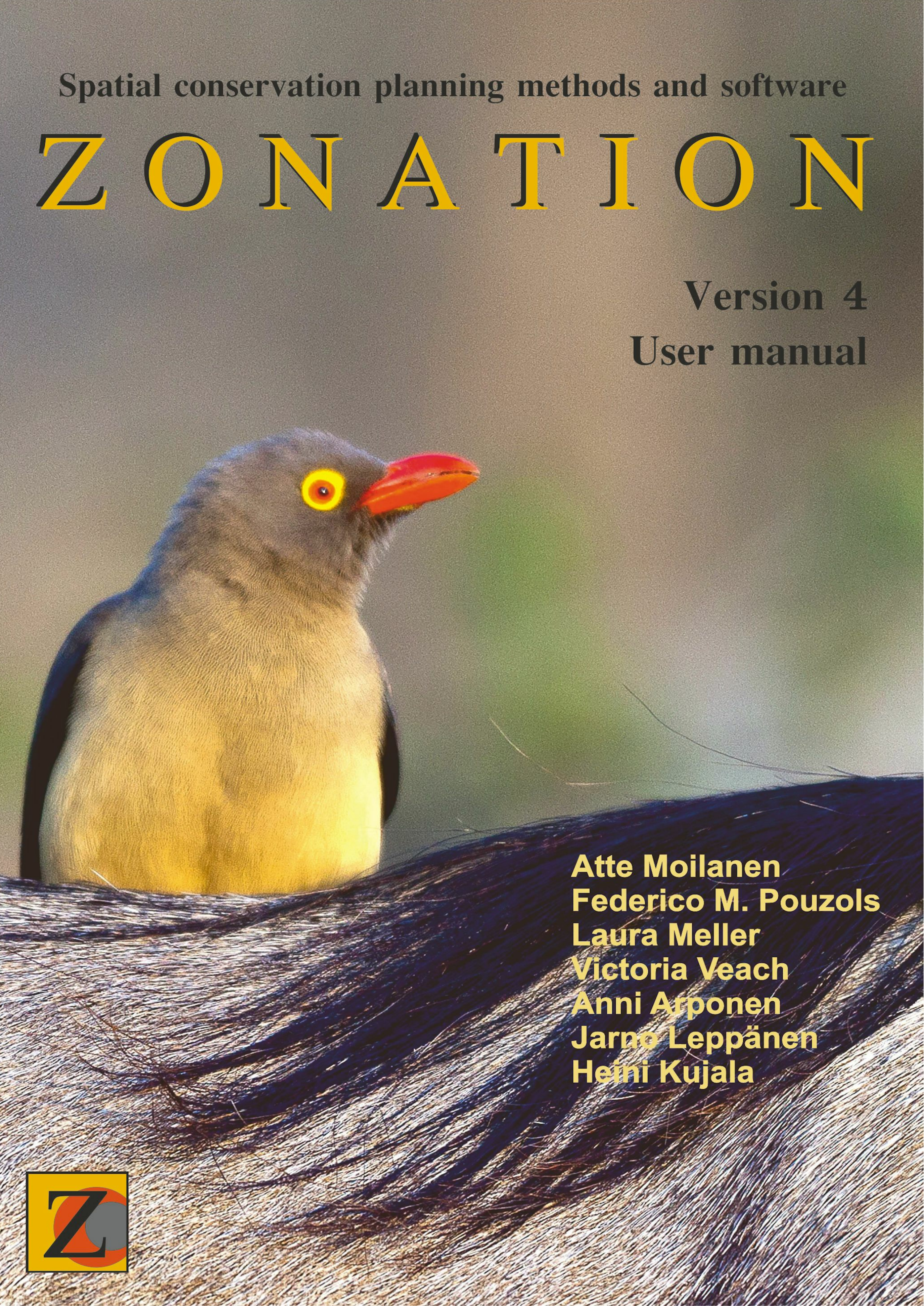
Spatial conservation planning methods and software

ZONATION

Version 4

User manual

Atte Moilanen
Federico M. Pouzols
Laura Meller
Victoria Veach
Anni Arponen
Jarno Leppänen
Heini Kujala



ZONATION
Conservation planning software

Zonation is methods and software for spatial conservation prioritization; it is a decision support tool for conservation resource allocation. Zonation identifies areas important for retaining habitat quality and connectivity for multiple species, habitats or ecosystems, indirectly aiming at long-term persistence of biodiversity. When including costs, Zonation produces cost-efficient solutions.

With suitable inputs, Zonation can (among other things) be used for answering questions about reserve design, reserve network expansion, impact avoidance, biodiversity offsetting and spatial targeting of habitat management or restoration. As Zonation operates on large grids, it provides a direct link between GIS, statistical distribution modeling and spatial conservation prioritization.

Keep an eye on our web site for new Zonation versions and other software!

http://cbig.it.helsinki.fi

ISBN 978-952-10-9919-9 (paperback)
ISBN 978-952-10-9920-5 (PDF)

Zonation - Spatial conservation planning methods and software. Version 4. User Manual

Do not use this software if you disagree with the disclaimer or conditions of use. Even though the Zonation software has been done with the best of intentions, it is quite beyond one small research group to ensure its correct operation under all operating systems and environments. Anticipating all potential combinations of erroneous input has not been possible. Therefore, use the software with care and make an effort to understand how the inputs connect to outputs.

