Running a Zonation Planning Project

By C-BIG, the Conservation Biology Informatics Group

Joona Lehtomäki
Atte Moilanen
Tuuli Toivonen
John Leathwick

Running a Zonation planning project

Joona Lehtomäki

Atte Moilanen

Tuuli Toivonen

John Leathwick

Table of Contents

Preface

A spatial conservation prioritization project using Zonation involves many different stages out of which the technical phase of running Zonation is just one. The exact underpinnings and features of Zonation have been extensively described elsewhere, most notably in the technical manual (Moilanen et al., 2014) and in many publications on the topic. Furthermore, setting up a simple Zonation analysis and getting the software up and running quickly has been described in the easy-to-approach document "A quick introduction to Zonation" (Di Minin et al., 2014). Until now, however, what has been lacking is a general introduction to the planning, construction and execution of a spatial conservation planning project using the Zonation, with a particular emphasis on the steps required both before and after the actual Zonation analysis. This document aims to provide that information. All of the aforementioned documents complement each other and in case you are developing your own Zonation analyses, we recommend keeping all of them nearby.

In **Chapter 1**, we begin by revisiting things you should consider well before even thinking about the nitty-gritty details of implementing and executing Zonation analyses. We will be looking at whether Zonation is the right tool for your job and the types of resources that you will require. In **Chapter 2,** we will discuss a very important aspect of all conservation planning and prioritization projects: setting the objectives. Without being explicit about your objectives there is a high chance that the results will not be as informative as you would like. In **Chapter 3**, we will be looking at how to translate a set of often relatively abstract and broad objectives into a conservation prioritization project using Zonation. In other words, we will be discussing how to use your data, expertise and tools to best meet your objectives. This will include consideration of spatial

data required for a prioritisation analysis, including descriptions of biodiversity features and other relevant factors, such as habitat condition and costs. In **Chapter 4**, we will explore key aspects of working with real data, the stage in which you will most probably spend the most time. In **Chapter 5**, we continue directly from Chapter 4 and look at how Zonation analyses are typically developed, starting with a simple model but progressively making it both more realistic and complex through the consideration of other factors. Here we will also share our experience regarding how to organize your projects. We will not be spending any time on the details of Zonation itself, as these are already well covered by the manual and the "Quick introduction to Zonation"-document. Instead, in **Chapter 6**, we will skip over the details of running Zonation and continue directly to the topic of visualizing, interpreting and validating the results. Finally, in **Chapter 7**, we will use several real-life planning examples to show how Zonation can be used as part of operational spatial planning both for conservation and for more general land use planning.

Throughout this document, we assume that readers will have general familiarity with both spatial conservation prioritization and Zonation, and provide pointers to background literature only where it is directly helpful. It is also worth pointing out that many of the issues that we discuss, related to both project planning and technical implementation, are not specific to Zonation, but could also be useful when implementing such analyses with other conservation planning tools. However, each real-life conservation prioritization project will invariably have its own quirks and peculiarities, so it is impossible to give detailed "how-to" instructions that would fit all occasions. Consider this document therefore a general introduction, which will give you a good starting point for using Zonation in your own projects.

Acknowledgements

Summary

Zonation is decision support software for spatial land use planning. It was originally developed for solving various problems around spatial conservation resource allocation, and it is capable of data rich, large scale, high-resolution spatial analysis. Today, applications of Zonation are not limited to conservation; it can also address variable problems in ecologically aware land use planning, including such as spatial impact avoidance and targeting of biodiversity offsets. The Zonation software implements a set of interrelated techniques in one package. It can in analysis account for many factors, such as multiple costs, habitat condition and ecological connectivity. While Zonation can flexibly address various spatial planning problems, it does not do statistical species distribution modelling (SDM) or spatial population viability analysis (SPVA). Nevertheless, outputs from SDM or SPVA can be used as inputs for Zonation.

If you have need for a spatial prioritization analysis, the first thing you need to do is to assess whether Zonation is the right tool for the job. Problems frequently addressed with Zonation include the identification of ecologically important areas for reserve network expansion or identification of the ecologically least important areas for impact avoidance of development projects. Zonation can be used for both scientific research and real-life planning. Explicitly defining the domain of use will help you with setting up the project.

Using Zonation in a planning project requires certain human and other resources. While details will heavily depend on the nature of you project, you will probably need a team of people with different skills. In addition to knowing how to use Zonation, experts in related scientific fields, project coordination, and spatial data management are often needed. In case your project aims at delivering real-life decision-support, engaging with relevant decision-makers and stakeholders is crucial. Very often, knowledge production among a group of people with varying backgrounds, objectives and expertise is one of the main outcomes of a planning project. Having sufficient technical (i.e. computer and storage) capacity is also required, as is having the actual spatial data for the Zonation analyses. Overall, the time required for running a Zonation planning project varies from few weeks to several years.

Once you have established that Zonation might be a suitable tool for you conservation planning project, you need to establish clear and explicit objectives for the project. This means first defining the high-level objectives such as "expand the reserve network by 5%" followed by translating these high-level objectives into components of the actual Zonation analysis. These components – which biodiversity features are used, how they are weighted, how connectivity is dealt with, are costs included, etc. – jointly constitute a model of spatial prioritization. This model in principle also defines what kind of data are needed, but in practice, the availability of data also restricts what kind of model can be constructed.

Essentially a Zonation analysis is an implementation of a particular model of spatial prioritization. Before doing the actual analysis, relevant data need to be acquired, verified, and pre-processed into the same spatial extent and resolution. This stage is frequently the most time-consuming part of a planning project, especially when inputs include many biodiversity features with high-resolution distribution information. Automating the pre-processing steps when possible and paying attention to data management planning can save you time and trouble.

Setting up and running Zonation requires planning as well. Instead of including all

the components of the model of spatial prioritization into a single Zonation analysis at one go, it is useful to develop a sequence of increasingly complex development variants. This way you can get a better view on how a particular analysis option, like the inclusion of ecological connectivity, influences the prioritization. While Zonation does not enforce a particular way of organizing your input and output files, we recommend following certain best practices such as placing individual development variants into subdirectories under your project folders etc.

Visualization and interpretation should directly serve the intended use of the results and these needs vary depending on whether the results are intended for scientific publication or real-life implementers. In addition to producing the standard output, Zonation is capable of performing useful post-processing operations on the results. Alternatively, the prioritization results can be post-processed with external software tools such as the R programming language. Using the different post-processing techniques, you can create whatever planning products best support the objectives of your project. Your results should also be subject to continuous verification between the different technical stages of the project, as well as final validation to check that your analyses actually address the high-level objectives you established in the beginning.

An overall successful Zonation planning project often requires running several individual planning stages in an iterative fashion. Running the Zonation software is just one, albeit an important, stage. Ultimately, having the capacity to run a successful Zontation planning project in any organization requires that all of the project stages can be addressed. In return, it is possible to achieve not only useful spatial conservation plans supported by data, but also increased knowledge sharing and understanding between all the stakeholders participating in the planning project. Frequently, this learning is as valuable as the prioritization outcome itself.

1. BEFORE YOU START

In this chapter, we discuss issues that will help you to decide whether Zonation might be a suitable tool for your analysis, focussing attention on the context within which you are planning to use Zonation, and the types of resources (data, human resources etc.) that you might need.

1.1 IS ZONATION THE RIGHT TOOL FOR THE JOB?

Zonation is a flexible tool that can be used for tackling a broad range of spatial conservation prioritization problems (see Box 1 and e.g. Di Minin et al. (2014) and references therein). However, the number of quantitative tools available to address such tasks has increased greatly during the last decade, leaving potential Zonation users faced with an important question: is Zonation the right tool for the job? We can partially answer the question by considering what Zonation *is not*. Zonation is often confused with other spatial approaches and software, including species distribution models (SDM), spatial population viability analyses (SPVA, and geographic information systems (GIS) — it does none of these tasks. While Zonation has several ways of including connectivity in the prioritization analyses (see for example 3.1.3), it is not a tool for analysing landscape-level connectivity either. So what is Zonation? It is a tool for balancing the requirements of many biodiversity features simultaneously in ecologically informed land use planning. A potential user might wish to compare Zonation with alternative and well-established software for spatial conservation planning and prioritization. These have been considered in some depth in Chapter 5 of "A quick introduction to Zonation" (Di Minin et al., 2014).

While Zonation is primarily a tool for spatial conservation prioritization, it is suited also for general land use planning and resource allocation. By extending the conceptual framework described in this document to include features describing land-uses other than conservation, Zonation can be a useful decision-support tool for zoning and natural resource planning. Ideally, of course, we want to take all of these aspects into account simultaneously. As with conservation planning, the use of Zonation for general land-use planning will require both the explicit formulation of objectives and the availability of suitable data required in analyses.

When Zonation is used in a real-life planning context, we would like to emphasize that it really is a decision-support tool, rather than a decision-making tool. Zonation analyses are only useful insofar as they can provide credible information that is relevant for the

---

Box 1. Typical spatial (conservation) prioritization problems addressed with Zonation.

1. Identification of a well-balanced set of ecologically important areas for reserve network expansion

2. Identification of ecologically least important areas for impact avoidance of development projects

3. Evaluation of existing or proposed conservation area networks

4. Spatial allocation of habitat maintenance or restoration

5. Targeting of financial incentives for conservation

6. Planning for climate change mitigation and adaptation

7. Planning of biodiversity offsets when impact avoidance does not suffice

planning problem at hand. As with other decision-support tools, Zonation will usually not provide a single correct answer, but rather a suite of alternative - and often closely related - answers that should always be interpreted in the context of the goals and objectives defined (see also Chapter 7). In our experience, the process of engagement between the analyst and various experts and stakeholders can be one of the most rewarding parts of a Zonation project.

Given the often highly context-specific nature of spatial conservation planning, it is difficult to give exact selection criteria for picking the right spatial prioritisation tool. Nevertheless, we have collated some guiding principles based on our own experience (Figure 1). We recommend that anyone considering using Zonation gives careful consideration to the background and details of each approach

to assess which tool might do the best job. Consideration should also be given to the practicality of interfacing between the selected prioritisation tool and other analytical tools that are required to complete the prioritisation project. For example, SDMs are often used to generate distribution estimates for the biodiversity features used in Zonation analysis, and most of the data preparation and visualization of results will be done using GIS software.

1.2 SETTING AN APPROPRIATE DOMAIN AND CONTEXT

Zonation is normally used within one of the following (partly overlapping) domains: curiosity-driven (research) science, issue-driven (regulatory) science (*sensu* van den Hove, 2007), and operational planning for conservation management. The distinction
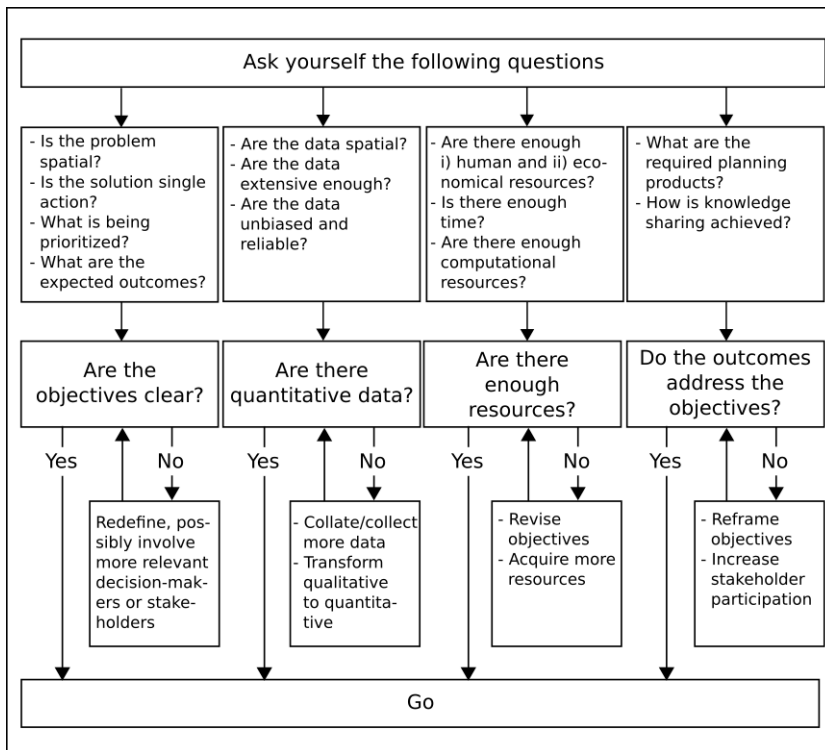
Figure 1. Relevant questions to ask when considering if Zonation (or any other spatial prioritization tool for that matter) is likely to be a useful tool for the job. The questions are best asked in parallel; objectives, available data and the resources needed are often closely intertwined. Note that requirements in scientific studies usually are lower than in policy-relevant on-the-ground applications.

Table 1. The three domains in which Zonation may be useful. Note that each real use-case will probably combine aspects from multiple columns. Table borrows from Jasanoff (1990).

	Curiosity-driven science	Issue-driven science	Operational planning
Objective	Scientific insight, novelty, and significance	Knowledge relevant for forming and assessing policies	Application of existing knowledge
Products	Published scientific papers	Reports and white papers, often unpublished	Operational plans including maps, priority lists etc.
Important knowledge production components	Credibility	Relevance, legitimacy	Relevance, legitimacy
Decision-making context	Does not necessarily have one	An existing context, can also aim at establishing a new process	Embedded in an existing decision-making context
Accountability	To scientific community and professional peers	To political decision-makers, general public	To political decision-makers, civil servants, stakeholders

between these domains is not always clear, as large operational analyses will often require the development of novel methods, and many researchers actively seek to maximise the operational relevance of their research. Having a clear understanding of the domain within which your project falls is helpful, given the consequent differences in emphasis (Table 1). For example, a Zonation project commissioned by an environmental authority might place more on "getting it right", because unsubstantiated results might compromise not only the utility of the results but also the credibility of the authority in the eyes of different stakeholders. From a technical point of view, this might be reflected in an increased emphasis on ensuring the quality of input data. We do not mean to imply that "getting it right" would not be important for curiosity-driven science and published papers, but simply wish to draw attention to the way that these emphases can differ depending on the analysis context. Demonstrating a scientific

principle does not require the illustrative example to be policy-relevant and based on high-quality data.

Furthermore, there is an associated difference between executing a spatial prioritization analysis project using Zonation and developing a broader spatial conservation process in which Zonation will be the tool used to identify priorities.

An analysis *project*:

- Has an objective or outcome to be accomplished, and the project ends when that objective is accomplished. These are often scientific projects, but can also be practical projects such as, "Expand the regional conservation area network by 100 hectares of forested areas".

- The objective might evolve during the project.

- Has a beginning and an end, although the beginning and end may not be well defined when the project starts and the end might be a long time in the future.

- The sequence of tasks is not normally repetitive (except for what is required to identify and correct problems in analysis) and may not be fully known at the outset of the project.

- Is frequently based on an existing set of data, which is used to best advantage.

- Usually involves relatively few stakeholders.

A broader *process* on the other hand:

- Has an objective that is defined around the ongoing operations of an organization. For example, "develop the conservation area network of Finland".

- Is a repetitive iterative sequence of tasks, which are typically known from the outset.

- Does not normally have an end defined at the time when Zonation is first used.

- Frequently involves multiple stakeholders and includes acquisition or improvement of data.

Although this difference might appear subtle, our collective experience indicates the importance of getting this right. In particular, the management of a prioritisation process as if it is simply an analysis project raises a serious risk that the overall goals will not be met. That is, the institutional and social framework required to successfully implement a prioritisation process require just as careful planning as the analysis component itself. For clarity and simplicity, we will here generally talk about running a Zonation analysis project except where the point is to specifically address issues related to using Zonation within a broader decision-making process.

1.3 BUILDING A TEAM

The number and people required for a project, and their skills will depend heavily on its exact nature. Take a moment and consider again the specifics of your project (see previous Section 1.2); this will help you to decide for example whether you need to involve decision-makers or stakeholders from the outset. Even if this is not required, in our experience, most projects benefit from having access to a diverse range of skills, reflecting the integrative nature of most prioritisation analyses. Therefore, you will typically want to assemble a team in which people with different skillsets will complement each other in different roles. In the following, we list the most common roles that need to be filled in a Zonation project. Note that the roles given here are approximate and often a single person can act in several roles.

A coordinator is someone who coordinates the planning and execution of the project and interacts with stakeholders and decision-makers. Since expert-elicited information often plays a big role in Zonation projects, good facilitation and communication skills are essential for a coordinator. A coordinator is especially important when constructing a process that supports operational conservation planning, and where achieving clarity and institutional agreement around project objectives is required.

A science expert is a person who has a background in one or several of the relevant scientific fields such as ecology, conservation, decision analysis, geography, etc. This person's main responsibility is to lead the construction of the model of spatial prioritization (see Chapter 3) in close collaboration with the rest of the team. While expertise with Zonation is not the key-asset of a person in this role, the science expert should have sufficient understanding of Zonation and its requirements to be able to work effectively with other team members.

A Zonation expert is person who has experience in operating Zonation and who has a good understanding of its conceptual underpinnings.

The main responsibility of this person is to lead the development of an analytical model within Zonation that addresses the high-level objectives of the project (Chapter 2). Furthermore, skills in data management and processing, usually with a GIS, will come handy.

Unless the Zonation expert is also an expert in data processing using GIS, then one or more *GIS-experts* will also be needed. Significant effort is frequently needed in the processing of spatial data into the format and content required by analysis.

Strictly speaking, having the roles defined separately as above is not always a prerequisite for running a successful Zonation project - sometimes a single person will have to manage all four - but the separation of responsibilities will help to distribute the workload and make the execution more efficient. Typically finding someone with enough experience with Zonation is the hardest thing; starting from scratch can be challenging, but is achievable given the amount of published knowledge in the scientific literature, help documentation such as what you are reading currently, and through collaboration and support from other experienced Zonation users.

1.4 BUDGETING FOR TIME AND RESOURCES

It should be clear from the outset that major spatial planning projects using Zonation cannot be implemented without significant expense, the highest of which are usually the salaries of people for the time they spend on the project. While minor scientific projects can potentially be done quickly in some weeks, substantial real-world applications are likely to take many months or years of cumulative time. Needs are highly case-specific and you must do your own estimation of the need and availability of resources, using the following as a guide.

Human resources: Having access to skilled people is the key to successfully running

Zonation projects. Depending on the type of the project, several people with different skill sets may be needed. Someone is needed who will set up Zonation analysis. Simple analyses can be copied from the Zonation tutorial material, but more experience is required when venturing into the use of more advanced features. Participation of someone with experience with Zonation will not hurt. Another significant task is the preparation of the input data layers, including both formatting and validation. Here, experience with GIS, scripting and data manipulation is particularly use, along with sufficient ecological knowledge to be able to spot errors when they occur. Depending on the size and scope of the project, coordination and communication may also be a crucial role. Substantial effort is often required to reach agreement on project goals, to negotiate data agreements, organize project meetings, create web site updates, and to communicate results to stakeholders and the public. Some people could possibly cover more than one of these skill areas needed.

