution that has some constraints, such as existing conservation areas (Leathwick et al., 2008; Moinian, 2012). One can ask, for example, “How much conservation value is lost due to constraints on cost or land availability?”

(vii) Targeting of incentive funding for conservation. Here the question is about how to allocate funding across different locations. The question could also be about which areas offer to accept in a reverse auction (Wilson et al., 2009). In addition to the priority rank map, the weighted range-size normalized richness score (also output by Zonation), can be utilized for this purpose. This measure specifies the weighted fraction of species distributions represented in the particular planning unit. This fraction, s, is an absolute measure of the conservation value of an area, and it can be directly used to scale incentive funding. Areas could for example be ranked by the \( s^2 / c_i \) ratio, where \( c_i \) is the price offer for the area.

(viii) Targeting of habitat maintenance of restoration (Thomson et al., 2009; Moilan et al., 2011b; Sirkia et al., 2012; Mikkonen and Moilan, 2013). This includes several topics that can be indirectly handled via relatively complicated analysis setups. One way forward is to first produce sensible scenarios about which maintenance/restoration actions would be sensible in what places. Not all actions make sense in all habitats; some protected areas would be in acceptable condition already and thus would not require any action while some areas would be unavailable for conservation, for example due to landownership. Second, the biodiversity features can be modified before entering them into the analysis to reflect how different maintenance or restoration measures would affect them if carried out; this is called retention analysis in Zonation (Moinian et al., 2011b, 2012). Prioritization can then indicate where maintenance/restoration would provide the greatest benefits, allowing for present distribution, connectivity, costs, and other such factors (Thomson et al., 2009; Moilan et al., 2011b).

(ix) Analysis in the context of climate change (Summers et al., 2012). In one analysis of this type, layers representing distributions of features both now (observed or modelled) and in the future (modelled based on climate scenarios) are entered into the analysis. The present distributions are linked to future distributions via connectivity transforms between distributions (Carroll et al., 2010). Weights can be decreased for the future and for the connectivity layers to reflect higher relative uncertainty for the future and/or connectivity (Kujala et al., 2013).

(x) Impact avoidance and offsetting. Impact avoidance is the first step in a mitigation hierarchy aiming at reducing the negative environmental impacts of economic development. Biodiversity offsetting is compensation for unavoidable damage caused by development. Offsetting can be done using Zonation; the damaged areas are “masked out” in a hierarchical analysis and compensation is sought by simultaneous optimal reserve network expansion (Moinian, 2012). This analysis should utilize the retention feature of Zonation (Moinian et al., 2011b; Moinian, 2012) to ensure that actions taken really produce compensating benefits.

2.8. Benefits, disadvantages, and threats

Stakeholder involvement can be crucial in several stages of the prioritization process. Fig. 3 summarizes benefits, disadvantages, and threats implied by the use of spatial prioritization, as seen from the perspectives of different stakeholders: conservation scientists, environmental administration, conservation NGOs, businesses and their lobbying groups, and private citizens. Which advantages or concerns are relevant would depend on the stakeholders, specific objectives of analysis, available data, and regional and/or national considerations such as environmental legislation and governance. In other words, the concerns of stakeholders would differ between real-world planning cases. Individual locally relevant factors might exist that are not included in Fig. 1. The online Appendix includes a more extensive table summarizing potential concerns of stakeholders (Table A1).

It is important to define who the decision-makers are and who the stakeholders are. Frequently, decision-makers would be working in the government administration responsible for the allocation of the public conservation resources. Stakeholders would frequently be groups who have a vested interest in the decision outcome, but no real mandate for making the decisions. Stakeholder involvement is present in several stages in the process proposed in Fig. 1. Stakeholder input is needed to build a competent and informative Zonation analysis, while at the same time stakeholders learn about the fundamentals of ecology, conservation biology, and conservation resource allocation. This is a great benefit in its own right, and also provides semi-mandatory background information for later stages in which the results are interpreted and translated back into recommendations for action. Equally important is to consider who the end-users of the analyses are and how recommendations for conservation implementation are generated. Final recommendations for conservation action fall outside the spatial conservation prioritization process described here and are part of the broader conservation decision-making context (Fig. 1). Information provided by the prioritization is conveyed to decision-makers or practitioners who then make use of it in decisions concerning resource allocation.
2.9. Resources required & risks of failure