Publisher

C-BIG Conservation Biology
Informatics Group
Department of Biosciences
University of Helsinki, Finland

Credits

Manual

Zonation project leader
Atte Moilanen

Technical editors
Victoria Veach (v4)
Laura Meller (v3)
Anni Arponen (v3)
Heini Kujala (v1 & 2)

Technical content
Atte Moilanen
Federico Montesino Pouzols
Victoria Veach

Cover Design
Aija Kukkala, Atte Moilanen
& the C-BIG

Cover Photos
Enrico Di Minin

Zonation v.4 software

Numerical core and new GUI
Federico Montesino Pouzols
Atte Moilanen

Thank you

We thank the numerous collaborators and Zonation users who have provided comments about various versions of the software and its documentation!

Special thanks to Brendan Wintle who generously provided sample data files from the Hunter Valley region to be included with the tutorial, and to Evgeniy Meyke who has kindly provided beautiful background photographs for the Zonation documentation.

Thanks again to all who have collaborated in the development of Zonation methods and applications!

We appreciate the support of our funders:

The Zonation project has been 2004-2012 supported by an Academy Research Fellowship to A.M. (2003-2008; grant 1206883); the Ministry of Environment via the Natural Heritage Services of Metsähallitus (2010 onwards); the EU 7th framework project SCALES (2009-2014; grant 226852); the EU/ERC project GEDA (2011-2015; grant 260393); and an Academy of Finland Centre of Excellence grant (2006-2011; grants 213457 and 129636) to the Finnish Centre of Excellence in Metapopulation Biology directed by Academy Professor Ilkka Hanski.

Helsinki, April 17, 2014, Professor Atte Moilanen,
Dept. Biosciences P.O. Box 65, FI-00014
University of Helsinki, Finland

Table of Contents

Part I Introduction

Part II Methods & algorithms

Part III ZIG - The Zonation software

Part IV Zonation Graphical User Interface

Part V Zonation analysis setups for common planning needs

Part VI Tutorial & Examples

Index

Part

I

1 Introduction

1.1 Aim & purpose

Zonation is cutting edge software for spatial conservation prioritization and broad-scale spatial conservation planning. It identifies areas or landscapes that are important for retaining habitat quality and connectivity simultaneously for multiple biodiversity features (e.g. species, land cover types, ecosystem services, etc.), thus providing a quantitative method for enhancing the persistence of biodiversity in the long term. Zonation can also account for costs, thereby allowing cost effective allocation of resources. Zonation can be used for traditional reserve or site selection, but it offers many other analyses for conservation prioritization. Typical analyses allowed by Zonation include:

- Identification of near-optimal connected reserve networks
- Expansion of existing reserve networks, effectively by simultaneous gap-analysis and filling
- Evaluation of existing or proposed reserve networks
- Identification of ecologically low-value areas for economic use (impact avoidance and biodiversity offsetting)
- Prioritization which can be used for many purposes including targeting of management, restoration or incentive funds

Zonation produces a hierarchical prioritization of the landscape that is balanced across many factors, meaning in the lingo of spatial prioritization that it accounts for complementarity. At the high level, the algorithm proceeds by iteratively removing the least valuable cells from the landscape while minimizing marginal loss of conservation value and accounting for connectivity needs and priorities given for biodiversity features. As a result, a nested sequence of highly connected landscape structures is obtained with the core areas of species distributions remaining last and previously removed areas shown as buffer zones. In this way, landscapes can be zoned according to their potential for conservation, and differing degrees of protection, maintenance, or restoration effort can be applied to different zones.

The Zonation software is intended for the analysis of biological data with the aim of aiding conservation decision making. In the broader sense, the output of Zonation should be seen as an analysis of conservation priority which feeds into a broader land use planning framework where political decisions are made. The purpose is not always to produce a detailed conservation plan for a large region, but it could be to identify priority areas of the landscape that could be subjected for more detailed analysis and planning together with stakeholders. Note that Zonation can also be used to identify the areas of the landscape where human activity would cause the least harm to biodiversity.

The Zonation software has been geared towards using large grids as input data, facilitating a direct link between GIS, statistical distribution modeling, and Zonation. It is particularly simple to input modeled species distributions, community turnover, or land cover types in Zonation. The Zonation software can process relatively large datasets on an ordinary desktop PC. Zonation v.4 is 64bit software and allows very large analyses on off-the-shelf

PCs that have large memory capacity.

1.2 The Zonation framework in a nutshell

Aim and purpose

- To provide a tool for broad-scale, high-resolution spatial conservation planning using primarily GIS raster data

Analyses

- Identification of optimal conservation areas
- Identification of least useful conservation areas for impact avoidance or offsetting
- Balancing of biodiversity conservation with alternative (human) land uses (New since 3.0)
- Replacement cost analysis for current or proposed reserves
- Combined community-level and species-level prioritization (New since 3.0)
- Planning modes that define alternative perceptions of conservation value:
    - Core-area Zonation
    - Additive benefit function
    - Target-based benefit function
    - Versatile generalized benefit function
    - Random ranking

(Biodiversity) features useable in Zonation include

- Species
- Ecosystems and similar high-level ecological entities (habitat types, vegetation classes, communities, etc.)
- Ecosystem services
- Distributions of alleles
- Any other data that describes the distribution of a feature on a grid. Features can represent other entities than elements of biodiversity.