Data: It is important to budget enough time for preparatory data work, as identifying, collecting and pre-processing the required layers of spatial data is usually the most time consuming step in a Zonation planning project. Data needs are generally directly dependant on the aims of the analyses. If the aim of the analysis is to identify key areas for the provision of ecosystem services, you will need spatially explicit data describing the distributions of different ecosystem services. If you wish to also take into account occurrences of rare species or land costs, you will also need to have data describing these. Having defined your data requirements it is well worth investing time to discover if someone has already produced such data, if necessary arranging for its purchase and/or permissions for its use. Having obtained all the necessary data, consideration needs to be given to the most appropriate spatial extent and resolution of the data layers to be used in the actual prioritisation analyses. The final step is processing of data into a standard raster

format used in by Zonation. This step is usually relatively systematic and quick. Usually the data selection, processing and evaluation are an iterative process that is linked to all parts of the Zonation process.

Hardware: do you need to buy a big computer? In most cases not (see section 5.3 for more details). In our experience, many data sets can be analysed using readily available, off-the-shelf laptop computers. By contrast, some of the bigger analyses we have seen did require desktop computers with some terabytes of hard drive space and RAM in the order of 300GB. If you need to buy one yourself, these cost in the range of 10-20k Euros (or Dollars), but many institutions have high-performance computers available to employees. While these costs might seem high, any "savings" via use of an inadequate computer could potentially be lost many times over as salaries paid when inefficiencies and delays in data processing cause delays in the project. Try to estimate what your likely requirements are, and then arrange access to a sufficiently capable machine with a little room to spare, i.e., allowing for additional data and other such contingencies. Additionally to hardware, you may need some GIS software for processing the input layers - just use your favourite brand of commercial or free GIS software.

The total cost? Human and hardware resources generally comprise most of the costs when running a Zonation project. Of these, salary costs are likely to be the highest. Time needed varies from a couple of weeks to many person years of effort. We have seen projects completed within a month, but others have continued on and off for a decade. Also, note that data processing and analysis usually requires several iterations as errors are discovered and corrected and analyses rerun. This should be allowed for in the project plan and plan for resources. In addition, although most costs are generated typically by the process of data acquisition and processing, the accumulated time required to run a large stakeholder group can also accumulate quickly. For further discussion see the open-access publication by Lehtomäki and Moilanen (2013), which discusses resources needed in preparation of the model of spatial prioritization, acquisition and preparation of data, analysis with Zonation, interpretation of results, and communication.

It makes a difference whether your project is a once-off study or part of a long-term contribution to operational environmental management. If the analyses are expected to be replicated regularly, extra effort and attention will likely be needed for documentation, safe data storage, knowledge transfer, data updates and possibly in on-the-ground validation and monitoring to verify the effectiveness of conservation management. Such needs will again be case specific.

2 SETTING OBJECTIVES

2.1 DEFINING HIGH-LEVEL OBJECTIVES

In the present context, high-level objectives are about what we want from or for biodiversity and ecosystem services in general. These could be complicated general requirement like "develop a well-balanced reserve network for the whole country", or "facilitate movement of species in response to expected climate change". Alternatively, they could be relatively simple such as "find two new bird protection areas of a size larger than 5000 hectares each". For operational processes, high-level objectives are set by someone who has legitimate power for making decisions in the society, such as elected authorities, usually guided by processes involving relevant stakeholders. In most cases, someone other than those running the spatial planning analyses thus sets this means high-level objectives.

High-level objectives are usually not directly useful for operational use and must be translated into an operational (Zonation) project structure and workflow (see Chapter 3). In other words, one must design an analysis that answers the high-level objectives in question. Designing such an analysis involves

much work: you must acquire sufficient data, involve the relevant stakeholders, understand and possibly include societal constraints in the analysis and so on. In general, high-level objectives will have major implications both for setting data requirements and for the defining the extent to which involvement will be required from stakeholder groups. While a scientifically motivated minor project can be implemented amongst a small group of scientists, large projects that lead to e.g. land use restrictions will generally require mandatory involvement with stakeholders (landowners, zoning authorities, environmental administration, environmental NGOs, etc.) from the outset of project. A number of well-described examples of projects implementing engagement with stakeholders can be found in the systematic conservation planning literature.

Goals should determine the data you need. For example, if the goal is to develop a national protected area network, then data required will need to cover biodiversity as broadly as possible. This in turn implies that significant effort will need to be expended in data preparation. Most likely, it will turn out that only partial data is available within the time available, and the question will frequently arise: are my data adequate? Fortunately the answer to this question is only rarely "no": outputs of Zonation analyses can be usually be utilized to a limited extent even if the available input data are incomplete. The provision here is that the answer is only a partial answer, applying only to those biodiversity components that are covered by data either directly or indirectly via surrogacy relationships. This reflects a general reality that while prioritization analyses, including those from Zonation, are frequently used to support land use decision-making, only rarely are these based on data that are complete and unbiased. This because the alternative course, of postponing decision making until all relevant data are available, is equally unpalatable, leading inevitably to the making of *ad-hoc* decisions in the absence of more robust, evidence-based decision making frameworks.

We emphasize that co-learning is one major outcome of implementing a spatial conservation planning project with a stakeholder group. Usually everybody involved learns something useful. Scientists may – and generally probably should – learn about stakeholder values and the availability and quality of data. Stakeholders might learn about the fundamental principles of biodiversity conservation and ecologically informed land use planning. Both may gain new understanding and appreciation about the state of the environment, about pressing conservation needs and the costs of actions, etc. In most cases, high-level objectives specified at the outset of the planning project are progressively clarified and made more explicit through a consultative process. In addition, perspectives and objectives may need to be updated as the project progresses, which may in turn lead to new analytical and data needs. The nature of the process is unavoidably somewhat iterative.

2.2 PLANNING AHEAD FOR PRODUCTS

Whatever the high-level objectives of the project are, it is well worth identifying the intended end-users of the results at the outset, and clarifying planning products that they require (e.g. Pierce et al., 2005). Here, it may be worth revisiting Section 1.2, and thinking carefully about the nature of the project in which you are planning to use Zonation? For scientific (curiosity-driven) research, findings published in an article are the usual products of the project. A research project leading to a publication is rarely revisited again, although later studies may build upon techniques developed there or other findings of the project. While it is important to document how the results were produced, there is rarely a need to re-produce them. For a project supporting operational conservation planning, the situation is usually different. While the same input data may be used for a research project, the final planning and prioritisation products will typically require further processing, visualization and documentation to be maximise their accessibility to a range

of end-users including management agencies, business stakeholders, and even the general populace. Furthermore, since operational planning often is a continuous and iterative process, there is frequently a need to revise and improve the analyses and update the planning products as new information comes to hand. Consequently, repeatability and transparency of the process increase in relevance.

Zonation automatically produces a variety of different outputs, including rank priority maps describing the spatial distribution of the conservation priorities, and performance curves (Section 6.2). While these outputs often address the high-level objective directly, more often than not they require additional processing to aid in their interpretation. For example, while a rank priority map can be used to locate and list the top-priority sites, in an operational setting it can also be important to know why a particular site has been given a high (or alternatively low) priority. The answer to the question can be complex, reflecting not only the biodiversity features occurring at a site, but also the manner in which they have been ranked by the Zonation meta-algorithm (including various optional components such as connectivity, habitat condition, weights etc.). It is therefore useful to consider (and budget for) what additional planning products will be required to support achievement of the general project objectives of the project. You should keep these considerations in mind throughout the whole prioritization project and where relevant, you should consolidate all details into planning products.

2.3 MEASURING SUCCESS

Once the high-level objectives have been set, it is also important to define how the success or failure of the planning is measured. Planning can be considered a failure if it does not address the defined objectives (i.e. validation [see Section 6.3] fails). After explicitly defining the objectives and assessing the different types of information needed in prioritization, it may become obvious that the analysis cannot be implemented with any confidence of success,

e.g. because of a lack of suitable data. Even if this is the case, however, undertaking the planning project might not be a complete failure because of the co-learning aspect described in Section 2.1.

How do you then assess whether you are getting it right? The prioritization analysis itself and the results it produces should be checked and validated at several stages of the project. While you are building the analysis, it pays off to do regular checks on the data and the intermediate results to catch any potential technical problems (i.e. doing continuous verification [see Section 6.3]). More generally, it is a good idea to involve experts in checking the results early on as they usually can quickly pinpoint problems and peculiarities in the results. Finally, on-the-ground validation can be done by comparing the results to independent validation data sets to assess whether the high-priority sites indeed contain valuable conservation features. However, exactly how the validation should be done is not always obvious, especially for the more complicated analysis types. For example, if a Zonation analysis includes multiple types of connectivity and e.g. projected land-use changes or even climate change, it may be difficult to validate the results based on a snapshot of the current conditions.

Finally, we note that there are several different levels at which the success of a planning project can be measured. An informative prioritization analysis is only one step in the overall iterative process that constitutes conservation planning. A technically successful conservation prioritization assessment, produced e.g. with Zonation, does not automatically lead to improved implementation and more effective conservation.

3 BUILDING A MODEL OF SPATIAL PRIORITIZATION

3.1 IDENTIFYING THE FACTORS RELEVANT FOR ADDRESSING THE HIGH-LEVEL OBJECTIVES

Defining high-level objectives, such as "developing a well-balanced reserve network for the country" is necessary for operative spatial conservation prioritization. However, such high-level objectives are not enough in themselves to develop an informative prioritization using a tool like Zonation. As well as specifying the desired high-level objectives, we must also identify the various work components required to achieve those objectives. Whereas the high-level objectives are typically defined in the socio-political domain, these lower level components are best identified and addressed in the technical-scientific domain (Ferrier and Wintle, 2009). We encourage the reader to take a look at Ferrier & Wintle (2009) for an excellent introduction on how the need for spatial conservation prioritization is defined and how that need can be serviced.

Let us now take a closer look at our example of developing a well-balanced reserve network for a given country. Our first step is to develop an explicit and detailed statement of what is required to meet this objective. In our desired well-balanced solution, which biodiversity features (species, habitats and/or ecosystem services) do we want to include? Should some biodiversity features be given more emphasis than others? If so, on what basis do we select data, set preferences, or define weights? Should we account for ecological connectivity with existing reserves? Is it necessary to account for the economic costs of establishing reserves, and should these include just the cost of land acquisition, or also include the opportunity cost of forgoing other land uses? Identifying all the relevant factors that will contribute to the final prioritization is by no means a simple task, particularly once requirements for relevant spatial data are also included. This might include not only data layers describing the distributions of biodiversity features (species or ecosystems), but also describing social values and preferences; hence we refer to the process as building model of spatial prioritization. Building such a model usually requires close collaboration with various experts and stakeholders, necessitating face-to-face contact through meetings and workshops, as well as the more mundane, technical tasks required for data processing. In our experience, this stage of the conservation prioritization process almost invariably takes up the most time and resources.

As said, data requirements are largely set by the project goals (i.e. the high-level objectives). In practice, most spatial conservation prioritization projects rely on data already available - very few have the option to collect new data. It is therefore inevitable that the model of spatial prioritization used in the analysis is constrained by the data available. Note that we do not just mean spatially explicit input data required by Zonation, but also the supporting data used to guide various aspects of the analysis. For example, connectivity methods in Zonation should ideally be parameterized using empirical data concerning the biodiversity features at hand, but since these data are often missing, we must most often rely on expert-derived estimates instead.

Implementing a Zonation analysis does not require, however, all of the data to be available at the project outset; you can also acquire and process data in stages. In this case, it is useful to define 1) analyses using only part of the data, and 2) an analysis using the complete and final dataset. In particular, it can be useful to build a pilot analysis using part of the data, or perhaps a spatial subset of the complete study area, and building a full set of Zonation setups based on these limited data (see Chapter 5). You can use this initial analysis to develop analysis structure and to gain insight into how do the different types of biodiversity features and analysis components influence priorities.

3.1.1 Biodiversity features (input data)

In most analyses, the biodiversity features (i.e. the selected input data) will have by far the largest effect on the priority ranking. This may seem like an obvious statement, but it is nevertheless worth pointing out, as is this: your ability to understand your results will largely be determined by your knowledge and understanding of your input data. You will have to engage, implicitly or explicitly, in a selection process when deciding what data to include. The data and weights you choose should address your specific objectives.

When the high-level objectives are not clear or when it is not easy to translate these objectives into specific goals, it may be tempting to include all available features in the prioritization. While this strategy may seem attractive, you need to bear in mind that by definition you cannot assign a high priority to everything (this is the same as having all equal!). In particular, the inclusion of a very large number of features, some of which are of low importance, may reduce the utility of your results by diluting the influence of the important features on your final rankings (for example, if your focus is on endemic species, why include information about non-endemics into analysis?).

The data used to describe biodiversity features in Zonation will generally be available as: 1) continuous occurrence data, 2) categorical data or 3) point observations.

1. **Continuous occurrence data.** This type of data is most commonly used with Zonation, noting that Zonation by default uses rasterized input data containing explicit values for each raster cell. A cell value indicates the local occurrence level of the feature in the cell. Often, these are modelled probabilities of occurrence (presence/ absence) or abundance (e.g., numbers of individuals predicted to occur in a standardised sample such as a vegetation plot or research trawl). They can also be

presence-absence (0/1) data if nothing more refined is available. For habitat data, the value in each cell could indicate for example an index of habitat quality.

A common question with this type of data is whether all the biodiversity feature layers need to have the same units and if they do not, should the values be rescaled to the same scale. The answer to both of these questions is "no", the data do not need to have the same units, nor do they require rescaling. This is because Zonation automatically applies a feature-specific normalization to all feature layers prior to commencing the ranking process.

2. **Categorical data.** Another common type of data used as input to Zonation is data in discrete classes, for example, data describing different habitats or land uses. In these layers, integer values have been used to label discrete classes such as different habitat types. To enter such data into Zonation, each class needs to first be separated into an individual feature layer. Then, there are two options. First, one can enter habitat layers so that the number in the grid cell represents habitat quality in the cell. The second option is that each habitat can be entered as a presence-absence (1/0) layer into Zonation. In this case, a separate condition layer is usually developed and linked to probably several presence-absence layers to model the fact that habitats have been influenced by human impacts and are in variable condition. The benefit of the second approach is that condition layers are easily modified to account for expected changes in land use: modifying a single layer is much easier than developing an entire new data set for each scenario.

3. **Point observation data.** While Zonation can work with point observations ("species of special interest" or SSI features in Zonation jargon), use of such data is generally not recommended. When point observations are included

in analysis, one frequently gets priorities concentrated at locations where point-observation data has been available. In most cases, you will probably fare better by using species distribution modelling techniques to create continuous estimate surfaces of occurrence. The one exception where SSI features can perhaps be useful is when they are used in conjunction with planning units (see 3.3.3.11 in Zonation manual). In this case the distributions of species within planning units may not be known with certainty, but SSI points can be used to indicate which species are associated with which planning unit; because all cells within each planning unit are ranked simultaneously, the problems caused by point observations being tied to individual pixels are not a concern. Note that most connectivity features of Zonation cannot be applied to point distribution features. In summary, prefer converting point observation sets into distribution grids.

Note that after reading in a layer, Zonation internally normalizes the numbers to sum to one. In other words, Zonation works on the fraction of the distribution of a feature in a cell. This has some implications for data preparation and weighting. For example, rarity alone is not necessarily a reason to give a feature a high weight: Zonation already fully accounts for this via the range size normalization. The range-size normalization also has implications for data processing. Care is in particular required when entering categorical data, such as habitat or land cover data. In particular, you need to avoid the creation of layers describing individual habitat classes that receive high effective weights because of their rarity, but which are not particularly valuable or special from any conservation perspective. If needed, the treatment of rare and common species can be influenced by the weights given to features. Using a weight equal to the original distribution size of the feature effectively cancels the range-size normalization automated by Zonation.

Unless individual biodiversity features are extremely rare in comparison to other biodiversity features or if there are only few biodiversity features, Zonation analyses are generally not very sensitive to the inclusion or exclusion of individual features, or for the change of the weight of an individual feature for that matter. For example, if you have 100 features in analysis, adding or dropping one feature is a 1% change in the mass of data and consequently you might expect the priority ranking only to change in the order of 1%. When the amount of layers in analysis increases, you will discover a trend that the priority rankings start converging: this is a result of many features pointing to the same places - the landscape is after all strongly structured by human impacts.

3.1.2 Preferences (weights)

Weighting of features always needs to be addressed in Zonation. Weights given to features influence their balance through the priority ranking. By default, all features have a weight of 1.0, but this does not mean an absence of weighting; rather, it means that the complete loss of the distribution of any individual species is seen as equal to the loss of any other species. The first form of weighting, while not necessarily recognized as so, occurs when features are initially selected. If the distribution of a feature is not entered into analysis in the first place, it is effectively given a weight of zero. Overall, it is desirable to obtain data for a good sample of features that jointly represent biodiversity broadly. All of these features should be entered into analysis with positive weights.

Typically, two types of scenarios are encountered in the weighting of species. In the simplest case all features are of the same broad type, say mammalian species. One could have the weight vary between species, for example by giving endemic species a weight of 5.0 and the rest a weight of 1.0. Factors that can be included in the weighting scheme include threat status, endemism, taxonomic distinctiveness, economical value,

and national priorities. A special factor is data quality: weights can be decreased for those features that have relatively poor-quality data. As a technicality, weights in Zonation are relative: weighting features as 1 and 5 is the same as weighting them 0.2 and 1.0. The weighting is always subjective and there are no fixed rules. See Lehtomäki and Moilanen (2013) for further discussion.

The second scenario that requires consideration is when biodiversity features are spread across more than one broad type. In our experience, it can be useful in this case to adopt a top-down or hierarchical approach to weighting. Let's assume that one has data for both the present and predicted future distributions of some taxa. We decide that (an arbitrarily chosen) 1000 weight units is divided between them so that the feature layers describing present distributions receive a combined total weight of 700 units of weighting, but the future distribution layers receive a combined total weight of 300, reflecting their lower certainty . Then, assume for the sake of illustration that there are three major types of features in analysis: species, habitats and ecosystem services. Let's say that species and habitats are seen as equally important and both are considered to be twice as important as ESS. This implies that species and habitats each receive an aggregate weight of $2 / 5 * 700 = 280$ and ESS half of that, $1 / 5 * 700 = 140$. (These sum to 700; one would in the same manner divide the 300 given to the data representing predicted future distributions.). Then, let's assume there are 1000 species, 200 of which are endemic and receive a relative weight of 5.0. Thus, the species have $800 * 1.0 + 200 * 5.0 = 1800$ relative weight units. The aggregate weight of 280 is then distributed between species in these ratios, i.e., each endemic gets a weight of $280 * 5 / 1800 = 0.7778$, and the remaining species receive a weight of $280 * 1 / 1800 = 0.15556$. You can verify the species weights sum to an aggregate of 280 units.