Resourcing is a topic that is immediately encountered when doing conservation prioritization for a real-world application. Table 3 summarizes the relative resources needed in different parts of the planning process. Factors to consider include time, money, availability of computational facilities, and availability of personnel that are competent in different phases of planning. Other relevant factors concern data, model development, and quantitative analysis. Unless data is available in the correct format, experience has shown that most of the time will most likely go for the collection and formatting of data. Pre-processing of data may also need specialist skills in GIS or species distribution modelling. Development of the ecologically based model of conservation value may benefit from the participation of several experts across different stakeholders, and consequently personnel demands are comparatively high for this stage. The computational spatial prioritization analysis itself is typically relatively straightforward and fast, assuming the availability of personnel who have prior experience in the design of analysis setups and in the technical aspects of operating the software.

3. Discussion

The framework summarized here does have limitations and potential pitfalls that are typical for any sort of conservation planning. If data availability or quality turns out poorer than expected, the utility of the analysis may be compromised. If data is taxonomically limited to start with, the analysis will naturally be informative to other groups only via (commonly unreliable) surrogacy relationships. The time and effort going into data collation and formatting may easily be underestimated, possibly leading to failure to satisfy the (unrealistic) expectations of stakeholders. One limitation of Zonation is that it is based on analysis of static biodiversity patterns, and the analysis process does not involve any dynamic process-based model of biodiversity. Dynamic features can be only partially accounted for by entering data for many time steps (Carroll et al., 2010). While more complex and realistic analysis frameworks may be desirable conceptually (Langford et al., 2011), in reality application of complex methodologies is generally compromised by lack of adequate data and expertise for implementation (Stoms et al., 2011).

Despite these limitations, the framework discussed here has been successfully used in operational conservation decision making, as have other related approaches that apply target-based planning on static biodiversity pattern data, such as the Marxan software (Possingham et al., 2000). Operational use is feasible when the data base is broad enough and of sufficient quality to be reliably informative for the planning problem at hand. Even if conservation prioritization analysis is not directly used for land-use planning, it could nevertheless assist, for example, in the targeting of survey effort, in the verification of the quality of focal areas, or in the specification of additional data needs. Combining quantitative analysis with further input from knowledgeable local experts makes sense because local experts will be aware of factors that have not been available in quantitative form for analysis and in this respect, Zonation should be regarded as a decision support rather than a decision making tool. In fact, many of the strengths attributed to multi-criteria analysis (MCA) branch of decision-support tools also apply for using Zonation. For example, stakeholders (including the experts) can learn how inputs and analysis options affect the outcome of the analysis, and recording the various stages of the process provide and explicit documentation on how a particular outcome was reached (Zerger et al., 2011). Especially the latter should be mandatory for any decision-support information. A further important role for experts is the verification of the quality of results and subsequent monitoring of conservation success, both of which are major components of the operational model of systematic conservation planning (Margules and Pressey, 2000; Margules and Sarkar, 2007).

Comparison of the present framework to target-based planning, the most commonly used analysis of SCP, is a topic that cannot be avoided. While in-depth comparison is beyond the scope of this discussion, we highlight several factors requiring consideration when comparing alternative analysis frameworks. In practice, the comparison is between software packages that must be used for doing the relevant analyses. A software package always implements solutions for some classes of planning problems. Software packages will differ, for example, in (i) the analyses they are presently capable of; (ii) the extent to which they are under active development and the analyses they will potentially be capable of in the future; (iii) the size of computational tasks that can be accommodated by the software; (iv) the detail and credibility of the scientific documentation of the approach; (v) public availability of software, documentation, and support; (vi) ease of operation; and (vii) the availability of worked or published examples of their use.