Data

- Data types:
    - Presence/absence data
    - Probabilities of occurrence
    - Abundance/density data
    - Any other (non-negative) numbers that represent occurrence levels of features
- Broad-scale grids with
    - Uncertainty of occurrence data
    - Mask layers specifying prior land uses, such as the present protected area network
    - Grids representing costs and opportunity costs
    - Landscape condition (New since 3.0)
    - Retention of biodiversity (New since 3.0)
- Point observation data
- Planning unit layers
- Administrative units (New since 3.0)

- Community similarity or connectivity matrices (New since 3.0)

Analysis features

- Species priorities via weighting
- Methods for managing connectivity needs of species
    - Distribution Smoothing (feature-specific)
    - Boundary Quality Penalty (feature-specific)
    - Boundary Length Penalty
    - Directed Freshwater Connectivity (feature-specific)
    - Interaction Connectivity (between two distributions of features)
    - Matrix Connectivity (between multiple features; New since 3.0)
    - Edge Effect Correction (New since 3.0)
    - Corridor connectivity (New since 4.0)
- Uncertainty analysis aiming at robust conservation decisions
- Prioritization over multiple administrative regions (New since 3.0)
- Automated post-processing analyses
- Utilizing increased memory capacity for versatile analyses (New since 3.0, major developments for 4.0)

1.3 Zonation compared to other reserve selection approaches

In this section we comment on the differences between Zonation and other commonly used approaches to reserve selection. This comparison is not meant to be exhaustive nor completely referenced, but rather to give an indication of the most fundamental differences (that we believe exist) between these methods. This section will naturally become outdated when new features are developed for Zonation and other conservation planning software packages.

1.3.1 Zonation

Input data. Zonation is targeted for use with large grid-based data sets. This implies that species distributions used within Zonation might be produced using some predictive statistical technique using environmental layers as predictors for species presence/ abundance. Data sets in the order of tens or hundreds of millions of grid cells can be analyzed. Observational data can, of course, be used as well.

Output. Instead of outputting the optimal set of sites for achieving targets Zonation produces (1) the hierarchy of cell removal throughout the landscape and (2) feature-specific representation loss curves. This kind of output has multiple advantages, including:

i. Indication of the importance of all cells, both inside and outside any given top fraction
ii. Curves show how well (relatively) individual species are represented at any given fraction of the landscape
iii. Curves indicate the relative value of the solution as well as the stability of the solution
iv. Easy and intuitive visualization

To elaborate on item (iii): If biodiversity feature representation levels are declining rapidly at the chosen landscape fraction, it means that the solution is not stable with respect to uncertainty in input data and small changes in the selected fraction and/or spatial pattern might have large consequences for biodiversity. If the biodiversity feature performances are stable at the chosen fraction, small changes in the fraction/spatial pattern are unlikely to have any significant effect on the solution quality.

Optimality. The optimality characteristics of Zonation have not been conclusively examined, but the present evaluation of this issue indicates that Zonation using additive benefit functions or the targeting benefit function (above the target) is very close to globally optimal. This is because the optimization problem is convex with these cell removal rules and can thus be solved using a gradient-like iterative heuristic (van Teeffelen and Moilanen 2008). Also, with the additive cell removal rules, the degree of sub-optimality goes down when the landscape size (number of cells) increases. Thus optimality is not a problem with the additive cell removal rules. The Zonation algorithm is deterministic, as opposed to stochastic, implying that it finds the same result in different runs.

An exception is the use of the Boundary Quality Penalty (BQP), which renders the problem non-convex (especially if some species benefit from fragmentation), and the degree of sub-optimality of the solution is unknown.

The core-area Zonation (CAZ) method has so far only been defined algorithmically, but not in an objective function form. The CAZ cell removal rule specifies a difference equation for conservation value but not the objective directly, and the degree of sub-optimality of results is unknown. However, no other implementation of this method is available.

Limitations. Zonation is primarily intended for binary select or not, protect or not, and restore or not problems. Unlike MARXAN with Zones, Zonation is directly not meant for the direct targeting of multiple alternative conservation actions.

1.3.2 Integer programming

Input data. Integer programming (IP) can accept arbitrary sites as well as grid cells as data. According to Williams *et al*. (2004), the data size limits of integer programming (IP) in 2004 were around 10 000 landscape elements, which corresponds to a grid size of about 100x100 elements. While larger data sets can now be processed using IP, Zonation v.4 can process grids of approximately >200 million elements, which is still likely to be too large for IP. Additionally, some Zonation problem variants are nonlinear (BQP; NQP), and such analyses cannot be approached using IP at all.

Output. Integer programming produces a globally optimal set of sites achieving targets. There is no prioritization through the landscape and no performance curves are produced.

Optimality. Integer programming guarantees a globally optimal solution to a simplified problem. The value of the global optimality of results is compromised by the requirement that both the objective function and constraints need to be linear (or that they can be linearized). In a sense you have the optimal solution to the wrong (simplified) problem. Integer programming is not applicable, at least not easily, to species-specific connectivity calculations on large landscapes (see also Moilanen 2008).

Literature

Moilanen, A. 2008. Two paths to a suboptimal solution - once more about optimality in reserve selection. *Biological Conservation,* 141: 1919-1923.

Williams, J., C.S. ReVelle, and Levin, S.A. 2004. Using mathematical optimization models to design nature reserves. *Frontiers in Ecology and the Environment,* 2: 98–105.

1.3.3 Stochastic global search

Stochastic global search includes techniques such as simulated annealing (SA; as in MARXAN) and genetic algorithms (GA).

Input data. In principle, stochastic global search can be run on extremely large problems with few constraints on the complexity of the problem. SA can handle larger problems than a GA because of the memory requirement for storing the GA population.

Output. A solution of typically unknown quality. In some cases it may be possible to devise an analytical method that provides bounds on solution quality (as in Moilanen 2005), which then changes the method from heuristic to approximation. (Heuristic: a method for which the quality of results is unknown; Approximation: a method for which the maximum degree of sub-optimality of the results has been quantified in a non-trivial manner.)