Turning to the habitat types, these will typically fewer in number than the species, implying that in this case each habitat would get a much larger weight than a single species. This is fine as distribution data is usually relatively good for habitat types, habitats frequently have wide distributions which reduces their influence on the priority ranking, and each habitat is probably acting as a surrogate for many species for which there is no distribution data. The division of weights would then proceed in the same manner for ESS, i.e., with greater weights given to layers describing current distributions than to those describing predicted future patterns.

How sensitive is the priority ranking to a change in the weight of a feature? For most analyses, changing the weight of an individual feature will result in little if any change, although representation of that feature might alter. This can be understood easily enough. Say, for example, that you include 500 features in the analysis. How much should the result change if the distribution of one feature is dropped out from analysis? Not much, because one feature is only 0.2% of the mass of features. While some localized changes in the priority ranking might occur, the overall patterns will generally show very little if any change. Remember that omitting a species equals setting its weight to zero. Consequently, making a moderate change to the weight of a feature will generally have minimal impact on the ranking, particularly when using large numbers of individual species layers. In contrast, if the analysis uses a single highly weighted cost layer (section 3.1.4) then changes in this layer will have relatively large effects. The same applies to a condition layer that has been linked to many habitats (next Section 3.1.3).

As a final remark, the inclusion of zero weighted layers can be very informative about analysis outcomes. Remember that while zero-weighted features do not influence the prioritization, Zonation will still track their performance through the ranking process. For example, by including a set of zero-weighted features in their own output group, you can assess how well positively weighted features in the analyses are able to act as surrogates for the zero-weighted species.

3.1.3 Ecological considerations: connectivity, habitat condition etc.

Simplifying a bit, the amount and quality of habitat available determines the carrying capacity of a landscape for a species. The distribution (aggregation, connectivity) of the habitat may influence how much of the habitat is actually occupied: habitat fragmentation may decrease the persistence of populations at the regional level. Therefore, habitat quantity, local quality and connectivity are fundamental variables of spatial ecology, all of which can and should be considered in a Zonation analysis.

The feature-specific distribution grids control the **amount of habitat**. It is assumed that an individual feature occur in any grid cell containing value greater than zero. Any areas that are clearly unsuitable for a species (e.g., the sea in an analysis of terrestrial plant distributions), should be represented as missing data in the input grid (instead of zero). Alternatively, they can be excluded using an analysis mask. Doing either has the potential to save considerable amounts of memory, particularly when a study area is of an irregular or convoluted shape.

Quality of habitat is a primary piece of information to use in spatial conservation planning. There are two alternatives here. First, variation in local habitat quality is indicated by the numbers (occurrence level, carrying capacity) assigned to each grid cell. These numbers do not have to be binary; they can be probabilities, abundances or any positive numbers (Section 3.1.1). Second condition layers can be used to further describe local habitat quality, an option that is particularly useful when analysing the distributions of habitats. Here, the distributions of major habitat types (forest, peatland, grassland, etc.) are generally represented as binary grids, with 1's or 0's used to indicate the presence or absence of the habitat in individual grid cells. Then, variation in habitat quality can be incorporated by linking a condition layer to the distribution grids. The condition layer has numbers varying continuously between zero and one, effectively modelling the proportion of habitat quality remaining. One advantage of using a condition layer is that one gets information about how much habitat condition has been lost. Also scenarios of land use change could be evaluated easily. The condition layer is one of the simplest and more useful features in Zonation. However, including the habitat condition in the analysis can have a significant effect on the results. This is because the values in the condition raster (between 0 [completely degraded condition] and 1 [no degradation]) act as initial multipliers in Zonation's cell-removal process and thus can drive the solution strongly.

Connectivity. Some form of connectivity is frequently used in Zonation analyses. Use of connectivity is logically mandatory when grid cells are small (such as one hectare), because then adjacent grid cells are strongly linked, both by alteration of physical conditions (e.g., microclimate), and by population dynamics of the resident species. Connectivity can perhaps be ignored when grid cells are very large, for example, in the order of 10 km2 or larger.

Zonation has nine different options for applying connectivity, all of which are described in the manual and in Lehtomäki and Moilanen (2013). These can be divided into (i) structural aggregation methods, such as the boundary length penalty (BLP), (ii) parametric single-feature connectivity methods, such as distribution smoothing or the boundary quality penalty, (iii) a pairwise connectivity interaction, (iv) directed connectivity in rivers, (v) connectivity amongst multiple partially similar habitats (matrix connectivity) and (vi) path-like connectivity. Note that many of these methods have parameters, which allow tuning of the spatial scale across which connectivity effects are applied. For example, knowledge about the size of the home range of the species can be converted into a spatial scale of connectivity by taking the radius of a home range approximated as a circle.

It goes beyond the purpose of this document to try to describe all the potential uses of

these methods, but here are a few pointers. The BLP can be safely combined with all other connectivity methods. The pairwise connectivity transform is useful e.g. when considering connectivity to an existing protected area network (Lehtomäki et al., 2009), or when linking present and future distributions of species in climate-change motivated analysis (Carroll et al., 2010). Matrix connectivity is frequently used when connectivity among multiple features is desirable. For example, a mosaic of different forests and wetlands could potentially provide a more viable opportunity for long-term biodiversity persistence, than forests and wetlands protected separately from each other.

3.1.4 Costs, opportunity costs and land use constraints (alternative land-use, costs, hierarchic masks etc)

Costs are often used in spatial conservation planning. In Zonation it is possible to account for not one, but multiple cost components simultaneously. These can represent, for example, direct land acquisition cost and/ or opportunity costs for various businesses or alternative land uses. It is your decision whether costs are important for your application. Note also that getting accurate information about costs is not always easy.

The motivation for including costs in a Zonation analysis is obviously the achievement of cost efficiency. The following outcomes are generally true for analyses that account for cost. (i) If the area of the solution (e.g., conservation area network or its expansion) is kept constant, then accounting for costs reduces the "ecological" quality of the solution. This is because the ecologically optimal fixed-area solution is achieved when the solution is based solely on biodiversity distribution data. Any constraints, including costs, can only reduce the ecological quality of the solution. (ii) Using costs will always result in higher cost efficiency, i.e., a higher ratio of aggregate benefits to aggregate costs. Maximising cost efficiency is nevertheless different from maximising ecological quality,

particularly when there is marked variation in costs between different types of biodiversity features.

Cost layers are entered into analysis in the same format of spatial raster data as the biodiversity feature distribution layers, but with the difference that cost layers are indicated by negative weights (instead of positive) in the biodiversity feature list file (a.k.a species list file). When costs are used, Zonation effectively attempts to resolve spatial conflict between ecological benefits and economic costs. In such a prioritization, lowest ranks will be assigned to areas that are of low ecological importance but that have high costs; conversely, highest ranks will be assigned to ecologically important areas with low costs. See Moilanen et al. (2011) for a worked example. The relative influence of cost layers in prioritization is tuned by changing the weights given to layers. If cost layers have extremely high negative weights then cost avoidance strongly drives the prioritization. If weights assigned to cost layers have low negative value, the prioritization is more of a trade-off between biodiversity representation and overall costs. What exactly counts as high and low weight will again depend on your data and application (Section 3.1.2).

In addition to cost layers, three other types of layers can be used to constrain the solution to achieve particular outcomes: 1) planning units, 2) hierarchical mask and 3) administrative areas.

When a **planning unit** layer is used, all grid cells occurring within a planning unit are removed and ranked at the same time, rather than the cell-by-cell ranking that is normally used.

Hierarchical mask assigns a "mask level" to each grid cell. One could for example have a file in which areas already protected are given the highest mask level (e.g. 2), areas available for protection are given an intermediate level (1), and areas unavailable for conservation are given the lowest mask level (0). Zonation will

produce a hierarchical ranking in which lowest ranks go to areas unavailable for conservation and highest ranks to present protected areas. Intermediate ranks will be assigned to areas available for protection, and will indicate the priority for expansion of protected areas. Hierarchical masks are needed for example when developing an expansion for a pre-existing conservation area network.

The division of the landscape into **administrative areas**, for example, districts within a region, or regions within a country. This type of analysis can be used when biodiversity and costs need to be accounted for separately within each administrative unit.

3.2 THE ROLE OF EXPERT KNOWLEDGE

Building the model of spatial prioritization for a Zonation project typically requires substantial amounts of expert judgement simply because the direct empirical data is lacking. All phases of building the model of spatial prioritization discussed above (sections 3.1.1 - 3.1.4) often rely – and sometimes heavily – on expert judgement. Paying attention to the reliability and accuracy of the expert-derived information is therefore something you will want to pay attention to.

Using expert knowledge in quantitative analysis is quite common in conservation science even if there are some serious known limitations. Most notably, expert judgement may be biased, poorly calibrated or self-serving all of which may lead to poor decision-making (Martin et al., 2012). Fortunately, there are ways to mitigate some of the well-known biases in expert elicitation and to increase the reliability of the analyses partly based on expert knowledge. At the outset of a Zonation project it is worthwhile to try to estimate the amount and type of expert judgment that is required and assess whether more formal expert-elicitation techniques could be used. Overcoming some of the limitations of expert judgement can be done through analytically testing the knowledge of experts, providing training and feedback for the experts and using

structured elicitation procedures such as the Delphi technique (Burgman et al., 2011).

Who are the experts then who you should be engaging? Once again, the answer to this question will depend on the specifics of your project and you will probably face a whole continuum of expertise from which to choose. Typically, the experts come from research and management organizations, private sector, NGOs and from organizations responsible for the actual decision-making. The level of expertise of any given single expert will of course vary depending on the subject matter and will range from simple understanding to substantial contributory expertise (Burgman et al., 2011). From practical point of view, however, more attention should go to examining knowledge claims critically rather than concentrating on formal definitions of expertise and who to include in the group of experts (Gregory et al., 2006). When considering the wider conservation decision-making process and implementation, it is also important that you are explicit about the roles of experts and decision-makers involved. In other words, make sure that experts are not presented with decisions that should really be made by decision-makers (i.e. people with a legitimate mandate to make decisions). In more Zonation-specific terms, be careful not to confuse defining the objectives of the prioritization with the factors needed to address those objectives. Doing so may decrease the transparency of the prioritization and ultimately the acceptance of the results.

4 DATA PRE-PROCESSING

4.1 DATA MANAGEMENT

All the generally accepted best practices of data management certainly hold true when designing, implementing and executing a Zonation analysis. In fact, there are several reasons why you should probably spend more time than on average on thinking how you organize, process, store and document your data.

Iterative pre-processing. It is often not clear at the early stages of designing and building the model of spatial prioritization how the data available should be translated into low-level factors needed to address the high-level objectives of prioritization. Therefore, the data pre-processing involved in building the model of spatial prioritization is typically iterative and experimental in nature, producing potentially large quantities of intermediate files and final input data.

Large amounts of input data. For anything but the most simple and localized analyses, the resolution and number of input features will stack up to a large amount bytes that needs to be stored on your hard drive. Most analyses will not end up consuming terabytes of disk space, but there is usually enough data involved to justify some advance planning.

Documentation. Especially when you are using Zonation to support operative decision-making, you will have to be able to show clearly the whole pathway of how the input data is processed in accordance to the model of spatial prioritization and how the results relate to the inputs. This task becomes excruciatingly difficult unless the processing (both pre and post) steps and the data provenance are well documented.

Repeatability. If you are interested in developing a conservation prioritization process rather than executing a once-off analysis, you will most probably want to repeat the analysis at some point in the future. Systematic and well-documented data management will increase the repeatability of your analyses.

Unfortunately, we cannot offer any silver-bullets for how exactly you should go about in planning for data management, but especially for larger projects it worth considering the usual stages of data management planning:

1. **Information about data and data formats.** What types of data are included? How will the data be acquired? How will the data be processed? What formats are used? How are version control and back-ups handled?

2. **Metadata on the content and formats.** What metadata are collected? Are specific standards used? How will the metadata be collected and stored?

3. **Policies of access, sharing and re-use.** Are there any obligations (e.g. legal, or from data owners and funders) related to data usage? How will the data be shared? What/who are the intended future uses/users of the data? How will the data be cited?

4. **Long-term storage and data management.** Which data are stored in the long-term? Which persons or organizations should be contacted with inquiries about the data?

5. **Budget.** How much resources (human and monetary) are available for data preparation, management, documentation and preservation? How will the costs be distributed?

For many projects using Zonation, coming up with a full data management plan may seem like overkill, but even thinking about these issues is certainly useful. A decent data management plan together with an emphasis on proper documentation of the various processing and analysis stages can also mitigate the effects of turnover in the team responsible for the implementation. With this, we refer to an all too common situation where the main responsibility of executing the data processing - and often Zonation analyses as well - is given to only a few people or a single person. Without the supporting data management plan and metadata documentation the capability to repeat and develop the analyses in the future is compromised should the key personnel move to other positions potentially in other organizations.

Box 2. Common pitfalls in data pre-processing.

Depending on the data and the type of Zonation analysis you are developing, several things can go wrong in pre-processing your data. Look out for the following common problems to avoid errors and to manage your Zonation projects more efficiently.

- Inconsistent resolution and extent between input features. Remember that Zonation requires that all of the input data you are using is in the same resolution (spatial grain). Zonation will make some checks while loading the input data, but there are cases in which features with different spatial resolution and extent can be loaded in without warnings. Generally, the biodiversity features are checked for consistency, but checks for e.g. different analysis masks may be less stringent.

- Misspecified NoData-values can cause problems later on in Zonation. Internally, Zonation uses GDAL to read in values from your input feature rasters. The NoData-value in your input rasters can be misspecified so that you think it is for example -1 whereas in reality it is something else (e.g. -3.40282346639e+038; different GIS software and programming libraries have different defaults for NoData-values). In this case Zonation will not know that -1 should be treated as NoData and erroneously treats cells with that value as legitimate data. Fortunately this type of problem is typically easy to notice as your results will have (priority rank) values in all cells across the land- or seascape.

- Coordinate reference systems (CRS) can also cause various problems in data pre-processing. While Zonation uses spatially-aware GDAL for reading in spatial data, it does not care about CRSs or projections. In other words, make sure that all input data are in the same CRS and that you correctly handle potential CRS transformations as a pre-processing step done in GIS. Note also, that with extensive spatial extents the choice of CRS and projection matters. Most notably, if you are working with global datasets you must take into account the fact the cells at different latitudes will be of different physical area.

- While in practice Zonation has no limitation to how large raster files (i.e. features) it can read in, having a large number of very large files (size gigabytes / file) will make managing your Zonation project more difficult. If this is the case, it is a good idea to consider different ways of reducing the raster file size. Many GIS software and spatial programming libraries (such as GDAL) will let you compress your rasters, but might not do it by default. We recommend using GeoTIFFs with compression. Explicitly setting the raster data type to the smallest possible can greatly reduce files sizes. If you have input features with values ranging between 0 and 255, use 8-bit unsigned integer type instead of 32-bit float and so on. Also, if data still remains unwieldy big, consider dropping resolution by summing e.g. 2x2 grid cells into a single cell of coarser resolution. If doing so, only the locations of features inside coarser cells are lost, but the correct amount (sum) is retained (see Section 4.2).

4.2 TYPICAL DATA PRE-PROCESSING STEPS

The exact nature of a data pre-processing sequence naturally depends on the specifics of the project and is hard to generalize. The input data can be based on variety of different sources including (but not restricted to) expert knowledge, modelling, and indirect (surrogates) and direct observational data. It is also possible to have the input data being derived from a combination of these source types. There are few steps, however, that are almost always encountered when pre-processing the input data.

1. **Data type conversion.** Zonation uses exclusively raster files as input data and so whatever type of spatial data is used originally, it has to be converted to raster format at some point of the pre-processing. The most common spatial type conversion encountered is converting vector (polygon) data to raster format. When converting categorical vector data, such as habitat or land cover data, into input rasters, there is a conceptual and practical difference in whether you convert the whole vector file into a single raster file with raster cell values indicating the class they belong to, or whether each class is converted into a separate raster file (see Section 3.1.1). While doing the conversion, you will have to pay attention to both of the following issues.

2. **Setting the geographic extent and resolution.** All the spatial input data for Zonation need to have exactly the same spatial extent and resolution. Be careful with this! For some input data Zonation will give you an error if this is not the case, but not always. If your data is derived from a separate modelling step, such as using species distribution models, then you should produce all the outputs of the modelling already in correct spatial extent and resolution. In other cases (including data already in raster format), most of the tools used to do the conversion let you

set the extent and resolution correctly.

But what then should these values be? Extent is usually to overall extent of your study area, which if not defined explicitly, is the spatial union of the extents of all of your inputs. Resolution is often trickier to define. Selecting between a finer and coarser resolution is essentially a trade-off between analysis detail and computational time. Note that Arponen et al. (2012) showed, that Zonation analyses based on the same data may produce different results when ran using different initial resolutions for the input data. Furthermore, they concluded that the resolution of the analysis should correspond to the relevant ecological (i.e. taking into account species dispersal capabilities, see Section 3.1.3) and management scales (e.g. the size of planning units used). If you have the data and hardware resources available and the increase in the computational time is not prohibitive, we suggest that you go for as fine a resolution as meaningful and possible. Even if using finer resolution becomes computationally intractable, it may still be better to convert your vector data in high resolution and then aggregate the data to a computationally more feasible lower resolution grid (see below).

3. **Determining the cell value.** As explained in Section 3.1.1, the absolute value of a raster cell in a given raster feature can only be meaningfully compared to the values of the other cells within that same feature grid. For the most typical input data types (binary presence/absence of a species range or a habitat, probability of occurrence), the interpretation of the range of values within a single feature is straightforward. If you are producing the input features using some modelling technique, such as a species distribution modelling software, you will want to check what the output values of that technique are and whether it would make sense

to apply some form of a transformation on the values before entering them into Zonation.

In case you have to aggregate your data to lower resolution for the actual Zonation analysis, some care must be taken when deciding the appropriate methods of determining the value of the aggregate cell (Lehtomäki and Moilanen, 2013). For the more common data types (probability of occurrence, coverage of habitat type) simply taking the sum of the values of higher resolution cells that fall within the lower resolution aggregate cell will be a good choice. Although the exact spatial arrangement of the higher resolution cells is lost in this process, at least full quantitative information from the higher resolution data is retained. In case your input biodiversity features are already some kind of aggregate quantities (such as species richness), you may consider using e.g. the maximum of high resolution cell values instead of taking a sum. One final consideration you should pay attention to concerns cells with missing data and cells with the value 0. Generally speaking, if you really have missing data in some parts of the input rasters, then these cells should be assigned an appropriate NoData value instead of 0. It does make a difference for example when connectivity transformations are in use. Also, missing data require little RAM memory in Zonation, making them preferable to 0's also from that perspective.