There are major differences between Zonation and other commonly used SCP frameworks including Marxan (Possingham et al., 2000), Marxan with zones (Watts et al., 2009), ConsNet (Ciarleglio et al., 2009) and C-plan (Pressey et al., 2009). C-plan is an interactive planning platform and as such differs markedly from these other approaches that all apply optimization. Zonation differs from Marxan, Marxan with zones, and ConsNet in that it produces a priority ranking through the landscape instead of a target-based solution. Zonation applies deterministic computation on large grids (which links well to statistical habitat modelling), while Marxan, Marxan with zones, and ConsNet apply stochastic optimization on a polygon-based description of the landscape (different solutions from stochastic optimization are frequently used to provide flexibility into planning). Marxan with zones is intended for allocation of alternative conservation actions, while Marxan, ConsNet, and Zonation primarily operate on binary planning problems. Furthermore, options for dealing with connectivity, uncertainty, environment types and ecological communities, administrative division of the landscape, etc., differ greatly between these software packages.

When assessing whether spatial prioritization using Zonation would be useful, we believe the following points merit consideration: (i) the spatial prioritization approach described here is most useful when using many biodiversity features to determine priority areas of the landscape without having well-justified individual targets available for each feature. Also, in particular the ABF version of Zonation analyses can be expected to produce higher return on conservation investment that target-based planning. Target-based planning may lose aggregate efficiency due to the nestedness of species distributions being ignored in target setting, leading to disproportional investment in features that occur in relatively species-poor and expensive areas (Di Minin and Moilanen, 2012; Laitila and Moilanen, 2012), (ii) Zonation has a high variety of analysis features, including many connectivity methods, and (iii) Zonation is applicable on very large grids and is thus suited for large-scale high-resolution analysis. In summary, despite similarities in the broad aim—assisting with spatial conservation decision making—there are major conceptual and practical differences between approaches, and these differences may well be relevant for anyone who wishes to apply systematic conservation planning or spatial conservation prioritization on real-world problems. Further material for comparison can be found from the original scientific literature describing each computational approach, from software user manuals, and from published examples of their use in real-world settings.

As a final observation, Zonation is not generally thought of as a method for target-based planning although the capability has been available for some time (Moilanen, 2007). However, recent work shows that Zonation is able to answer a combined targets + benefit-based overrepresentation problem thus providing higher return on conservation investment than pure target-based
minimum set approaches (Laitila and Moilanen, 2012). An additional advantage of Zonation is that individual targets do not need to be defined a priori for each biodiversity feature of interest. To conclude, we expect that the present work should be helpful to conservation scientists and managers who have the need to apply the Zonation framework to assist with real-world large-scale, high-resolution conservation decision analysis. Numerical analysis techniques implemented in Zonation have been described elsewhere. Here we have concentrated on what happens before and after the numerical analysis itself, two stages of analysis that cannot be avoided in any real application.

Acknowledgements

We thank the ERC-StG grant 260393 (GEDA), Finnish Ministry of Environment and the Academy of Finland centre of excellence programme 2012–2017, grant 250444, for support. We thank Rafael Loyoia, Priscilla Lemes, and Victoria Veach and two anonymous reviewers for constructive comments that substantially improved this work.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2013.05.001.

References

Cabeza, M., Moilanen, A., 2006. Replacement cost: a practical measure of site value after the numerical analysis itself, two stages of analysis that cannot be avoided in any real application.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2013.05.001.

References

Cabeza, M., Moilanen, A., 2006. Replacement cost: a practical measure of site value after the numerical analysis itself, two stages of analysis that cannot be avoided in any real application.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2013.05.001.

References

Cabeza, M., Moilanen, A., 2006. Replacement cost: a practical measure of site value after the numerical analysis itself, two stages of analysis that cannot be avoided in any real application.


Proctor, S., McClean, C.J., Hill, J.K., 2011. Protected areas of Borneo fail to protect forest landscapes with high habitat connectivity. Biodiversity and Conservation 20, 2693–2704.