Optimality. The degree of sub-optimality will be highly dependent on (1) the size of the data, (2) the complexity of the problem (for example, the presence of nonlinear connectivity components), and (3) the details of the implementation of the optimization algorithm. SA and GA are not standard algorithms (except for the high-level meta-algorithm). They can be varied in numerous ways, particularly, in terms of how they generate the new solutions to evaluate. If the search starts far from the good regions of the search space, it actually is not guaranteed that the good regions are found at all. Good convergence with large problems is absolutely not guaranteed. Multiple runs from different starting points are required to test for indications of convergence, and if multiple runs reliably converge to a very similar result, then this indeed is an indication that the solution is likely quite acceptable in terms of optimality. The performance of these methods is probably rather good with data sets with thousands or tens of thousands of sites, but at the million-element scale the performance of these methods is poorly known. Relative performance probably degrades when problem size increases, which is opposite from what is actually expected for Zonation, at least with the additive cell removal rules.

Optimization is a large field of science, and there is a lot of literature about this topic. See the references below for examples of the use of stochastic optimization on nonlinear reserve selection problems. Also check the MARXAN reserve selection software user manual and references therein.

Literature

Moilanen, A. 2005. Reserve selection using nonlinear species distribution models. *American Naturalist,* 165: 695-706. AND in particular its electronic appendixes A-C.

Moilanen, A. and M. Cabeza. 2002. Single-species dynamic site selection. *Ecological Applications,* 12: 913-926.

Moilanen, A. and I. Ball. 2009. Heuristic and approximate optimization methods for spatial conservation prioritization. Pp. 58-69 in Moilanen, Wilson and Possingham (Eds.), Spatial Conservation Prioritization*,* Oxford University Press.

1.4 A typical Zonation work flow

There are two parts to the Zonation work flow. First, the aims of the analysis need to be specified and the data needed to answer the aims must be obtained. Second, a Zonation analysis that corresponds to the stated aims must be designed and executed. Here we present a short discussion about specifying aims and data, followed by a more detailed instruction of executing a Zonation analysis.

1.4.1 Specification of aims, ecological model and data

The open online publication Lehtomäki and Moilanern (2013) provides a summary of executing a Zonation project. Zonation could also be applied inside the broader operational model of systematic conservation planning (see Margules and Sarkar, 2007; Pressey and Botrill 2009).

The aims of any analysis should be clearly stated before beginning computational spatial conservation prioritization. Only after the aims are specified, data and a Zonation analysis that supports the purpose can be developed. Specifying the aims and developing the data sounds simple, and at the conceptual level it may indeed be. In reality, specification of aims and obtaining relevant data can be 90% of the work, with the computational analysis itself taking only 10% of the effort. The following steps may be relevant to help with this stage.

1. **Determine the broad aims of the project**. Is the question about the design of a new reserve network, about the expansion of an existing network, or about identification of areas where economic activity would harm biodiversity first? What is the target of analysis; what species, habitats, or ecosystem services will be included in the analysis? Should costs be accounted for?

2. **Develop what we call 'the ecological model'**. What ecological information describes biodiversity and conservation value in a manner that is relevant for the planning need? What data is available to answer the planning question? What species distribution information is available? What habitat and ecosystem distribution data is available? Is there information about habitat condition? What, if anything, is known about present and future human impacts and threats? What are the costs and opportunity costs incurred by

stakeholders? Also consider how features will be weighted and what connectivity responses will be used. Such questions are relevant when the ecological model is developed.

3. **Develop data to answer the needs of the ecological model**. Obtain relevant data layers and format them so that they work in a Zonation analysis.

Note that development of the aims, ecological model, and data is best done among a planning group consisting of knowledgeable experts representing different stakeholders. It is rare that one party alone has all the knowledge and data relevant for analysis, and multiple parties often need to participate in the planning.

Conceptually, the development of the aims, ecological model, and data are separate steps, but in practice these stages are intertwined and need to be developed together. Perfect data will never be available. Therefore, the planning process is more about identifying data that answers the broad aims of the study as well as possible rather than answering a conceptual goal with a perfect ecological model and data. Consequently, analysis aims and the details of the ecological model may need to be modified according to what data becomes available. The amount of effort and workforce available for data collation can also influence what kind of analysis can be implemented.

Limited data results in an analysis with limited value. The better the ecological model and data, the more weight should be given to the Zonation analysis in ultimate decision making. As an easy example, say that a reserve network is desired that protects all biodiversity of the region. Then assume that excellent data is available for birds, butterflies, mammals, and vascular plants. The analysis is fully relevant for these feature groups only. Coverage for all other major taxa, such as insects, reptiles, lichens, etc. would need to be accounted for via expert knowledge. Note that surrogacy relationships could be relevant here, but, unfortunately, the general findings in scientific literature suggest that surrogacy relationships are unreliable between most pairs of taxa. For example, one could not automatically assume that birds would act as a reliable surrogate for plants.

Literature

Lehtomäki, J. and A. Moilanen. 2013. Methods and workflow for spatial conservation prioritization using Zonation. *Environmental Modelling & Software*, 47: 128-137.

Margules, C. R. & Sarkar, S. 2007. Systematic Conservation Planning. First Edition. Cambridge University Press, Cambridge.

Pressey, R. L. & Bottrill, M. C. 2009. Approaches to landscape- and seascapescale conservation planning: convergence, contrasts and challenges. *Oryx,* 43, 464–475.

1.4.2 Getting a Zonation analysis running

This section outlines a typical sequence of steps that would be done for the Zonation analysis of one data set. Section 5 of this manual includes information relevant for common planning setups.