4.3 AUTOMATING PRE-PROCESSING TASKS

The most time spent in executing a Zonation planning project is typically spent in preparing and pre-processing input data. This is because the complex interplay between setting of the high-level objectives, identifying the factors relevant for addressing these objectives and the actual data available for doing so. This process is almost inevitably iterative in nature with the different pre-processing steps being repeated with varying frequency. Taking the more long-term perspective that is necessary especially for operative planning, the prioritization

analyses will typically have to be repeated at given intervals because objectives might change, data get updated and so on. Therefore, it is often a good idea to automate the pre-processing tasks as much as possible given the resources available for a project. As many of the pre-processing tasks involve manipulating geospatial data in different ways, the exact way of implementing automation depends on the tools you are using for that manipulation. Below we quickly describe two approaches that we have used. While other software than those listed here exist, these examples should broadly cover the most typical uses.

Automation using a GIS. Using a GIS software, such as ArcGIS and QGIS, is probably the most common way of pre-processing any kind of spatial data. Doing things manually is certainly feasible if you have to do it once or twice, but manual approach gets quickly unwieldy if the tasks have to be repeated several times or if there are thousands of layers to begin with. Using automation tools built-in to the software may thus be a good choice. For example, ArcGIS (version 10.2.1 at the time of writing) has an integrated facility called model-builder which lets you chain individual tools together using a graphical user interface. These chains, or models, can later be repeated relatively easily and applied to different data and possibly using different parameters.

Automation using a programming language. The functionality of many GIS software packages can also be controlled programmatically using a high-level language such as Python (Python Development Team, 2014). Furthermore, many programming languages have bindings to popular open source geospatial libraries such as GDAL (GDAL Development Team, 2014) that can be leveraged to do the heavy lifting. The programming approach usually gives you more intricate and flexible control over the tools, but it also requires more know-how from the team responsible for implementation.

Automation does not, however, come without a cost. Developing automated pre-processing

toolchains can require significant amounts of expertise and time from the project team. Furthermore, if the tools are to be used in the future they require active maintenance which again requires dedicated resources. Besides time saved on repeating the pre-processing tasks, having automated toolchains has more advantages as well. An important, but often overlooked feature of having automated toolchains (such as model-builder models or scripts) is that they provide detailed documentation about what was actually done to the data before it was fed into Zonation. This may sound like trivial, but doing things manually rarely leaves a trace about what was done and unfortunately people often do not document what they did. Notice also that automation is not restricted to pre-processing of data; as we show later, many aspects of running Zonation and post-processing of results can also be automated.

5 SETTING UP AND RUNNING ZONATION

5.1 DEVELOPING ZONATION VARIANTS

Before you start implementing your Zonation analysis (i.e. developing the actual input and configuration files needed to run Zonation), it is a good idea to revisit your objectives. Sketch down your model of spatial prioritization if you have not done so yet, and take a look at it. Try to identify the important input data and the different analysis components such as connectivity and weights. The analyses including all relevant data and analysis components are called "production runs". These potentially complicated analysis runs are the ones that should inform land use decision making. However, Zonation production runs are almost never developed in one go. Rather, they are developed in stages through what we call "development variants" or "variants" in short. There are several important reasons why projects are done like this and we strongly recommend following this approach.
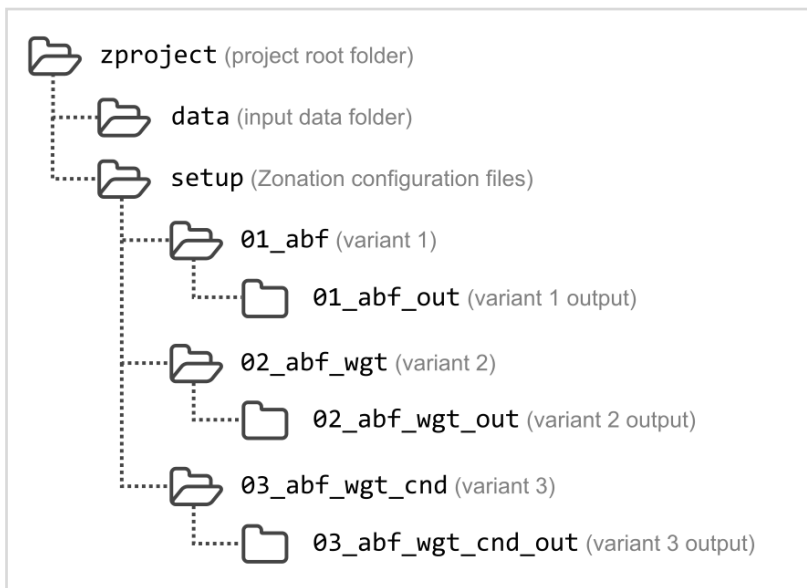
Figure 2. An illustration on how to organize the files for your development variants in a systematic way in your Zonation project. Subdividing your data into appropriate subfolders may also be a good idea.

Development variants proceed from simple to complicated, and their number ranges in our experience from several to some dozens. The main logic behind this incremental development can be summarized by the following reasons.

Catching errors in input data or Zonation settings. Each development variant brings in new data or adds structural analysis components via Zonation settings. If you apply lots of data and analysis settings in one go, the likelihood of you not noticing strange outcomes is increased. When data is entered in blocks or analysis settings are switched on one by one, it is comparatively easy to notice if something weird has happened. For example, we once noticed stripes appearing in priority rank maps, which was caused by unnoticed errors in habitat models. Sometimes nothing happens to aggregation in the priority rank maps after some form of connectivity has been switched on. In this case, frequently the parameters of connectivity are wrong or the option is not truly on.

Getting started early. Especially with big projects, it takes a long time before all input data layers (species distributions, habitat distributions) have been updated, acquired, processed and verified. It makes no sense to wait until this point before Zonation analyses are initiated. While you are finalizing the data, it is possible to develop structurally correct analysis setups, enabling rapid completion of analyses after final data is ready. Effectively, you can develop proficiency with Zonation while still fighting with data.

Sensitivity analysis. Development variants can be used to understand the sensitivity of the final product to data components, parameter values, or analysis settings such as connectivity.

There are no fixed rules for how the development variants should be constructed, but the following illustrates an order we have frequently found to be operationally sensible. Note, however, that the sequence does not

include all the potential features you can include in your model of spatial prioritization and the consecutive Zonation implementation. For an exhaustive description, consult the Zonation manual.

1. Typically a first analysis is done with partial data (e.g. species only) and without any special analysis settings. The motivation of this analysis is to see that data layers are in format that Zonation understands and that a simple analysis seems to run as expected.

2. After the initial entry of some input features (see 3.1.1), rest of the major data blocks (e.g. habitats, ecosystem services) should be added in. At this stage a sensible weighting between features can also be adopted (see 3.1.2). This analysis starts to give an indication of what to expect, but important considerations like connectivity or costs are still missing.

3. Still early into the analysis, habitat condition should be linked to layers that describe (binary) distributions of habitat types.

4. After the basics, items 1-3 are operational, some form of connectivity should be included. This could involve distribution smoothing for species or matrix connectivity for partially similar habitats or both (see 3.1.3). Also the boundary length penalty could be applied.

The analysis now has biodiversity features, sensible weights for them, habitat condition, and basic connectivity features set up. The next variants start getting close to production variants. Note that there can be several of those: for example analyses with and without costs give different but meaningful outputs.

5. If the aim of the analysis is expansion of a protected area (PA) network, a hierarchical mask for the PA network is used next. Then the highest-priority areas outside existing PAs are the areas that most cost-

effectively fix deficiencies of the existing PA network. Effectively, Zonation does a gap analysis and gap filling in one go. If the aim is development of a PA network from scratch, a hierarchical mask is not needed.

6. If developing on top of an existing PA network, one could also apply interaction connectivity to include connectivity to the existing PA network in priorities (see Lehtomäki et al. 2009 for example). This should only be done if connectivity to existing PAs is a priority.

7. Accounting for costs. One or more cost layers can be entered into analysis; these are distinguished by giving them negative weights in the feature list file. If costs are used, Zonation tries to resolve conflicts between ecological benefits and costs; see Moilanen et al. (2011) and Kareksela et al. (2013) for examples.

Typically, you can expect to come up with 2-4 production runs, each providing information valid in its own right. Usually production runs are done with and without considerations that are in partial conflict and which may represent different preferences of stakeholders. These could be, for example, analysis with and without costs, with and without ecosystem services or with and without connectivity to the existing PA network.

As a final consideration, we note the possibility of doing runs either pixel-based (default) or doing them on planning units (PLUs). PLUs can represent, for example hydrological catchments, land parcels, spatially distinct patches of different habitats, etc. If analysis is done on planning units, the priorities come out as slightly different from the averages of pixel-based analysis. This can be, for example, because the occurrence of a very rare feature in a pixel can raise the priority of an entire PLU, whereas in pixel-based analysis and in the absence of connectivity only the pixel itself is influenced. When working with PLUs, one should pay attention to both the mean priority rank of each PLU and the so-called "distribution sum", which is the summed fraction of occurrences of all features in the PLU.

To illustrate, a very small PLU can be of very high relative priority but relatively unimportant absolutely, whereas a very big area can be relatively low mean priority but very important in the absolute sense. Both numbers need to be considered when evaluating PLUs. Favouring mean ranks is justified when cost-efficiency is the rule and the habitats in question are such that local-scale connectivity is sufficient. When count of areas is more important than costs (area size), one should shift attention more to the distribution sums.

5.2 ORGANIZING A PROJECT

In the previous section, we explained how it is often useful to develop the final analysis versions (production runs) by starting from a simple version and adding data and analysis components preferably one at a time (so called development variant). We also gave some guidance on the order in which data and analysis components are typically added into the development variants. Zonation imposes no strict specifications on how exactly you should name and arrange your input files, but over the years, we have settled on some best practices that we will go through next. For a detailed description on what these input files are, take a look at Chapter 3.3 in the Zonation manual or Chapter 8 in "A quick introduction to Zonation" document.

As starting point, it is convenient to organize all your Zonation input files in a single root folder. This folder constitutes your project folder (see "project root folder" in Figure 2). This folder will house all the necessary input files you need. Write down plain text descriptions of your development variants and devise a coding scheme you can use to create informative file and folder names in the development variants. Note that the input rasters are best organized under a separate subfolder underneath the project (see "data" in Figure 2). Then, you

Box 3. Common pitfalls in setting up and running Zonation.

While technical execution of Zonation analyses is not the focus of this document, here are a few pointers to things that relatively commonly go wrong in Zonation setups. Check these out first if an analysis does not seem to work as expected.

- Parameter name spelled wrong in the run settings file (i.e. the .dat file, see Zonation manual section 3.3.2.3). This causes the option specified by the parameter to not be used in the analysis without producing a warning. The run_info.txt file is a standard Zonation output, showing information that may help you verify which options were used and which were not.

- Parameter listed under wrong section (section headers are in square brackets, e.g. [Settings]) in the run settings file. Note that also the sections headers are case-sensitive. A misspelled section header will cause all the options defined under that section to be silently ignored.

- Zonation stops right in the beginning and an error message is given but you can't see what is wrong. If the error message says something about a file name, you probably have incorrectly spelled either a file or folder name. Check for odd characters or letters that are easily confused, (1, i, l, o, O, 0 etc.). Alternatively, check that folder names work correctly in absolute or relative paths - you might need to add or fix the path to the folder where the files are. If you are using relative paths ("../../" etc.), make sure you are referring to the correct level in the folder hierarchy.

- Parameter for dispersal kernel wrong by orders of magnitude. If, for example, the unit of measurement in an input raster file is meters but connectivity parameters are entered per kilometer, the scale of the connectivity response is wrong by three orders of magnitude (see e.g. Zonation manual section 3.3.2.2). If you switch on a connectivity option and either (i) no change occurs in the priority ranking, or (ii) you get highly aggregated or circular patterns, you might wish to check your connectivity settings or parameters.

- •One or many input feature(s) include(s) no data at all, which may lead to odd errors. You should never use a feature layer that has no data, i.e. features with only zeroes and missing data in them. This may be noticed in some but not all contexts in Zonation. You can check this by searching for features that have a distribution sum of zero reported in the run_info.txt file. Check section "Loading feature (e.g., species) data layers" and look for feature files which have y (distribution sum) with value 0 in:

  ```
  Loaded biodiversity feature file #N, file_path, non-missing cells: x, their sum: y
  ```

- Priority rank maps display geometric patterns that intuitively do not appear to be associated with your biodiversity feature distribution data. Check input files as chances are that there has been an error in the development of input files. Dubious geometric patterns such as lines or squares might also appear in analyses that have large areas with low variation in the input features – such as an analysis with only a small number of presence/absence features.

- When a new set of biodiversity feature layers, connectivity, costs, or any other options are entered into an analysis (or switched on), you can always expect a sensible change in the priority rank map. The change is not always large, but it should be there. If no change is visible, the change has most likely not taken place correctly. You may have edited, for example, a different settings or feature list file than you thought you were editing.

can reference the same input rasters from all analysis variants using relative paths. We do not recommend that you replicate input rasters under individual analysis variants as space requirements multiply and data maintenance can become awkward.

Figure 2 displays a coding scheme based on a numeric prefix and 3-letter code to describe the input data and analysis components used. For example, `01_abf` refers to 1st development variant which is using additive benefit function methods in Zonation as a ranking method. This coding scheme does not specify what input biodiversity features are used, but you can include this in the name as well (e.g. `hab` for habitat data and `spp` for species data if you have both). `02_abf_wgt` builds on the 1st development variant and adds weights (`wgt`). The next variant could be called e.g. `03_abf_wgt_cnd` where `cnd` would stand for taking into account the ecological condition of habitats. While the benefit of this kind of naming scheme is that it makes it easier to remember what is included in each development variant, it has at least one obvious disadvantage: variant names can get unwieldy long if you have many development variants. Note also, that not every input data type and analysis component needs to go into the name; only the critical ones that help you to understand the development of the variants towards the production runs are recommended.

After you have decided which coding scheme serves your purpose the best, create a subfolder for each of your development variants using the coding scheme. Again for convenience, the development variant folders are best placed in a subfolder (see "setup" in Figure 2). If you have all your variant folders in a subfolder, it is easier to move or copy them if need be, or place them in a version control system without the need to necessarily include the data folder.

For each of your development variants - as for any Zonation analysis - you need at absolute minimum the input biodiversity features (input rasters, not shown in Figure 2), a biodiversity

feature list file (spp-file) and a run settings file (dat-file). In addition, you will also need a Windows batch file (bat-file) that defines which spp- and dat-files are used, where the outputs are placed and some other parameters. Having a large number of development variants means that you can potentially end up with a large number of input files. Therefore it useful to place all the configuration files related to a specific development variant into that variants' subfolder (Figure 2). Some development variants will share identical configuration files, but in our experience some redundancy caused by using identical files with different names is a small price to pay for clearer organization. Outputs produced by a development variant are also best placed in that variant's subfolder. Note also that when you get the simple development variants set up, you can usually create a new variant by just copying the previous development variant's subfolder, replacing the name of various input files, editing the existing configuration files and potentially adding new configuration files.

Should the input data go into your project folder as well? Everything else being equal, we believe this preferable. As with all files, Zonation does not care where they reside on your system as long as the path definitions in the configuration files are correct. However, if you place all your input files, including the biodiversity feature rasters, into subfolders under your project folder and use relative file paths instead of absolute file paths, your project is self-contained and portable. This does not matter if you are planning to work exclusively on a single computer, but if you ever need to distribute your project to other computers, say to use more computational resources, then a self-contained project folder with relative file paths will save you a lot of trouble.

5.3 TECHNICAL ASPECTS OF RUNNING ZONATION

Zonation v4 runs on both on Windows (e.g. Windows XP and Windows 7) and Linux (tested only on Ubuntu 12.04 and 14.04 LTS) operating systems . Typical hardware requirements vary

per application, and they can and should be anticipated in advance.

Simplifying, computer resources should not be an issue with smallish applications up to a combination of some hundreds of features and a one million element landscape grid size. It should be possible to run such an application with pretty much any present day off-the-shelf laptop with a minimum of 4GB of RAM (random-access memory). Largest problems that have to our knowledge been run with Zonation are studies with tens of thousands of features and/or landscape grid sizes in the hundreds of millions. While exact computational requirements cannot be estimated, the following can usually be used for guidance.

First, it is the combination of (effective) landscape grid size, number of features in analysis, fill factor of the landscape grid and some analytical settings that determine memory needs. The size (product of dimensions) of the landscape grid should not exceed a billion elements. We will illustrate resource estimation with an example.

RAM requirements. A full grid of information takes 8 bytes per grid cell, meaning that a 100 million element grid requires 800 MB (megabytes) of RAM. A 10M element grid would need 80MB and so on. If a computer has 128 GB (gigabytes, 1000 MB) of RAM, a problem with 100+ of such layers can be analysed. It is possible, however, that the landscape grid is not full of information (an issue called fill factor). For example, assume that a narrow island like New Zealand is bounded in a box of dimensions of 2000 x 4000 elements, implying grid size of 8 million elements. However, if the analysis is about the terrestrial areas of New Zealand, then these areas cover only a small fraction of the grid (say 10%) and the rest is sea that should be coded as NoData. In this case, the fill factor would be 10% and ten times more layers could therefore be included in analysis compared to what would be needed if the grids were full of information. Furthermore, Zonation only stores information where

features have occurrences, further reducing memory demands. Some analysis options add to RAM requirements. For example use of matrix connectivity temporarily doubles the memory usage for the layers that the transform is applied on, possibly leading to problems with runs that might otherwise work ok. You can check the memory usage of the Zonation process from the Windows task manager or its equivalent in Linux. Under no circumstances should the RAM memory of the computer be fully used up - if this happens, computations slow down so much that runs will effectively never finish.

Computation time. Computation time increases linearly with features and more than squared as function of landscape size. In our experience computation times start being a problem when the landscape has tens of millions of elements (grid cells) of information and hundreds of layers. If the memory usage of the Zonation process starts to be tens of GB, runs start taking hours. When memory usage is hundreds of GB, computational times grow to days even on comparatively high-powered desktop workstations or small servers. There are several general issues to note.

1. If the landscape is big (many tens of millions of elements) the acceleration factor of analysis (warp factor) can and should in most cases be raised to 5000 or 10000.

2. Server computers frequently have replicated memory buses, which allows more than one Z processes to be run in the time of one.

3. Having many cores on the computer does not help for a single run, because it is the memory bus speed that is the bottleneck.

4. Running a Zonation project frequently involves running many runs during which the analysis setup is developed, data errors are corrected or data is updated. Consequently, running Zonation projects is not an issue when single runs take

seconds, minutes or hours. If this is the case, a large number of development runs can be done online or automatically overnight or over a weekend. When individual runs take more than a day, problems start emerging as cumulative computation times accumulate and runs do not complete overnight. When single runs take in the order of a week, one has to very carefully consider which runs can be afforded. Also, rerunning all due to a data update will take many weeks - not a desirable situation. Therefore, if single runs seem to take weeks or months, one should probably aggregate the landscape to a lower resolution: accumulating 2x2 blocks of grid cells into bigger grid cells will reduce memory needs to ~30% and the computation time by more than an order of magnitude.