(i) Get the simple analyses running

1. Install Zonation. Make sure the program is working with the example files provided.

2. Decide the concept of conservation value, implemented via the cell removal rule.

3. Produce a new settings file, biodiversity feature list file, etc. for the data and check that the basic analysis (without aggregation methods or uncertainty analysis) runs.

4. Try variants of the basic analysis by adding unequal biodiversity feature weights, aggregation methods, uncertainty analysis, and some other potentially relevant analysis features. Solution comparison can be used to check the difference that the addition of each new feature can cause in the solution. These preliminary analyses can be run using high warp factors (100-1000) to reduce run times.

(ii) Identify a realistic base-analysis. Zonation v.4 provides a large number of options about which considerations to include in an analysis. All of them have some specific data requirements, and applying them all is hardly a meaningful thing to do. The choice of analysis features will depend on the specific planning context, objectives, and data availability. Further, there are endless options of how to weight species or other biodiversity features, which aggregation method to use with what exact parameter values, and defining parameter values for other analysis features. It is not possible to run all combinations of everything and it is not useful to do so. Therefore, after getting the basic Zonation running, decide the most reasonable options for the analysis. Things that need to be decided include:

1. Determine the most appropriate analysis setup to match the aims. Analysis setups for different planning problems are outlined in section 5.

2. Determine what would be the best combination of analysis components. Should cost, uncertainty, interactions, landscape condition and retention, alternative land uses, landscape dynamics, etc. be considered?

3. Decide how to account for the connectivity needs of biodiversity features (species) in the analysis. In general, aggregation should be included, especially if the raster cell sizes are small, such as a few hectares. This is because population dynamics of nearby cells are strongly linked with small selection units. If cell sizes are very large (e.g., 10x10 km cells) then aggregation could plausibly be omitted under the assumption that a cell is big enough to include viable populations.

4. Determine weights for species or other biodiversity features. Equal weights is the default option but there may be good reason to favor particular features by giving them more weight. See Lehtomäki and Moilanen (2013) about the setting of weights in Zonation.

(iii) Sensitivity analysis around the base analysis. At this point the most appropriate analysis options have been identified. The next steps are:

1. Run the base-analysis, preferably using a relatively low warp factor.

2. Run variants around the base-analysis varying a single analysis feature at a time (it is probably not possible to run all combinations of everything). This is essentially a sensitivity analysis, which is done by varying weights, aggregation, and uncertainty analysis settings within reasonable bounds. An advisable strategy for complex analysis setups is to start from a simple setup and adding one new component at a time, continually checking the impact of that component. The addition of each data component and major analysis feature should result in a change that makes sense. If the change does not make sense, or no change at all is observed, it is possible that there is some error in the data or that the analysis feature has not been correctly switched on. For example, switching connectivity on should result in observable aggregation of priorities in the rank map. Use solution comparison to see how big a difference various options make.

3. An analysis of selection frequency may provide useful summary information over analyses. Selection frequency analysis can be done using the merged map feature of the Zonation v.4 GUI.

At this point it should be clear how different planning options influence the analysis and solutions.

(iv) Interpretation and post-processing of output. Identify priority areas for conservation (or the least important areas) from the solution. From this, management landscapes can be identified and their statistics can be checked to find out why different areas are important. What are the biodiversity features that occur there?

1.5 Software installation and quick start

Installation

The installation packages for windows include:

- The Zonation program (**zig4.exe**)
- The user interface (**zig4gui.exe**)
- The **zig4load.exe** and **zig4run.exe** utilities
- A user manual (**.pdf**) and tutorial files, also in html format browsable from the graphical interface.

The GNU/Linux packages include the same binaries. Zonation.compressed packages with these executable files and/or installer packages for 32 and 64 bits windows systems can be downloaded from the C-BIG website (cbig.it.helsinki.fi). (Note that with subsequent Zonation version the names of executables may become updated.)

Quick start

The following instructions serve as a guide to run the basic Zonation for those who have already familiarized themselves, at least to some extent, with the program. For more detailed instructions and additional analyses please see sections 3.2. Running Zonation, 3.3. Input files & settings, 3.5. Post-processing analyses & options, 4. Zonation Graphical User Interface, and 5. Analysis setups for common planning needs.

1. To run the program at least the following input files are needed:
    - Biodiversity feature map files. These basic raster files (e.g., .asc or .img or .tif) are exported from GIS programs. These files define biodiversity feature (e.g., species, habitats, communities, ecosystems, etc.) distributions in the landscape. The program can incorporate any kind of species distribution data, such as presence-absence, probabilistic or abundance data, or species-specific population connectivity surfaces, etc., as long as data for all species is in the same format. It can also include files describing the suitability of the landscape for alternative uses. Zonation can also use point distribution data.
    - Biodiversity feature list file (.spp). The names of all the biodiversity feature files must be listed in a separate file with each feature on a separate row with the feature-specific parameters before the file name (see section 3.3.2.2 for more detailed descriptions). The biodiversity feature list file tells the program which distribution files will be used in the analysis.
    - Run settings file (.dat). Defines the settings in the analysis.
    - A batch file (.bat). Contains the command line call to Zonation.