5. The vmat-feature can be used to save the state of Zonation after initial data transforms (connectivity, condition, etc.) have been completed. This reduces the time needed to initialize Zonation into minutes even with large problems, thereby saving days of time with large problems that require loading and processing thousands of very large grids. Please see the manual for details.

6. Zonation has not been developed to distribute on a cluster. This means that even if you have access to a computational cluster, it does not help. A typical cluster constitutes of a large number of nodes with relatively low-performance central processing units (CPUs). Zonation will ever use only one core, so a cluster will not help. Nowadays many service providers, such as Amazon or Google, offer cloud-based virtual machines with reasonable CPU and RAM capabilities. These providers usually charge based on time used, so this is becoming increasingly an affordable option. The most straightforward and simple approach - given that you have the funding available - is to get a powerful-enough desktop workstation.

Storage requirements. Hard drive space is not usually a problem, but again can be anticipated. Hard drive space is needed when the landscape grid is very large, implying that output files (e.g. the priority rank raster) can be files of gigabytes in size. For large analyses, avoid using inputs or outputting results as ASCII grids, as these take much more space than packed binary formats such as compressed GeoTIFF-files. One should also be careful with the "save transformed features" setting, as use of this might cause one large grid to be output per each feature, thereby requiring even terabytes of free hard drive space. In general, check output file sizes after first Zonation runs and make sure they are not excessive in comparison to hard drive space available. Thus, you can verify that you have space for your Zonation analysis development variants.

6 WORKING WITH THE RESULTS

6.1 POST-PROCESSING THE RESULTS

Some of the outputs generated by Zonation, such as the rank priority map in jpeg-format or the performance curves created by the Zonation GUI, are immediately useful. Often, however, the results need additional post-processing to better suit your needs. There are of course endless ways how you can post-process the numeric results (e.g. the rank priority raster and the performance curves), but nevertheless there is a set of common post-processing tasks that we have found useful. Note that here we use "post-processing" to refer to any activities that process the Zonation results or combine Zonation results with other sources of information. Any activity that aims at producing informative visualization or numeric outputs from the results we save for the next section. Below we have divided post-processing tasks into two separate categories: tasks done using Zonation (including the GUI) and tasks done using some other software tools.

6.1.1 Post-processing using Zonation

Zonation (from version 4.0 onwards) has several useful automated post-processing functions that you can use to get more information about the solutions or compare the solutions between different runs. "Landscape identification" is a set of automated post-processing procedures that can be set up (before running the analysis) using particular configuration files (see Zonation manual Chapter 3.5). The original landscape identification analysis can be used to generate so called management landscapes (e.g. Mikkonen and Moilanen, 2013) for a selected top fraction of solution. In this top fraction, you can set the distance and similarity between spatially distinct areas to identify sets of locations that can serve as management units. It also possible to use a mask file to constrain the landscape identification analysis to a specific part of the landscape. For each of the units identified, Zonation produces extra statistics about the occurrence of features within the unit. These statistics include information, for example, on the proportions of distributions for each feature in the unit.

Even if you are not directly interested in identifying management landscapes, the landscape identification can be configured to use a pre-defined unit mask. In this analysis, statistics are produced separately for each pre-defined spatial unit. For example, you could be interested in knowing what proportion of the distribution of each biodiversity feature is contained within protected areas, municipalities or within units of some other cadastral delineation. These statistics can then be combined with the spatial data describing the units to produce additional information that helps understand why the priorities are the way they are. Coming back to the example of protected areas, you could provide managers with more information on why particular protected areas are seen as high priority. This type of additional examination is very useful for a broad set of analysis; in fact, we recommend it as a standard procedure whenever meaningful units can be defined for the calculation of statistics.

Zonation is also capable of comparing the spatial overlap between different solutions, using the landscape comparison post-processing functionality. It is possible to configure the options so that the overlap between solution A and solution B can be compared for a given fraction. You could compare, for example, solutions A and B for the top 10% fraction and produce a new output raster that would indicate which cells in the landscape are included in the top 10% in both solutions. By using the "merged map" functionality (see Zonation manual Section 4.3.4) in the Zonation GUI you can further analyze which parts of the landscape are in a given top fraction in solution A, solution B or in both.

You can do a more thorough cross-comparison of different solutions by using the solution loading functionality in Zonation (see Zonation manual Section 3.5.2). The solution load enforces a pre-defined ranking on a different set of data; the ranking may have been developed outside Zonation or using Zonation with different data and/or settings. This type of cross-comparison will not produce information about the spatial overlap of solutions, but it does enable the detailed comparison of the performance of two solutions. You might be interested in how much does the inclusion of a particular connectivity measure influence the performance of your solution (i.e. quantifying trade-offs). Another example where such cross-comparison might be useful is if you are interested in studying whether a set of biodiversity features acts as a surrogate for another. This type of analysis can also be easily implemented by entering all features into the same analysis run and using zero weights for selected taxa.

6.1.2 Post-processing using other software tools

For visualization and interpretation, you will frequently want to import the spatial outputs, most notably the priority rank rasters, into a GIS software. Most of the post-processing tasks that might be applied to Zonation results

are beyond the scope of this document, but the single important, albeit semi-obvious, point we want to make is that there is nothing special about the raster files produced by Zonation. You can work with them as you would with any other raster-type spatial data. Nevertheless, it is useful to keep in mind the following technical considerations:

1. Avoid using the ASCII raster file format (default in Zonation < v4.0) if at all possible. While there is certain appeal in the simple file format, ASCII rasters can become very large and reading in ASCII rasters has its quirks in different GIS software. For example, by default ArcGIS (at least versions up to 10.2.1) will assign ASCII rasters with the integer data type. This means that importing an ASCII-format priority rank raster will not produce the intended output as ArcGIS thinks the raster is an integer raster whereas it actually is a float raster!

2. Zonation knows nothing about coordinate reference systems (CRS). Since it is using GDAL internally for reading and writing raster files, in most cases your CRS will be preserved in the outputs as well. Nevertheless, it is a good idea to check the outputs for correctness. Furthermore, if your prioritization covers very large areas, such as the whole globe, you need to consider the projection effects (i.e. the raster cells will be of different size at different latitudes).

3. The non-spatial tabular outputs (e.g. performance curves and outputs from landscape identification) have a format that is primarily designed to be human-readable rather than machine-readable. In other words, reading and parsing the files into other software platforms takes a little work.

In case you are using R programming language, you might be interested in an R-package "zonator" (Lehtomäki, 2014). This package has functionality for (i) automatically reading in a whole Zonation project and associated results as R data structures, (ii) controlling Zonation programmatically and (iii) post-processing and visualising Zonation results. Zonator makes it easy to integrate Zonation analyses into your other workflows that use R.

6.2 VISUALIZING AND INTERPRETING THE RESULTS

Visualizing the results is an important step in any Zonation analysis. You will want to look at your results immediately after the analysis has finished to verify that the analysis seems to work technically. Visualization is also needed to interpret the results. The basic interpretation includes things such as analyzing the spatial patterns of priorities and the performance of different biodiversity features. At some point, you may also want to take a closer look at the results and do further quantitative analysis either by using the facilities in Zonation or some external tools. This section introduces the basics of visualizing Zonation results as well as important points related to interpreting what you are seeing. If you are working with a real-life conservation planning project, visualizations alone may not be enough to provide the support that is needed. Instead, the results alongside with informative visualizations can be developed into planning products that are relevant and easy enough to understand by the intended users. The exact form and function of such planning products will vary case-by-case, but some considerations are given in Section 7.2.

6.2.1 Standard Zonation outputs and their interpretation

There are two main outputs produced by Zonation: 1) Priority rank maps and 2) performance curves. Almost all other results are derived from these two outputs and consequently the interpretation of the results is based on them. Here we give a short description of each of the outputs, for a more thorough explanation please see the Zonation manual (Moilanen et al., 2014) and the "Quick

introduction to Zonation" document (Di Minin et al., 2014).

The **priority rank map** is the first main Zonation output visualization. By default, Zonation outputs a raster file with the value in each cell indicating the relative conservation priority of that cell. These values are derived from the order of iterative cell ranking (removal). Each raster cell has a value between 0 and 1, meaning that values close to 0 were removed first (low relative conservation value and priority), while high values close to 1 were retained towards the end (high relative conservation value and priority). Assigning a suitable colour scale to these values gives rise to a rank priority map in which the colours indicate the conservation priority over the area of interest (Figure 3). This type of map is quite easy and intuitive to interpret for most people, but it is good to bear in mind that the map – as all Zonation results – shows conservation priorities relative to the data and the model of spatial prioritization that you used. The priority rank map does not necessarily tell you anything about the absolute conservation value over your area of interest, although usually it does: it makes little sense to run a Zonation analysis using data that has no relevance for conservation.

Performance curves, the second main output of Zonation, are automatically produced and exported as text files for each feature (or group of features) during a Zonation analysis. The content of performance curves files can be visualised as a curve that quantifies the proportion of the original occurrences retained for each biodiversity feature, at each top fraction of the landscape chosen for conservation (Figure 4). Performance curves start from 1.0 because the full landscape includes the full distribution of the biodiversity features. At the other end, no areas are chosen,
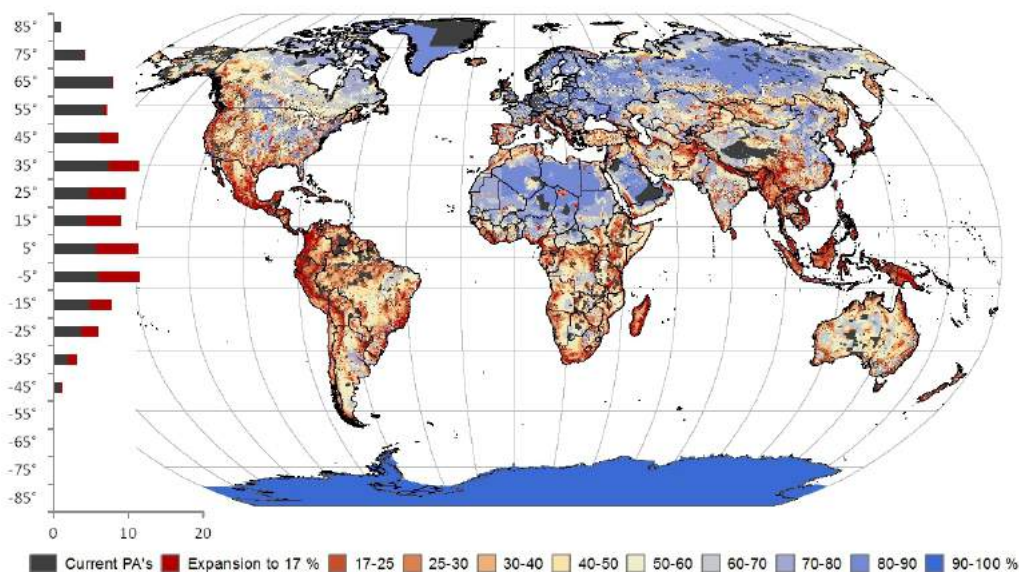
Figure 3. An example of a priority rank map (from Pouzols et al. (2014)). This priority map shows which areas are the globally the most important (and least important) for protected area expansion, based on analysis taking into account the distributions of c. 24 000 terrestrial vertebrates. The red color indicates higher priority for protection. The original pixel values (0-1) have been converted to percentages (0-100%) corresponding to the priority fraction of landscape to be protected. In terms of visualization, the colors have been selected to print out well both on screen and on paper. The projection has been set to equal area (Eckert IV) to retain the comparability at different latitudes. The latitudinal bar chart describes how large a fraction of the current protected area network (grey) and the top priority areas (red) is located in that latitudinal bin.

and correspondingly, the protection level for the feature is zero. Because the number of these curves can be high, it is common to average and visualize curves across feature groups, such as taxonomic groups. Average curves usually are concave because the initial aggregate losses for biodiversity are low (low-priority areas, such as densely populated urban areas contain relatively little biodiversity), but aggregate losses unavoidably accelerate when moving to high-priority areas with high feature richness, rarity and local occurrence levels.

The priority rank map has direct correspondence with the performance curves - top-priority areas selected from the priority rank map include feature representation summarized by the respective performance curves (Figure 4). For example, if we select the top 10% of the priority rank map, we can evaluate the corresponding representation for each biodiversity feature or for broader groups through the performance curves.

6.2.2 Creating visualizations

You can create visualizations of Zonation results in two ways: Using the Zonation GUI or some external tool. The Zonation GUI provides multiple tools for visualizing and inspecting the results as priority maps and
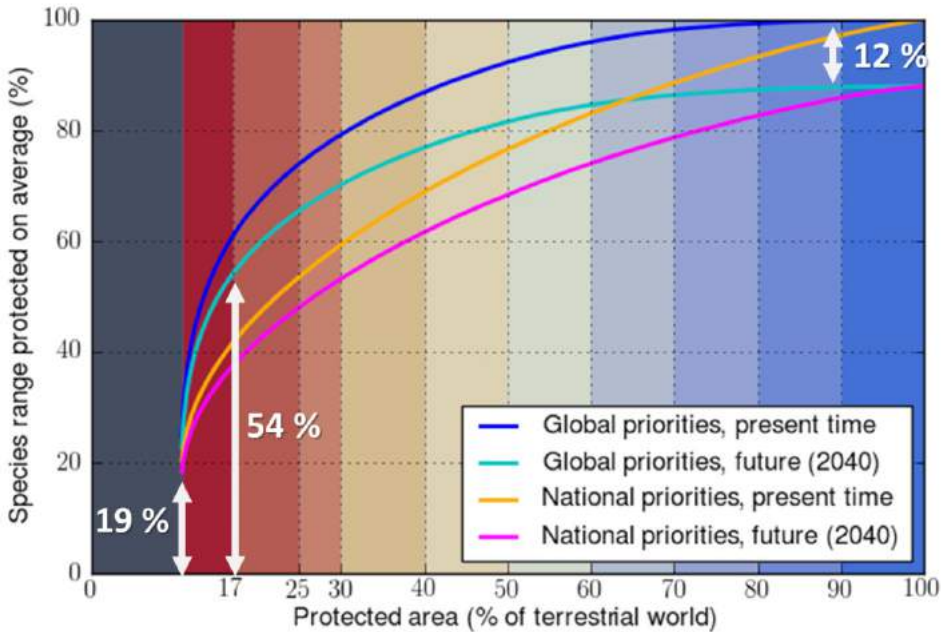
Figure 4. An example of performance curve visualization (from Pouzols et al. (2014)). The figure shows how the biodiversity features perform on average in the priority ranking presented on a map in Figure 3. In this visualization, the background colouring corresponds to those used on the map for different proportions of the terrestrial area. The average performance across the 24 000 input features is shown with a turquoise line (Global priorities, future 2040). To compare the performance of this Zonation prioritization result with three other variants, mean curves for all have been plotted to the same figure. The figure reveals that the current protected areas (11 % of the surface area) protect on average 19 % of the input biodiversity feature distributions. The next 6 % of land surface area (up to 17 %) in the prioritization order would cover on average 54 % of the species distributions. The figure also shows that land use change expected by 2040 reduces the species distributions with 12 % from their current extent (impact of the use of a so-called condition layer in the analysis).

performance curves (Di Minin et al., 2014). If you are using the GUI (Zonation can also be run from the command line), examining both the rank priority map and the performance curves in the GUI should be your first stop in visualizing your results. You can also compare the performance of different features or groups in your analysis by using the interactive plots functionality in the GUI. The merged map tool allows you to make simple visual and quantitative comparisons between the distributions of priorities in different runs. In case you are running Zonation from the command line only, Zonation will produce an image file of the priority rank map using a pre-set colour scale, but the performance curves you will have to plot yourself. Consult Chapter 4 in the manual and Chapter 10 in the "Quick introduction to Zonation" document for an in-depth description of Zonation's built-in visualization capabilities.

For a more detailed analysis of the results, you need to be able the dig deeper into the results themselves. You can do this outside the Zonation environment by working with the numerical outputs produced by Zonation (see Section 3.4 in the Zonation manual). The priority rank rasters (files ending in *.rank.tif) and weighted species range size rarity rasters (files ending in *.wrscr.tif) are usually visualised using a GIS tool such as QGIS or ArcGIS. The performance curve files are numerical tables that you can processes and visualize using some type of statistical software such as R or Excel. The textual outputs created by Zonation, such as the performance curves, are primarily designed to be human- rather than machine-readable. This means that getting the data into some other software takes a little care, but can be done. If you use R, it might be worthwhile to get familiar with package zonator introduced in Section 6.1.2. Zonator includes functionality for automatically parsing all the common Zonation outputs into useful data structures in R. It can also easily visualize many of the outputs and it provides some analytical tools for the outputs.
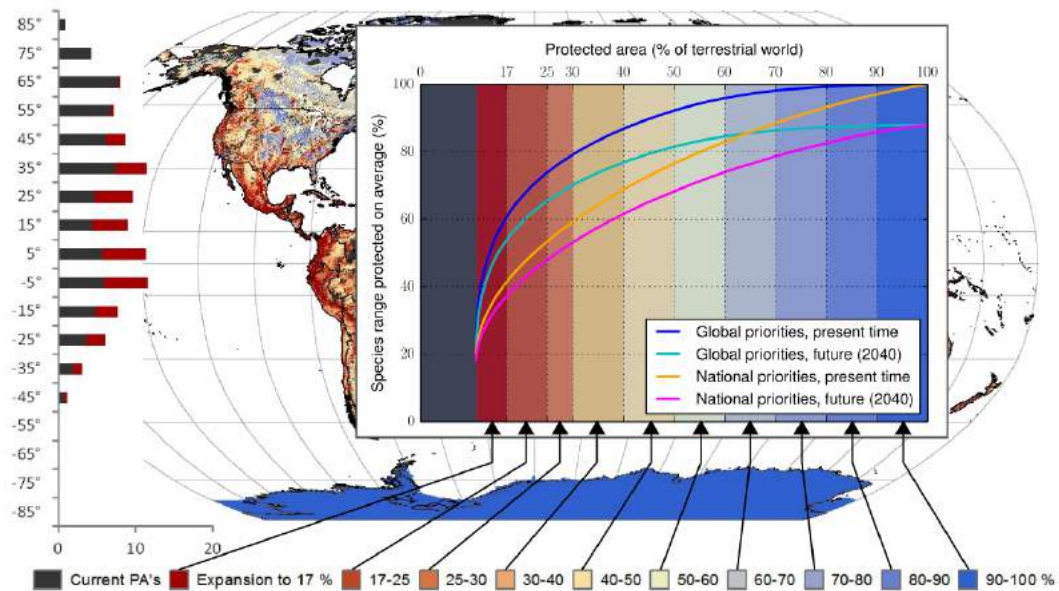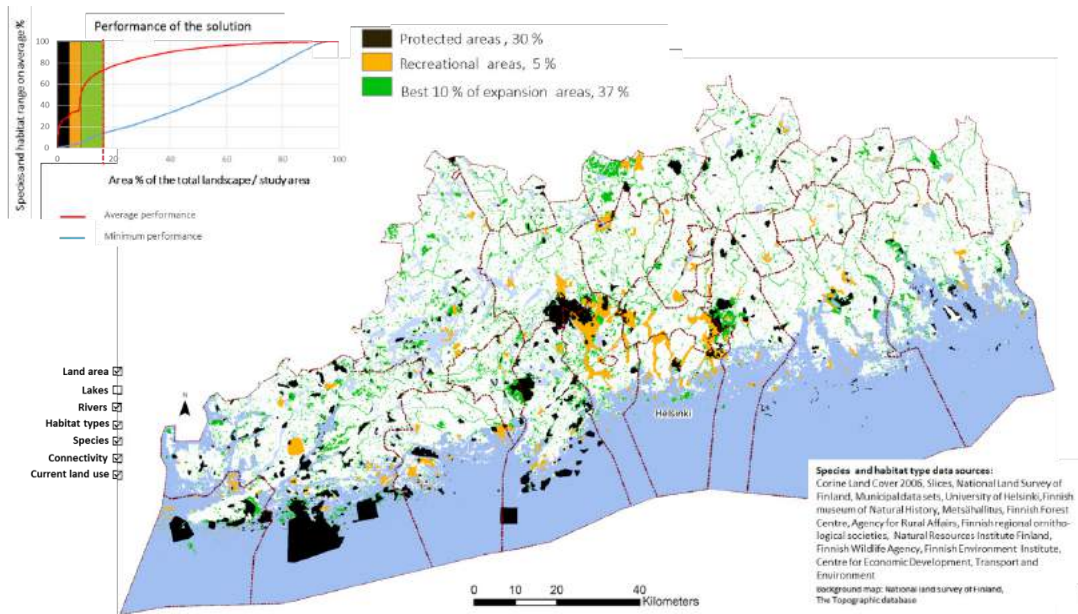

Figure 5. Example of a visualization showing how the rank priority map and the performance curves produced by Zonation can be visually linked. For the explanation of the priority rank map and the performance curves in this figure, see Figure 3 and Figure 4.

As mentioned, the priority rank map and the performance curves always go hand in hand and almost every representation of Zonation results should include both. If you can, it may also be a good idea to link the rank priority maps with the performance curves to emphasize that they are closely connected (Figures 5 and 6).

It is good to remember that while the initial visualizations of the Zonation results can be done quickly, fine-tuning them to meet specific communication needs usually takes more time than you would expect. In addition, normal rules of cartographic and statistical communication apply here. Therefore choosing the presentation style for the intended audience and media is important (see Figure 6 for an example). Equally important is to aim at consistency between different representations of the same results (e.g. maps and performance curves).

6.2.3 Visualising differences between Zonation variants

Assuming that you have followed the suggestions of developing different analysis variants, you may be wondering which variants should go into the final visualisations. In case you have arrived at one or a few production runs (i.e. the "final" versions), then those are typically the ones that you base your visualizations on. Note however, that sometimes it is useful to highlight particular features of the analysis using development variants. Say you were developing your Zonation analysis using a sequence (as in Section 5.1 starting from page 35). The first type of connectivity (distribution smoothing) is introduced in variant 4 and the second type of connectivity (interaction connectivity) in variant 6. Variant 7 has all the analysis components in it and can be considered a production run. While your final visualizations would probably be based on variant 7, you might still want to create a separate map from variant

Figure 6. An example of a visualization showing how different area reservations in a plan (protected areas and recreational areas) perform in biodiversity protection and where additional green areas should be planned. The visualization combines performance curves and the map, and tries to make some relevant settings in the Zonation process visible (connectivity, masking with waters and the source data). Adapted and translated from Kuusterä et al. (2015).

4 to demonstrate the effect of connectivity. Which analysis components you should want to demonstrate is up to you. However, try to keep it simple. Including everything will make the product large and complicated, and might confuse the main message.

There are many ways you can visualize the differences between Zonation variants. You can compare, for example, the performance curves of different variants (e.g. Figure 4). Another common option is to make a map series showing the spatial variation in the priority rankings (Figure 7). If you are interested in a particular fraction of the priority ranking (e.g. the 17 % of the landscape in Figure 8) you can examine the spatial overlap and/ or similarity of different variants in terms of the fraction of interest. For example, the Jaccard index (Warman et al., 2004) and fuzzy similarity index (Mas et al., 2012) can be useful

in inspecting the similarity between priority patterns of different variants.

While group-based performance curves can be very informative, you need to think a little about what exactly you want to visualize. If grouping functionality is enabled, Zonation gives you the following statistics on the representation levels for each group: minimum, mean, weighted mean and maximum. In many cases, you are probably interested in the mean representation levels, but in some cases looking at, for example, the minimum representation level (i.e. the biodiversity feature with the worst performance in that group) would make sense. Summarising information into a statistic such as the mean will also hide other types of potentially useful information. For example, it may be useful to visualize the distribution of individual
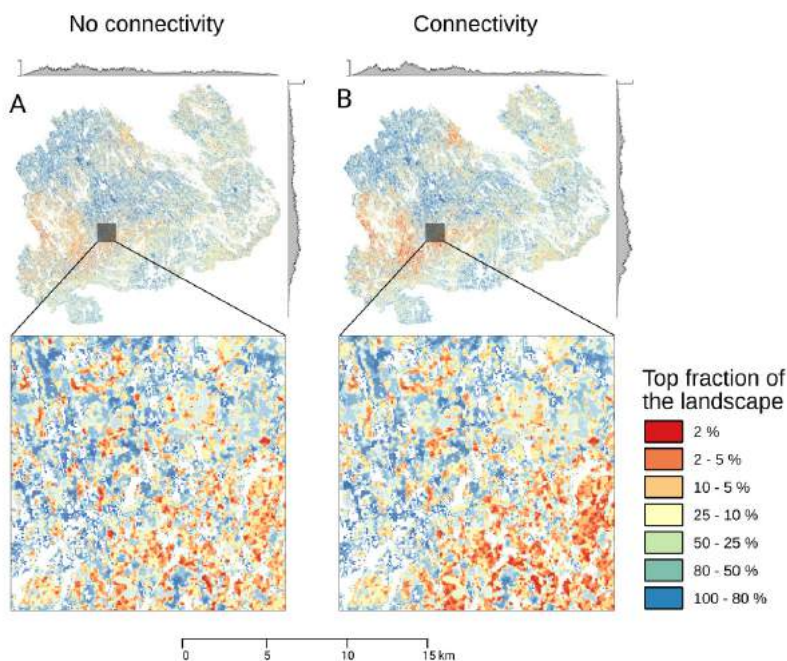
Figure 7. An example of visualisation intended to allow the user to evaluate outputs of two different Zonation variants spatially. Panel A shows a variant without connectivity and panel B with connectivity between different forest types in the landscape. The visualization and the close ups demonstrate how the overall pattern stays the same between the runs, but the more detailed pattern of priorities changes. The marginal plots on top and on the left side of each panel can facilitate the comparisons. They show the count of cells in the top 10% of the landscape along both latitudinal and longitudinal gradients. Adapted from Lehtomäki et al. (2015).
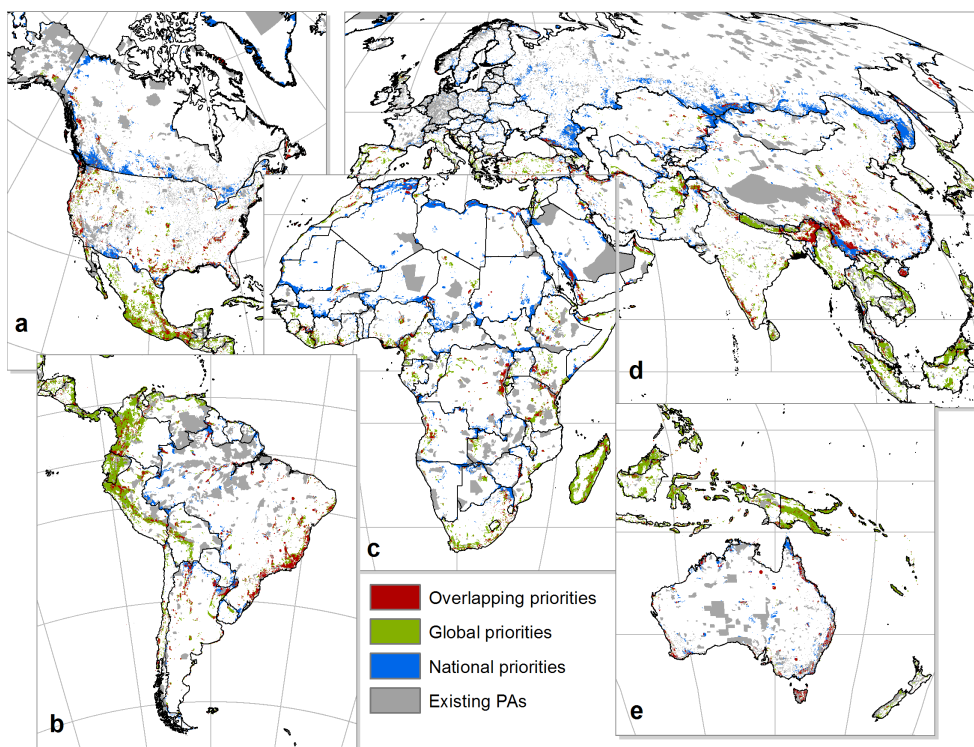
Figure 8. Example of visualisation to compare the top 17% of two different Zonation runs, one made globally and the other made nationally using the administrative units feature of Zonation. In practice, the continuous priority maps have been converted into a binary 0/1 surface before the visualisation. The most distinctive colour (red) has been chosen for overlapping priorities (value 1 in both Zonation variants). Figure from Pouzols et al. (2014).
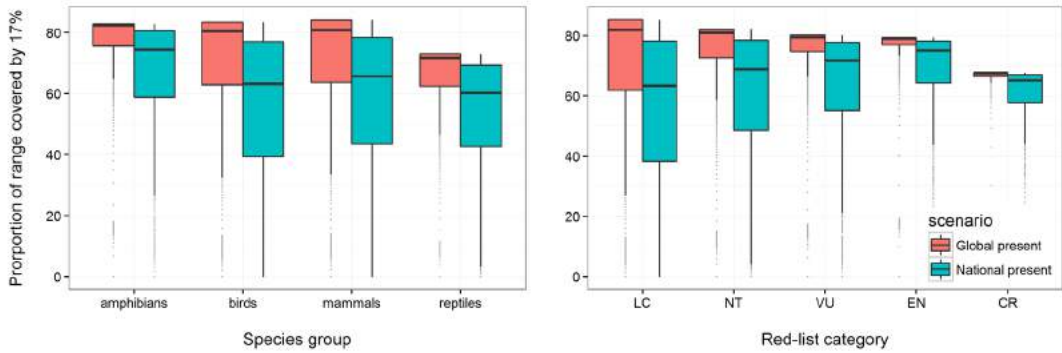
performance levels within the group from the feature-based performance data (Figure 9).

6.2.4 Providing context for interpretation

Providing sufficient context for the results is crucial for the usability of the results. It is important to bear in mind that rank priority maps and performance curves tell very little about how they were produced and it is unfortunately easy to interpret them out of context. In our experience, maps produced with a specific model of spatial prioritization easily start to represent "generic" conservation value of the areas covered by the analysis. In other words, a prioritization based on the occurrence of valuable forest habitats can be implicitly and mistakenly expanded to represent a broader set of biodiversity features (e.g. peatlands, freshwater habitats etc.). It is

therefore very useful to combine additional information to your visualizations providing information on what is the context of the analysis and its results.

Often the intended users, such as land-use planners or other implementers, also need more detailed information on individual planning units. Combining the information produced by the post-processing (Section 6.1.1) analyses to selected planning units (cadastral units, forest stands, individual protected areas or whatever other spatial delineation) will give direct and accessible information on a given location's priority rank. This information is especially useful for illustrating the range-size normalization that is built-in into Zonation. For any given planning unit it is easy to check what features occur there. What is less obvious, however, is how large a fraction of the

Figure 9. Boxplots are a useful way to visualize and compare the distributions of feature-specific performance levels within groups. This example presents the distributions of species ranges covered by expanding the global protected area network to 17% of the global terrestrial area based on Zonation prioritization (Pouzols et al., 2014). Results are grouped by species in different taxons on the left and by IUCN Red-list categories on the right. The colour indicates one of two scenarios corresponding to those in Figure 4.

overall occurrences of that feature is contained in the planning unit. Both these types of information, what features and what fractions of their overall occurrences are within the planning unit, are given by suitably tuned post-processing analysis. The information can then be combined to other spatial data in a GIS or turned into other planning products such as unit-specific description cards (Figure 10). This type of a planning product also reminds whoever is using it about data and analysis features that were used, thereby summarizing the all-important planning context.

Including information on the input data and some key parameters to the visualisations will also remind the user that the results cannot be generalised beyond their context. An example of such a visualization is given in Figure 6, where the data sources and some settings (use of connectivity) are presented as an integral part of the visualisation. Whenever possible, it is good to try to visualize also the analysis process together with the results to make the user of the results aware of choices made. This can be done, for example, by using flowcharts that reveal the input data sets, different steps of the process, and the differences between Zonation analysis variants (Figure 11).

Site 1: [SITE_NAME]

[SITE_DESCRIPTION]

High-priority areas
- Zonation priority area
- Other priority area

© OpenStreetMap contributors
Cartography licensed under CC BY-SA

More information: www.someaddress.org
Contact person: John Doe, john.doe@email.org

The density of biodiversity features relative to the whole area of interest	17.2
Priority rank average	98.2 / 100
Total area	410 ha

INFORMATION BASIS OF THE SITE DESIGNATION		
Habitat features	**Fraction of the whole landscape ‰**	**Fraction of site's features %**
Forest habitats	1.4	0.4
Woodland key-habitats	0.76	0.2
Valuable grasslands	1.85	0.5
Peatlands and mires	0.6	0.2
Valuable nature types	1.6	0.5
Coastal habitats	11.9	3.4
Valuable wind and coastal deposits	300.4	86
Wetlands and small water bodies	2.96	0.9
Species features		
Threatened species (except Lepidoptera)	1.1	0.3
Threatened Lepidoptera	25.4	7.3
Valuable bird areas	0.1	3
Cervid population densities	0.7	0.2
Large carnivores	0.24	7
Habitat features total ‰	321.5	
Species features total ‰	27.7	
Features total ‰ / %	349.1	100

Figure 10. An example of a site information card created for top each priority area in a real-life planning case in Finland. The card proves information on why the site has been selected as a top priority area, and what biodiversity features it contains at high densities compared to the study area in general. In addition to the broad scale priority maps and performance curves, such visualizations are useful when decisions over individual land parcels will be implemented on the ground. This example card was adapted from Kuusterä et al. (2015) and has been simplified for illustrative purposes.
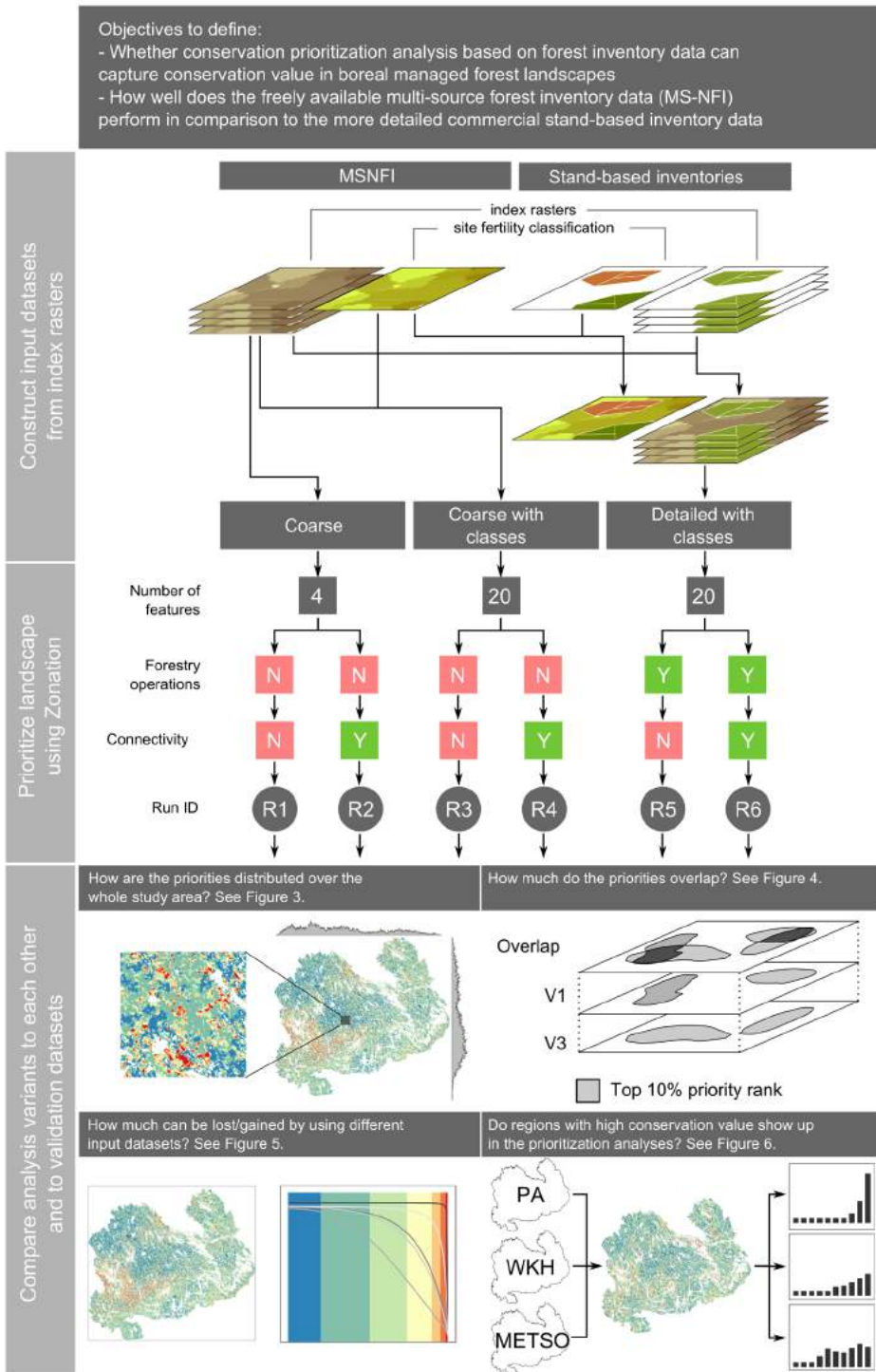
Figure 11. An example of a flowchart visualising a Zonation analysis presented in a scientific article (adapted from Lehtomäki et al. 2015). The flowchart works with the text and describes in detail the different input data sets, Zonation variants and the outputs.