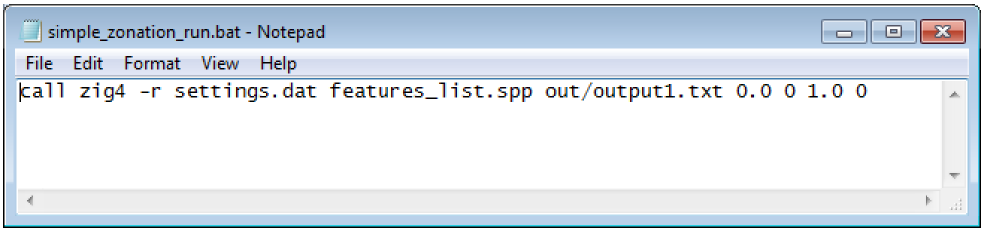
Please use decimal points, NOT commas, in all input files! Zonation has originally been developed under environments that assume the use of decimal dots. To avoid surprises, it may be helpful to change the computer settings to a country which assumes decimal dots. The project maker facility of the GUI can help in setting up a simple zonation run.

2. Select the suitable cell removal rule in the run settings file. This rule specifies how conservation value is aggregated across features and space. (A simple ABF, rule #2, is often a good starting point.)

3. Type the following command line to notepad:

```
call zig4 -r settingsfile.dat specieslistfile.spp
outputfile.txt 0.0 0 1.0 0
```

and save the file in .bat format in the same directory as all the other files. In this command line, give the names of the settings file and species list file and define a suitable name for the output files. See section 3.2.1 for explanations of the four numbers in the call. The extension of this file may need to be manually changed to .bat (instead of .txt) in the file manager.



```
call zig4 -r settings.dat features_list.spp out/output1.txt 0.0 0 1.0 0
```

4. Open the Zonation GUI (**zig4gui.exe**). Right-click on the **Project View** and use the dialogue window for selecting the correct batch file. Initialize the Zonation run by right-clicking the *.bat* file name in the Project View and selecting **Queue**.

5. Monitor the run by double-clicking the run in the GUI Process View. An output map showing the progress of cell ranking will be drawn in the Map of the visual output. Representation curves are drawn in the Runtime plot of the visual output. A memo of the run is written in the Text section of the visual output. See section 4.3 for a more detailed description of visual outputs.

See section 3.4 for a detailed description of the output files that Zonation produces. In summary, it automatically produces at least eight output files with all analysis variants:

- .jpg and .emf maps of the landscape ranking showing the priority rank order with different colors. See section 4.3.1 for a detailed interpretation of the colors.
- A .curves.txt text file containing a list of species and relative weights used in the analysis. This file also includes columns of data reporting how the proportion of distribution remaining for each species develops through the landscape ranking.
- A .rank.asc raster file representing the order of cell removal (the priority ranking). This file can be used to produce map images in GIS software.
- A .prop.asc raster file contains additional information related to the priority ranking.
- A .wrscr.asc raster file. This file contains a weighted range size corrected measure (= weighted range size rarity) of conservation value for each cell that can be used as a scoring measure of value for cells. Another name for this measure is weighted range size rarity map.
- A .txt file. This file is written during the run. It keeps track of the files, settings, and analyses used during the run and is useful for tracing errors and checking that everything happened as it was supposed to. Note that some error messages or warnings may appear here.
- A .run_info.txt text file keeping track of input files, analyses, and settings used. This file will be created after the program has been closed.

Depending on analysis variant, there may be additional output files.

1.6 New features

This section lists the new features and small additions that are included in Zonation v. 4 and v. 3.1 that are not available in earlier versions.

Added to v. 3.1 since v. 2

(1) New analytical features

- **Alternative land-uses.** Allows consideration of multiple land uses to alleviate conflicting interests. The raster layers are input as biodiversity feature files and assigned negative weights (sections 2.9 and 5.3.4).
- **Community level analysis.** Focuses on community composition. Overlap in species composition is accounted for by similarity expansion (sections 2.8 and 5.3.1).
- **Landscape condition and retention.** An important part of the community level

analysis (sections 2.10 and 5.3.5).

- **Administrative units.** Allows planning over multiple administrative regions that have different priorities (sections 2.12 and 5.3.8).
- **Matrix connectivity.** Allows connectivity between multiple partially similar habitat types. Decreased matrix connectivity of cells on the edges of the planning region is corrected with the edge correction operation (sections 2.4.5 and 5.1.6).
- **Removal mask layer.** Has a new format and improved function. Multiple levels of prioritization hierarchy are now possible (section 3.3.3.9).
- **Analysis area mask.** Allows selection of a subset of cells within the planning region to be included in the analysis. In other words, unwanted external areas can be cut away from the analysis without modifying all input files (section 3.3.3.10).
- **Automated post-processing analyses.** Identifies management landscapes and compares two solutions in terms of overlap of the best or worst areas of the landscape (section 3.5.1).

(2) New software features

Added to v. 3.1 since v. 3.0

- **New interactive plotting window.** It is now possible to interactively generate combined plots of the following general results: curves for species, administrative units, groups of species, and groups within administrative units (see section 4.4). This interface also allows for visualizing histograms of occurrence levels of species for a given top fraction, as well as plots of distribution vs. representation levels (for groups).
- **New merged map window.** Can be used for visualizing combinations of input and output maps (see section 4.3.4).
- **New output files.** Contain information about administrative unit specific results (ADMU.XX.curves.txt), as well as per group results within administrative units (ADMU.XX.grp_curves.txt). The format of the general curves file (.curves.txt) has been slightly modified as well, as described in section 3.4.3.

Main additions to v. 4 since v. 3.1

- Much more efficient memory processing, allowing analyses orders of magnitude bigger than before. This comes also with comparable speed optimizations.
- New method to build corridors.
- Possibility to use arbitrary kernels in all connectivity methods.
- Additional plots and features in the interactive plots window of the GUI.
- New output files for transformed layers from 7 different methods.
- It is now possible to create Zonation setups from scratch through a graphical interface.