6.3 VERIFICATION AND VALIDATION

In assessing whether a Zonation conservation prioritization analysis is producing informative results, we borrow the concepts of verification and validation from the field of software engineering. *Validation* is the process of assessing whether our spatial prioritization model answers to our problem definition, i.e. how well the prioritization addresses a real-world need (Figure 12). *Verification* on the other hand is concerned with checking whether the implementation of the Zonation analyses corresponds to the model of spatial prioritization and how error-free the implementation is (Figure 12). In other words, validation answers to the question "Are we building the right system?" and verification to "Are we building the system right?" (Easterbrook, 2010).

You should do verification after every stage of the prioritization project (see also Figure 13). You should check, for example, whether pre-processing produced sensible and technically correct inputs for Zonation. Did Zonation produce sensible results, or were there perhaps some errors in the process? If the interpretation of the results does not make sense, is the problem in the input data or in the Zonation analysis configuration? Much of this type of verification is manual checks of various kinds, but it is possible to automate some of the verifications. For example, automating the data pre- and post-processing programmatically (see Chapter 4 and Section 6.1) reduces the risk of manual errors and makes verification easier. For the most common things to keep an eye on in terms of verification, see Box 2 and Box 3 ("common pitfalls").

The fact that the prioritization analysis is technically correct does not mean that the results are useful. Whereas verification is typically a relatively objective and technical matter of checking that the prioritization analysis does what it is supposed, validation involves broader and more subjective assessment. The results of an informative spatial prioritization should of course
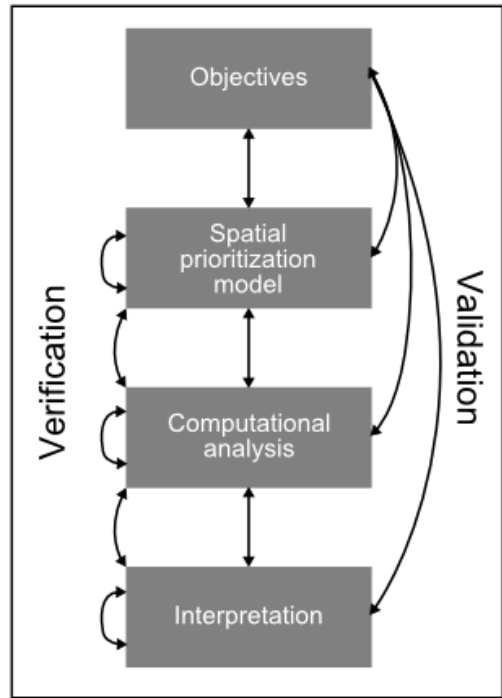
Figure 12. The concepts of validation and verification in relation to different stages of spatial conservation prioritization (see also Figure 13).

correspond to the observed occurrence of biodiversity in nature. For a relatively simple analysis, which does not include any complicating factors such as connectivity or costs, the accuracy of the identified priorities can sometimes be evaluated against independent validation datasets and expert judgement. Introducing abstract components (such as connectivity) into the analysis makes validation harder. This is another reason why the sequential development of Zonation variants is a useful strategy; validating just the final production variant(s) can be difficult because of multiple interacting factors, whereas the validity of simpler development variants is typically easier. Perhaps the most common question you will hear from anyone looking at the results in more detail is "Why is this location a high-priority?" or its complement "Why isn't this location a high-priority?" Using the landscape identification post-processing in Zonation (see Section 6.1.1) is useful here, because it gives insight into the underlying data and helps you to understand

what is driving the priorities in the solution. What was inside an area that increased its priority?

Judging whether the model of spatial prioritization truly captures the occurrence of biodiversity features, our preferences and other relevant considerations is not an easy thing to do. Nevertheless, validation is crucial in evaluating the degree to which the results are useful and applicable. In the longer run, you will also need validation to improve your model of spatial prioritization and its implementation adaptively and iteratively. Even when the prioritization is not perfect e.g. due to incomplete data, one can still ask if it is informative in some sense and a reasonable complement to expert judgment.

7 PUTTING THINGS TOGETHER: ZONATION PROJECTS IN PRACTICE

## 7.1 THE STRUCTURE AND SEQUENCE OF A ZONATION PROJECT

By now, it should be obvious that each conservation prioritization project - those using Zonation included - will be more or less unique and describing the exact structure and sequence of a project is difficult. However, you can always find the stages described in the previous Chapters of this document and in Figure 13 in a Zonation project. While the relative importance and the time taken at each stage will vary depending on the domain and context of the project, the most time and effort is almost invariably spent on building the model of spatial prioritization (Figure 12). In particular, collating and pre-processing the input data needed for the computational analysis typically takes a significant amount of time. What may come as a surprise for many new users of Zonation, the actual stage of computational analysis using Zonation usually takes up a relatively small fraction of the total time spent on the project.

The most notable difference between the structure and sequence of things presented in Figure 13 and real life is that in real life the different stages seldom happen so linearly. In general, the stages given do follow each other in the order given in Figure 13, but there are usually some types of feedback loops involved making the execution of the project iterative rather than strictly linear. For example, the available data often has a great impact on the model of spatial prioritization and potentially even on the objectives: it is pointless to construct an ideal prioritization analysis for which data do not exist. In practice it also often the case that as the project progresses, some new data becomes available while some data originally planned to be used turns out to be of unacceptable quality. Similarly, creating the Zonation development variants (Section 5.1) is often highly iterative as it is not clear from the outset which data and what Zonation options should go into which variants.

Verification and validation (Section 6.3) are especially important forms of feedback in executing Zonation analyses. Doing proper verification between the major project components (building the model of spatial prioritization, computational analysis and interpretation, see Figure 13) guarantees, to the extent possible, that problems and inconsistencies are captured as soon as they arise. Trying to verify the different project components only in the end of the project is going to be not only difficult, but also time-consuming - a simple problem identified late may lead to a "redo all" situation. Validation (i.e. the process of assessing whether the spatial prioritization model addresses the objectives) on the other hand is notoriously difficult and time-consuming (and subjective), and is often simply not done. However, to the extent that it is possible, we believe that you should always carry out validation especially when the work involves real-life on-the-ground land use decisions.

The final point about Figure 13 is that it also shows how spatial conservation prioritization is embedded in the broader context of conservation decision-making. The fact that some components of this context, such as setting objectives and creating recommendations based on the prioritization results, are shown in Figure 13 does not mean that there are no others. Nevertheless, broad elaboration of decision analysis and decision-making is beyond the scope of this document.

In the following sub-sections, we summarize three case studies that demonstrate what conservation prioritization projects using Zonation can look like. The first two case studies have direct relevance for conservation and land-use decision-making, whereas the third case study is an example of a purely scientific study (although with potential practical implications too).
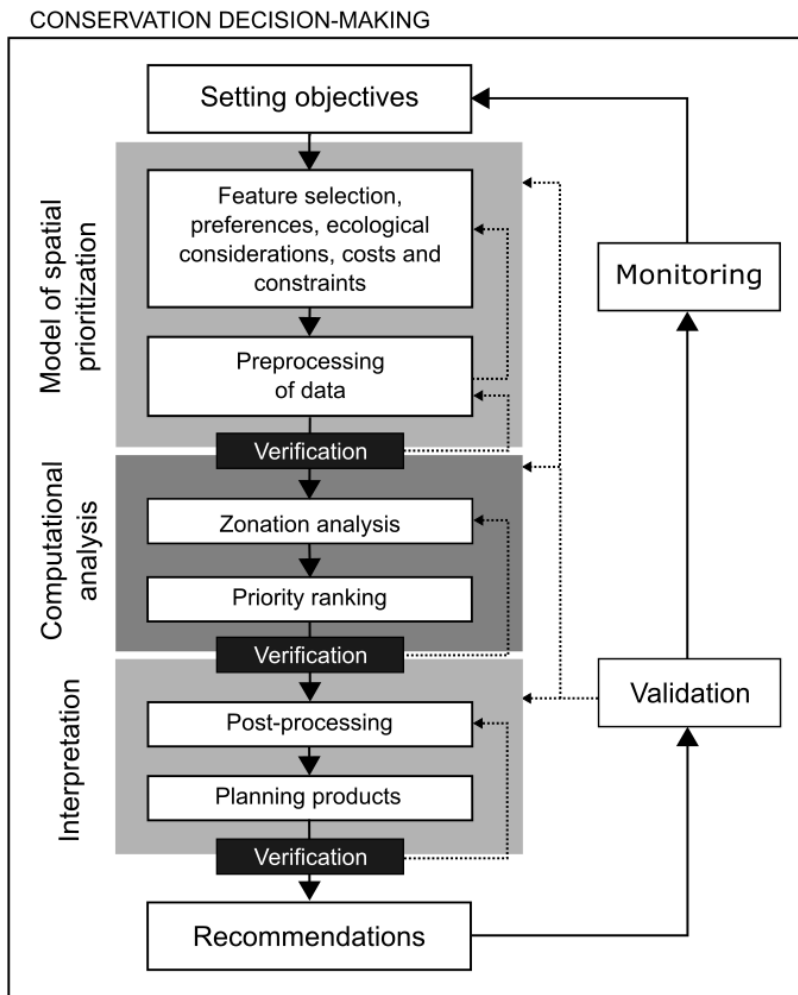
Figure 13. A schematic representation of the stages of a spatial conservation prioritization project. All stages of the process happen in the broader context of conservation decision-making, or, general land use planning when it involves ecological considerations. Dashed arrows mark the feedback paths in case verification or validation fail (see also 6.3). This figure was adapted from Lehtomäki and Moilanen (2013).

7.1.1 Example 1: Planning for green infrastructure in the Helsinki-Uusimaa region, Finland

Context and objective:

In Finland, 18 regional councils are responsible of the highest level of land use planning. The Helsinki-Uusimaa region, which includes the capital district of Finland, is one of the fastest growing metropolitan areas in Europe. It consists of 27 municipalities and is home for some 1.6 million inhabitants. The Regional Council of Helsinki-Uusimaa leads the ongoing process of the Regional Plan 4 for the region. The goal of the plan is to create preconditions for a competitive region that offers well-being for its inhabitants and does not exceed the limits of sustainable development. Planning for green infrastructure is one of the five topic areas of the plan. In this process, green infrastructure is viewed from the perspective of biodiversity, ecological networks, recreational networks, natural resources and ecosystem services.

To support knowledge-based planning of the green infrastructure, the regional council carried out a development project supported by the Ministry of Environment. The project supported the development of two approaches to support green infrastructure planning at the regional level: 1) Green Frame method (Kopperoinen et al., 2014) for evaluating ecosystem services and 2) application of Zonation in the context of land use zoning at the regional level.

> **The objective:** To create a spatial prioritization of high nature values sites with relatively low human disturbance.

The team:

The project required broad expertise and the commitment of various stakeholders. The broad objectives of the project were set in close collaboration between the planners of the regional council and a stakeholder group with representatives from municipalities and various NGOs. The Zonation team (coordinator at the regional council and zonation and scientific expertise at the University of Helsinki) interpreted the broad objective to formulate a Zonation project. An environmental expert team (approx. 25 participants) was called together to help define the details of the analysis project, including development of the model of spatial prioritization (data sets, their weights, connectivity settings, etc.) and validation of results.

Data:

The data used in the analysis were classified into four "baskets": 1) Habitats, 2) Species, 3) Land cover including noise and population, and 4) Protected areas and land use plans. Habitat and species layers were used as surrogates for biodiversity at large, whereas harmful land cover, noise and population were used to lower the value of areas. Protected areas and land use plans were used to restrict the Zonation prioritizations in hierarchical analyses. In the end some 70 informative data layers were used as input in the analysis. Many of these layers already were a surrogate for a broader set of biodiversity features. Original data was collected from several institutes and NGOs. It arrived in various source formats and was rasterized to 100 meter resolution grids in the national EUREF-FIN coordinate system.

Zonation setups:

Zonation setups were built gradually starting from local biodiversity value (habitats and species), then adding a weighting scheme, human impacts as a condition layer to reduce the local biodiversity value, connectivity settings to identify well connected areas, and finally introducing hierarchical masks.

ZONATION SETUPS

Input features   Tens of data layers of habitat distribution and condition and species distributions obtained from various sources.

Weights   Set together with the environmental expert group, first deciding the overall weights for data baskets for habitats and species, then for each layer with in these baskets (e.g. forest habitats or rural habitats) and then finally distributed to individual layers.

Connectivity transformed layers were included in analysis with half the weight of the actual data layers.

Condition   Intensive land use and noise were converted into conditions layers to model their detrimental influence on local biodiversity.

Connectivity   Distribution smoothing was carried out as a pre-processing step for some of the species data layers.

Matrix connectivity set between different habitat types, to identify well connected forest areas regardless of the detailed forest type (e.g. deciduous and coniferous forests were considered equally important to each other in connectivity calculations). The relevant connectivity scale was defined as 2 km based on expert judgement in the environmental expert group.

Negative interaction connectivity between harmful land use and biodiversity features (habitats and species) was used to lower the biodiversity value near harmful land use.

Cost   No cost information was used in this analysis.

Hierarchical mask   The performance of the current protected areas and protected area plans were tested by using them as hierarchical masks in the analysis.

Cell removal rule   The Additive Benefit Function (ABF) analysis variant was used, because the biodiversity layers used in analysis were thought to be surrogates of broader biodiversity.

Results and implementation:

The prioritization identified the most important areas for green infrastructure from the perspective of biodiversity and expansion of conservation areas. The results showed that while the current protected areas function fairly well as areas of high biodiversity value in the Helsinki-Uusimaa region, there is potential for a small but valuable expansion of the protected area network. Also, there is a high potential to maintain important habitats species if intensification of land use is directed to semi-degraded areas, which are comparatively plentiful in the area. The regional council used the final prioritizations to mark some 130 new areas as "high nature value sites" in the plan. The plan is still in the making, so final evaluation of the implementation of the results can be done only later.

Some observations:

Finland has fairly liberal data policy and advanced data infrastructure for environmental spatial data sets. Nevertheless, data negotiations and pre-processing of the data was very laborious and consumed most of the time resources in the Zonation project. Lack of data for example from the marine areas made it difficult to make a fully holistic analysis of the green infrastructure for the area. The result was nevertheless considered as a significantly useful input to the planning process. Introducing of ecosystem services into the analysis was tested at the end of the project. Based on preliminary analysis however, high biodiversity value and high ecosystem service values were only partially overlapping in the area, leaving clear scope for future investigation. The commitment of two broader groups was crucial for the success of this project. The stakeholder group that participated in setting the objectives was able to critically evaluate the outcome of the analysis. The expert group was crucial in iteratively developing the Zonation setups.

7.1.2 Example 2: Identifying native biodiversity priorities in the Waikato Region, New Zealand

Context and objective:

Primary responsibilities for biodiversity conservation in In New Zealand rest with the Department of Conservation, a central government agency set up under New Zealand's Conservation Act. However, under New Zealand's Resource Management Act (1991) some responsibilities for management of indigenous biodiversity are devolved to Regional Councils, of which there are sixteen nationally. A number of these councils have either already adopted, or are in the process of adopting, spatial conservation planning approaches to the identification of priority sites for management of indigenous biodiversity, reflecting both the limited funding available for such work, and a desire to manage resources in an efficient manner. Conservation goals for most councils are aligned around New Zealand's Biodiversity Strategy, in which the third of four high level goals, identifies the need to (i) maintain representative examples of a full range of native habitats and ecosystems in a healthy functioning state, and (ii) maintain and restore viable populations of all indigenous species (Anonymous 2000).

The objective: To identify a set of high priority sites that collectively represent a full range of the terrestrial and freshwater ecosystems occurring within the Waikato Region. This region is located in the northern half of the North Island, encompassing an area of approximately 25,000 square kilometres. Project goals were set by staff of the Waikato Regional Council, ensuring that they were consistent with the high level objectives specified in the Regional Policy Statement, a statutory document that sets out a vision for sustainable land management within the Region (http://www.waikatoregion.govt.nz/Council/Policy-and-plans/Policy-effectiveness-reports/Biodiversity-and-natural-heritage/).

The team:

The project required broad expertise comprising both professional staff from the Regional Council (ecologists, spatial analysts, land managers), and two independent consultants who contributed specific skills not available within the Council. Results were evaluated by a project team consisting of the consultant responsible for running Zonation analyses working together with Regional Council staff, with occasional input from other local experts and stakeholders.

Data:

Terrestrial data used in the analysis consisted of (i) a digital map of the potential natural ecosystem cover for the Region as mapped by an external consultant ecologist, and (ii) digital mapping of the current distribution of native-dominated ecosystems based on analysis of satellite imagery, supported by field validation by Council staff. Coverage was restricted to sites in which native species remain dominant, comprising a little under 9000 square kilometres, or 36% of the Region. Riverine data consisted of a river environments classification in which the selection, transformation and weighting of input variables was guided by a parallel set of biological data (Leathwick et al., 2011); lakes were divided into five morphogenetic classes with artificial reservoirs treated separately.

Separate condition layers were used for terrestrial, riverine and lacustrine ecosystems. The terrestrial condition layer was constructed by combining two components, the first describing the predicted condition in the absence of management, and the second describing any gains for biodiversity resulting from recent conservation management. The first of these was calculated as the product of three different factors: (i) the proportion of non-edge habitat in each contiguous ecosystem polygon calculated using a

50 m internal buffer, (ii) the average human population pressure as described by the most recent New Zealand census, and (iii) the spatial distribution of partial logging of native forests to remove merchantable timber species. The individual estimates were scaled so that the

Estimated gains from biodiversity management (mainly control of introduced vertebrate pests, both browser and predators) were used to modify the initial condition estimates, with the gains scaled so that the final combined condition estimates ranged from 0.02 to slightly less than one. Highest values (> 0.9) occurred in extensive mainland areas of native-dominated ecosystems subject to sustained control of introduced pest species, or on pest-free offshore islands. River and lake condition estimates were taken from the Freshwater Ecosystems of New Zealand database (http://www.doc.govt.nz/our-work/freshwater-ecosystems-of-new-zealand/). These estimates were compiled from descriptions of the catchment cover of native-dominated ecosystems, inputs from mines or industrial sites, nitrogen inputs (which are elevated in catchments subject to pastoral farming), the distributions of introduced pest fish, and/or the negative effects of dams on flow regimes and fish migration.

An underlying set of 21,435 watersheds for individual river segments was used as a planning unit layer, with upstream-downstream connectivity between planning units defined by the river network topology; a total of 1244 river catchments occur within the study area. Additional planning units were manually added around the coast to cover small areas with insufficient runoff to generate the creation of a river segment.