Outlook for the next version

Zonation 4 includes major improvements from Zonation 3, in terms of computational capacity, features, bug fixes, and useability of the GUI. At this time no guidance is given as to the expectation about the timing of the next major release of Zonation.

Part

II

2 Methods & algorithms

2.1 References

Two recent useful, open online references

Moilanen, A., B.J. Anderson, F. Eigenbrod, A. Heinemeyer, D. B. Roy, S. Gillings, P. R. Armsworth, K. J. Gaston, and C.D. Thomas. 2011a. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

Workflow for spatial conservation prioritization with Zonation

Lehtomäki, J. and A. Moilanen. 2013. Methods and workflow for spatial conservation prioritization using Zonation. *Environmental Modelling & Software*, 47: 128-137.

The original Zonation methods studies

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and C. D. Thomas. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences,* 272: 1885-1891.

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Important methodological studies (excluding connectivity)

Why Zonation produces higher return on conservation investment than target-based planning

Laitila, J. and A. Moilanen. 2012. Use of many low-level conservation targets reduces high-level conservation performance. *Ecological Modelling,* 247: 40-47.

Planning for impact avoidance and/or biodiversity offsetting

Moilanen, A. 2013. Planning impact avoidance and biodiversity offsetting using software for spatial conservation prioritization. *Wildlife Research*, 40: 153-162.

Kareksela, S., Moilanen, A., Tuominen, S. and J.S. Kotiaho. 2013. Use of Inverse spatial conservation prioritization to avoid biodiversity loss outside protected areas. *Conservation Biology*, 27: 1294–1303.

Balancing local and global representation over multiple administrative regions

Moilanen, A. and A. Arponen. 2011. Administrative regions in conservation: Balancing local priorities with regional to global preferences in spatial planning. *Biological Conservation*, 144: 1719-1725.

Balancing representation and landscape retention in community-level conservation prioritization

Moilanen, A., Leathwick, J.R. and J.M. Quinn. 2011. Spatial prioritization of conservation management. *Conservation Letters*, 4: 383-393.

Accounting for distributional uncertainty in Zonation, the distribution discounting method

Moilanen, A., B. A. Wintle, J. Elith and M. Burgman. 2006. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology*, 20: 1688-1697.

Reviewing and explaing the concepts of spatial prioritization (open online; not Zonation-specific)

Kukkala, A. and A. Moilanen. 2013. Core concepts of spatial prioritisation in systematic conservation planning. *Biological Reviews*, 88: 443-464.

Connectivity

BQP: One feature-specific method for generating reserve network aggregation

Moilanen, A., and B. A. Wintle. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

An extension of the BQP method to freshwater systems with different connectivity requirements upstream and downstream

Moilanen, A., Leathwick, J. and J. Elith. 2008. A method for freshwater conservation prioritization. *Freshwater Biology*, 53: 577-592.

Interaction connectivity; a single-species conservation prioritization study with multiple connectivity considerations

Rayfield, B., Moilanen, A. and M.-J. Fortin. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

Connectivity between habitat types (matrix connectivity) applied in expanding current protected area network

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

Edge adjustment for matrix connectivity

Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. and A. Moilanen. 2012. Effects of connectivity and spatial resolution of analyses on conservation prioritization across large extents. *Conservation Biology,* 26 (2): 294-304.

Corridor connectivity

Pouzols, F.M. and A. Moilanen. 2014. A method for building corridors in spatial conservation prioritization. *Landscape Ecology*, 49: 789-801. doi: 10.1007/s10980-014-0031-1.

Distribution smoothing, info-gap uncertainty analysis

Moilanen, A. and B. A. Wintle. 2006. Uncertainty analysis favours selection of spatially aggregated reserve structures. *Biological Conservation*, 129: 427-434.

Distribution smoothing and analysis of surrogacy power

Franco, A.M.A., Anderson, B., Roy, D.B., Gillings, S., Fox,R., Moilanen, A., and C.D. Thomas. 2009. Surrogacy and persistence in reserve selection: landscape prioritisation for multiple taxa in Britain. *Journal of Applied Ecology*, 46: 82-91.

Other general methods

Community-level conservation prioritization [conceptual, not Zonation-specific]

Arponen, A., Moilanen, A., and S. Ferrier. 2008. A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology*, 45: 1436-1445.

Community-level conservation prioritization and directed connectivity

Leathwick, J.R., A. Moilanen, S. Ferrier and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143: 984-991.

The benefit function approach to reserve selection [general, not Zonation-specific]

Arponen, A., Heikkinen, R., Thomas, C.D. and A. Moilanen. 2005. The value of biodiversity in reserve selection: representation, species weighting and benefit functions. *Conservation Biology*, 19: 2009-2014.

Basics of the information-gap decision theory for reserve selection uncertainty analysis

Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Y. Ben-Haim. 2006. Planning for robust reserve networks using uncertainty analysis. *Ecological Modelling*, 199 (1): 115-124.

Replacement cost analysis

Cabeza, M. and A. Moilanen. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation*, 132: 336-342. [conceptual, not Z-specific]

A replacement cost analysis and consideration of cross-taxon surrogacy in conservation planning

Kremen, C., A. Cameron, A. Moilanen, S. Phillips, C. D. Thomas, H. Beentje, J. Dransfeld, B. L. Fisher, F. Glaw, T. Good, G. Harper, R.J. Hijmans, D. C. Lees, E. Louis Jr., R. A. Nussbaum, A. Razafimpahanana, C. Raxworthy, G. Schatz, M. Vences, D. R.Vieites, P. C. Wright and M. L. Zjhra. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science*, 320: 222-226.