Zonation setups:

Setting up the analysis was complicated by the conflicting connectivity requirements for terrestrial and lake ecosystems versus riverine ecosystems. In particular, omni-

directional connectivity is appropriate for terrestrial ecosystems, allowing for recognition of linkages across catchment boundaries; by contrast, bi-directional linkages are appropriate for riverine ecosystems, reflecting the importance both of headwater environments on the ecological integrity of downstream river segments, and of coastal river segments environments in controlling the reproduction and survival of migratory fish species that spend most of their lives in rivers but reproduce in marine environments.

Terrestrial smoothing was specified using a connectivity matrix that was applied to all terrestrial and lake ecosystems. Matrix values were set to reflect the likely positive gains resulting from adjacency between different pairs of ecosystems, often in an asymmetric fashion, e.g., primary forests were assumed to have a strong positive influence on adjacent secondary vegetation, than vice versa, reflecting their value in providing propagules to foster natural succession back to forest.

Riverine smoothing was specified so that the value of all river segments showed

ZONATION SETUPS	
Input features	Data layers describing the distributions of 72 terrestrial ecosystem types, 6 lake types, and 36 river types.
Weights	Terrestrial weights set to reflect successional status and degree of habitat loss compared with estimated pre-human extent. Most ecosystems were given a weight of '1', but wetland classes and forest classes of warm, mild and cool climates, i.e., those suffering greatest loss from forest clearance and wetland drainage, were given a weight of 2; rivers were given a weight of 2, and lakes a weight of 1.
Condition	Separate condition layers were used for terrestrial, lacustrine and riverine ecosystems; these were based on relevant introduced pest and land use factors having a negative impact on biodiversity; condition estimates for terrestrial ecosystems were moderated by recent biodiversity management actions.
Connectivity	Terrestrial and lacustrine ecosystem smoothing carried out using matrix connectivity, using a kernel distance of 1km for forest ecosystems and 500 m for all other ecosystems. Riverine connectivity implemented using directed connectivity taking account of loss of both upstream and downstream planning units.
Cost	No cost layer was applied in this analysis.
Hierarchical mask	No hierarchical mask was used at this stage, although one could be employed in the future to estimate the representation delivered by areas currently receiving management.
Cell removal rule	Additive Benefit Function (ABF) was used throughout the setup to prioritize areas of overall highest value for biodiversity.

similar levels of decline in value as planning units in their headwaters were removed. By contrast, coastal river ecosystems were most heavily penalised for removal of planning units downstream of them, this penalty gradually reducing with progression from coastal to far inland river environments, and in which marine connectivity becomes progressively less important. Lake values were reduced in proportion to loss of their upstream catchments as for rivers.

Primary terrestrial ecosystem classes were given higher weights than the three secondary ecosystem classes; as a consequence, the three secondary classes had lower levels of representation in high-ranked sites, except where they were adjacent to primary terrestrial ecosystems or to lakes. Weights for rivers and lakes were set after iterative testing to assess their degree of representation relative to terrestrial ecosystems groups when using different weights.

Results and implementation:

The prioritization identified a subset of the region that maximises representation of a full range of biodiversity features, assuming that approximately 30% of native-dominated ecosystems within the region are likely to receive some form of management. Sites falling within the top 30% were delineated using a set of vector polygons, with buffers used to associate together small fragments lying within close proximity (<500m) either to each other or to larger areas of high value habitat. Accompanying tables describe the biodiversity character of each of the larger polygons (> 1000 ha) in extent, and the overall representation of different ecosystems. Conservation management required to maintain native biodiversity in good condition at each of these sites will be assessed in due course. However, as this project is still in progress, reporting fully on the implementation of results is not possible at this stage.

7.1.3 Example 3: Conservation prioritization for global protected area expansion under land use change and national decision making

Context and objective:

Protected areas are one of the main tools for halting the ongoing global biodiversity crisis caused by habitat loss, fragmentation, and other anthropogenic pressures. According to the Aichi Target 11 adopted by the Convention on Biological Diversity (CBD), the protected area (PA) network should be expanded to at least 17% of the terrestrial world by the year 2020. At the time of the study, 11% of terrestrial land was under protection.

The objective: To identify 1) where the priority areas for expanding the protected areas network from 11 to 17 % would be, and 2) how much the expansion of the network could improve the level of protection of species listed by the IUCN. Additionally, the analysis aimed to estimate 3) the impact of land use change and 4) the influence of global coordination in planning (national vs. global planning). The primary focus of the work was scientific.

The team:

The work required expertise from different fields and the research team consisted of academic researchers with different backgrounds ranging from conservation scientists (defining the objectives) to geographers (spatial data collection and manipulation) and to computer scientists (development of a new version of Zonation to handle extensive analyses).

Data:

Biodiversity data consisted of roughly 25 000 distribution maps of terrestrial vertebrates (IUCN, 2013) and 826 maps of ecoregions (Olson et al., 2001). Land use scenarios for the years 2000 and 2020 were transformed via expert knowledge into a numerical land

use naturalness layer (van Asselen and Verburg, 2012; used as condition layers in the analyses). The World Database on Protected Areas (IUCN and UNEP-WCMC, 2013) was needed for protected area delineations (hierarchical analysis) and Global Administrative Areas (Global Administrative Areas, 2012) was used to delimit spatial units for the "national level" prioritizations (administrative units analysis). The original polygon data was rasterized with a 1.7 km resolution in a geographical longitude / latitude coordinate system. For proper analysis, variation in the cell sizes was taken into account by weighting the cells with a cell size layer.

Zonation setups:

Several different Zonation setups were tested and run with several data resolutions for sensitivity analysis. The table below presents the Zonation setup used in the final analyses.

ZONATION SETUPS

Input features
25 000 layers of biodiversity features (IUCN Red List), land use maps for condition transformations (Pouzols et al., 2014), protected areas to create a hierarchical analyses (expansion starting from the current protected areas).

Weights
The weights were set according the IUCN classification of the species. The more threatened the species, the higher the weights.

Condition
Two condition layers were applied in two parallel analyses: land use and condition at present and projected for 2020.

Connectivity
No connectivity transformations were applied, due to the computational intensity of the analysis as it was.

Cost
No cost information was applied in this analysis.

Hierarchical mask
The performance of the current protected areas was tested by using them as hierarchical mask in the analysis

Administrative units
Countries were used as administrative units for analysis runs demonstrating the impact of national planning and decision-making

Cell removal rule
The Additive Benefit Function analysis variant (ABF) was used in the final analysis setup to prioritize areas of overall high value for biodiversity. Core Area Zonation (CAZ) was used in test phase.

Sensitivity analyses
Sensitivity was analyzed with respect to 1) original data resolution (by running the analysis with several input data resolutions), 2) data quality (by introducing random error to species distribution maps, 3) weights (by varying weights of the species groups).

Results and implementation:

Four analysis variants emerged as the main results: 1) Global prioritization for the present land use, 2) Global prioritization for the future land use, 3) National prioritization for the present land use and 4) National prioritization for the future land use. These analyses were contrasted to demonstrate the potential of the protected area expansion and the impact of land use change and national scale planning. The results were published in Nature (Pouzols et al., 2014) and as an interactive browsing service in http://avaa.tdata.fi/web/cbig/gpan. The resulting maps and statistics are also shared via that site.

Some observations:

Selection of the data and rasterizing the extensive amount of data layers was one of the most laborious tasks. The high resolution analysis at a global scale required a lot of computational time, so planning the runs well in advance was needed. Testing the setups and the runs with coarser scale data would have been useful. In the review process, very much time was used to discuss the "right" rasterizing scale of the IUCN polygons. At the end of the publication process, Nature required written use permissions for all the data sets that were available for non-commercial use only.

7.2 CREATING PLANNING PRODUCTS

Many of the standard outputs created by Zonation are directly useful for scientific communication or informing conservation decision-making and management. For planners and implementers who have not been part of the prioritization project, the outputs may seem rather technical and it may be difficult to understand their creation correctly. This is why sometimes much interpretation and redesign is needed to turn the outputs into more developed planning products.

Here, we use the term planning product to refer to a product that is based on Zonation results and is meant to address the objectives of the prioritization project from the end-users perspective. Good visualization of the Zonation results is usually a precursor to a planning product, but planning products are tailor-made to a particular purpose and integrate more additional information than simpler visualizations. This information can be actual spatial data or descriptive (textual) information documenting and explaining the results and their interpretation in more detail.

All the examples in the previous section include planning products of some type. In the case of planning for green infrastructure in the Helsinki-Uusimaa region (7.1.1), the Regional Council devised a classification scheme for high conservation value sites. Based on this classification, they created information site cards that provide much useful information on not only where the sites are, but also on what makes them valuable (see Figure 10). Similarly, the Regional Council of the Waikato region (7.1.2) was provided with planning products combining the spatial information of top-priorities with additional data tables describing the characteristics of each location in more detail. The example of conservation prioritization for the global protected area network expansion (7.1.3) was not done to directly support any particular planning process, but rather to highlight the importance of taking into account spatiotemporal land use change and global collaboration (or the lack of it). Nevertheless, in addition to the scientific publication reporting the findings, the authors also created an interactive website (http://avaa.tdata.fi/web/cbig/gpan) that makes it easy to examine the results from different perspectives. Importantly, the website is designed to highlight particular aspects of the results and their potential ramifications instead of providing as much details of the results as possible.

Starting to think about the planning products when the analyses are done can already be too late. Ideally planning products should be thought about throughout the whole project and the relevant information from all the stages should be incorporated into the products

As a final consideration, it is worth pointing out that even if the intended users are not likely to repeat the Zonation analyses in the future, it is still worth packaging Zonation setups and distributing them with the planning products. This way it is later always possible to go back to the original analyses if need be. For anyone familiar with Zonation, these files (preferably accompanied by the input data used) provide the necessary details on the analyses as far as the technical implementation goes. They do not tell, however, much anything about the objectives, the model of spatial prioritization and data pre- and post-processing. This is why ample documentation and textual descriptions in the planning products are needed.

7.3 CAPACITY BUILDING

Unless the Zonation project you are planning is guaranteed to be a one-off event, you might want to pay some attention to how the project could be repeated in the future. If you are working in an organization that is considering using Zonation in a particular process or even as part of developing new decision-support processes, you will certainly want to consider what it takes from an organizational perspective.

Chapter 1 has already given you an idea on what are the general considerations and the expected investments in terms of time and money for an organization. The points made in Chapter 1 are largely based on our own experience on working with various collaborators within and beyond academia. Especially in organizations working with operative environmental management, introducing Zonation into the toolset

used may seem like a mostly technical challenge. Having the technical capacity to run Zonation analyses is of course required, but as this document has hopefully demonstrated, many other skills are required and most of the time is spent in doing something else than running Zonation. This means that using Zonation, and doing spatial conservation prioritization in more general, requires building of capacity from data manipulation to communication of the results and everything else between.

Team work is also about having partly overlapping capacities. At best, everyone from data manager to the ones writing the final report would understand the basics of Zonation analysis. At best, the planning and implementation of a prioritization project is a group effort where everyone is to some extent involved in each step. In an ideal situation, also those dealing with implementation are actively engaged in the prioritization project and can develop a good understanding about the analyses while they are being developed and executed. Co-learning has already been mentioned in this document several times and here it once again proves to be one of the crucial outcomes of a spatial prioritization project.

References

Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E., Moilanen, A., 2012. Effects of connectivity and spatial resolution of analyses on conservation prioritization across large extents. Conserv. Biol. 26, 294–304. doi:10.1111/j.1523-1739.2011.01814.x

Burgman, M.A., Carr, A., Godden, L., Gregory, R., McBride, M., Flander, L., Maguire, L., 2011. Redefining expertise and improving ecological judgment. Conserv. Lett. 4, 81–87. doi:110.1111/j.1755-263X.2011.00165.x

Carroll, C., Dunk, J.R., Moilanen, A., 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. Glob. Chang. Biol. 16, 891–904. doi:10.1111/j.1365-2486.2009.01965.x

Di Minin, E., Veach, V., Lehtomäki, J., Pouzols, F.M., Moilanen, A., 2014. A Quick Introduction to Zonation. Conservation Biology Informatics Group, Helsinki.

Easterbrook, S.M., 2010. The difference between Verification and Validation [WWW Document]. URL http://www.easterbrook.ca/steve/2010/11/the-difference-between-verification-and-validation/ (accessed 11.4.15).

Ferrier, S., Wintle, B.A., 2009. Quantitative approaches to spatial conservation prioritization: matching the solution to the need, in: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Spatial Conservation Prioritization: Quantitative Methods & Computational Tools. Oxford University Press, Oxford, p. 304.

GDAL Development Team, 2014. GDAL - Geospatial Data Abstraction Library, version 1.10.1. Open Source Geospatial Foundation.

Global Administrative Areas, 2012. GADM database of Global Administrative Areas [WWW Document]. URL http://www.gadm.org/

Gregory, R., Failing, L., Ohlson, D., Mcdaniels, T.L., 2006. Some Pitfalls of an Overemphasis on Science in Environmental Risk Management Decisions. J. Risk Res. 9, 717–735. doi:10.1080/13669870600799895

IUCN, 2013. The IUCN Red List of Threatened Species [WWW Document]. URL http://www.iucnredlist.org

IUCN, UNEP-WCMC, 2013. The World Database on Protected Areas (WDPA) [WWW Document]. URL www.protectedplanet.org

Kareksela, S., Moilanen, A., Tuominen, S., Kotiaho, J.S., 2013. Use of Inverse Spatial Conservation Prioritization to Avoid Biological Diversity Loss Outside Protected Areas. Conserv. Biol. 27, 1294–1303. doi:10.1111/cobi.12146

Kopperoinen, L., Itkonen, P., Niemelä, J., 2014. Using expert knowledge in combining green infrastructure and ecosystem services in land use planning: an insight into a new place-based methodology. Landsc. Ecol. 45, 823–841. doi:10.1007/s10980-014-0014-2

Kuusterä, J., Aalto, S., Moilanen, A., Toivonen, T., Lehtomäki, J., 2015. Uudenmaan viherrakenteen analysointi Zonation-menetelmällä. Helsinki.

Leathwick, J.R., Snelder, T., Chadderton, W.L., Elith, J., Julian, K., Ferrier, S., 2011. Use of generalised dissimilarity modelling to improve the biological discrimination of river and stream classifications. Freshw. Biol. 56, 21–38. doi:10.1111/j.1365-2427.2010.02414.x

Lehtomäki, J., 2014. zonator: Utilities for Zonation spatial conservation prioritization software. R package version 0.4.0. Helsinki.

Lehtomäki, J., Moilanen, A., 2013. Methods and workflow for spatial conservation prioritization using Zonation. Environ. Model. Softw. 47, 128–137. doi:10.1016/j.envsoft.2013.05.001

Lehtomäki, J., Tomppo, E., Kuokkanen, P., Hanski, I., Moilanen, A., 2009. Applying spatial conservation prioritization software and high-resolution GIS data to a national-scale study in forest conservation. For. Ecol. Manage. 258, 2439–2449. doi:10.1016/j.foreco.2009.08.026

Lehtomäki, J., Tuominen, S., Toivonen, T., Leinonen, A., 2015. What Data to Use for Forest Conservation Planning? A Comparison of Coarse Open and Detailed Proprietary Forest Inventory

Data in Finland. PLoS One 10, e0135926. doi:10.1371/journal.pone.0135926

Martin, T.G., Burgman, M.A., Fidler, F., Kuhnert, P.M., Low-Choy, S., Mcbride, M., Mengersen, K., 2012. Eliciting Expert Knowledge in Conservation Science. Conserv. Biol. 26, 29–38. doi:10.1111/ j.1523-1739.2011.01806.x

Mas, J.-F., Pérez-Vega, A., Clarke, K.C., 2012. Assessing simulated land use/cover maps using similarity and fragmentation indices. Ecol. Complex. 11, 38–45. doi:10.1016/j. ecocom.2012.01.004

Mikkonen, N., Moilanen, A., 2013. Identification of top priority areas and management landscapes from a national Natura 2000 network. Environ. Sci. Policy 27, 11–20. doi:10.1016/j.envsci.2012.10.022

Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J., Thomas, C.D., 2011. Balancing alternative land uses in conservation prioritization. Ecol. Appl. 21, 1419–1426. doi:10.1890/10-1865.1

Moilanen, A., Arponen, A., 2011. Setting conservation targets under budgetary constraints. Biol. Conserv. 144, 650–653. doi:10.1016/j.biocon.2010.09.006

Moilanen, A., Pouzols, F.M., Meller, L., Veach, V., Arponen, A., Leppänen, J., Kujala, H., 2014. Zonation spatial conservation planning methods and software v. 4, user manual. Helsinki.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J. a., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., George, V.N., Jennifer, A.D., Wesley, W., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. Bioscience 5, 933. doi:10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Pierce, S.M., Cowling, R.M., Knight, A.T., Lombard, A.T., Rouget, M., Wolf, T., 2005. Systematic conservation planning products for land-use planning: Interpretation for implementation. Biol. Conserv. 125, 441–458. doi:10.1016/j.biocon.2005.04.019

Pouzols, F.M., Toivonen, T., Di Minin, E., Kukkala, A.S., Kullberg, P., Kuusterä, J., Lehtomäki, J., Tenkanen, H., Verburg, P.H., Moilanen, A., 2014. Global protected area expansion is compromised by projected land-use and parochialism. Nature 516, 383–386. doi:10.1038/nature14032

Python Development Team, 2014. Python Language Reference, version 2.7. Python Software Foundation.

van Asselen, S., Verburg, P.H., 2012. A Land System representation for global assessments and land-use modeling. Glob. Chang. Biol. 18, 3125–3148. doi:10.1111/ j.1365-2486.2012.02759.x

van den Hove, S., 2007. A Rationale for Science-Policy Interfaces. Futures 39, 1–19. doi:10.1016/j.futures.2006.12.004

Warman, L.D., Sinclair, A.R.E., Scudder, G.G.E., Klinkenberg, B., Pressey, R.L., 2004. Sensitivity of Systematic Reserve Selection to Decisions about Scale, Biological Data, and Targets: Case Study from Southern British Columbia. Conserv. Biol. 18, 655–666. doi:10.1111/j.1523-1739.2004.00538.x

Zonation is decision support software for land use planning, including uses such as traditional design of conservation area networks or spatial impact avoidance. It is capable of data rich, large scale, high-resolution spatial prioritization. Whether using Zonation for scientific research or in real-life planning, running a successful project involves several project stages that often proceed in a somewhat iterative fashion. This document provides an overview of what those stages are, and what types of issues should be considered when planning to use Zonation. This information is intended for any individual or organization that is considering making use of Zonation.

The topics addressed here are those encountered before and after the Zonation prioritization analysis itself:

- Budgeting for time and resources
- Setting objectives and planning how to meet them
- Building a model for spatial prioritization
- Data requirements and pre-processing
- Setting up and organizing Zonation input files
- Visualizing and interpreting the results
- Creating planning products

Zonation projects are also demonstrated through a set of case-studies that range from national to global scales.