A replacement cost analysis with several levels of financial cost constraints applied to marine protected areas

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

Different applications with a specific characteristic

Habitat restoration and landscape dynamics

Thomson, J.R., Moilanen, A., McNally, R., and P. Vesk. 2009. Where and when to revegetate: A quantitative method for scheduling landscape reconstruction. *Ecological Applications*, 19: 817-828.

Planning for climate change

Carroll, C., Moilanen, A., and J. Dunk. 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. *Global Change Biology*, 16: 891-904.

Kujala, H, Moilanen, A., Araújo, M.B., and M. Cabeza. 2013. Conservation planning with uncertain climate change projections. *PLoS One,* 8: e53315, doi:10.1371/journal.pone.0053315.

Identifying least suitable areas for conservation to expand urban areas

Gordon, A., Simondson, D., White, M., Moilanen, A., and Bekessy, S.A. 2009. Integrating conservation planning and land-use planning in urban landscapes. *Landscape and Urban Planning*, 91: 183-194.

Balancing local and global representation over multiple administrative regions

Moilanen, A. Anderson, B.J., Arponen, A., Pouzols, F.M., Thomas, C.D. 2012. Edge artefacts and lost performance in national versus continental conservation priority area. *Diversity and Distributions*, 19: 171-183.

Balancing biodiversity and carbon

Thomas, C.D., Anderson, B.J., Moilanen, A., Eigenbrod, F., Heinemeyer, A., Quaife T., Roy, D.B., Gillings, S., Armsworth, P.R., and K. J. Gaston. 2013. Reconciling biodiversity and carbon conservation. *Ecology Letters*, 16-S1: 39-47.

Prioritization inside a national Natura 2000 network

Mikkonen, N. and A. Moilanen. 2013. Identification of top priority areas and management landscapes from a national Natura 2000 network. *Environmental Science & Policy*, 27: 11-20.

Planning in the urban environment

Bekessy, S.A., White, M., Gordon, A., Moilanen, A., Mccarthy, M.A.,and B.A. Wintle. 2012. Transparent planning for biodiversity and development in the urban fringed. *Landscape and Urban Planning*, 108: 140-149.

One more single-species application with muliple connectivity components

Sirkiä, S., Lehtomäki, J., Lindén,H., Tomppo, E. and A Moilanen. 2012. Spatial conservation prioritization of capercaillie (Tetrao urogallus) lekking landscapes in South-Central Finland. *Wildlife Biology*, 18: 337-353.

Rare example using extensive genetic data

Taberlet, P., Gugerli, F., Zimmermann, N.E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Mirek, Z., Moilanen, A., *et al*. (54 authors). 2012. Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters,* 15: 1439-1448.

Case-study that uses many different types of data

Sharafi, S.M. , A. Moilanen, White, M., and M.A. Burgman. 2012. Integrating environmental gap analysis with spatial conservation prioritization: A case study from Victoria, Australia. *Journal of Environmental Management*, 112: 240-251.

Case study showing how Zonation works better than target-base planning

Di Minin, E. and A. Moilanen. 2012. Empirical evidence for reduced protection levels across biodiversity features from target-based conservation planning. *Biological Conservation*, 153: 187-191.

National prioritization study that used larger data than had even been used at the time

Kremen, C., A. Cameron, A. Moilanen, S. Phillips, C. D. Thomas, H. Beentje, J. Dransfeld, B. L. Fisher, F. Glaw, T. Good, G. Harper, R.J. Hijmans, D. C. Lees, E. Louis Jr., R. A. Nussbaum, C. Raxworthy, A. Razafimpahanana, G. Schatz, M. Vences, D. R. Vieites, P. C. Wright , M. L.Zjhra. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science,* 320: 222-226.

Nice marine planning example

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters,* 1: 91-102.

Z-relevant Book Chapters

Moilanen, A. and I. Ball. 2009. Heuristic and approximate optimization methods for spatial conservation prioritization. Pp. 58-69 in Moilanen, Wilson and Possingham (Eds.), Spatial Conservation Prioritization, Oxford University Press.

Moilanen, A., Kujala, H. and J. Leathwick. 2009. The Zonation framework and software for conservation prioritization. Pp. 196-210 in Moilanen, Wilson and Possingham (Eds.), Spatial Conservation Prioritization, Oxford University Press.

The manuals

This version: Moilanen, A. Montesino Pouzols, F., Meller, L., Veach, V., Arponen, A., Leppänen, J., Kujala, H. 2014. Zonation spatial conservation planning methods and software v.4, User manual, 288 pp. cbig.it.helsinki.fi/software/zonation

Moilanen, A., Meller, L., Leppänen, J., Montesino Pouzols, F., Arponen, A. & H. Kujala. 2012. Zonation spatial conservation planning framework and software v. 3.1, User manual, 287 pp. cbig.it.helsinki.fi/software/zonation

Moilanen, A. and H. Kujala. 2008. Zonation spatial conservation planning framework and software v. 2.0, User manual, 136 pp.

Moilanen, A. and H. Kujala. 2006. Zonation spatial conservation planning framework and software v. 1.0, User manual, 126 pp. Edita, Helsinki, Finland.

For those who would wish to familiarize themselves more broadly with recent literature concerning spatial conservation planning, we recommend using Web of Science (or a similar search facility) with key words such as, **reserve selection**, **spatial conservation prioritization**, **reserve network design**, **site selection algorithm**, **area prioritization**, **systematic conservation planning**, **spatial conservation planning**, and **spatial optimization**. Journals such as *Biological Conservation*, *Conservation Biology*, *Conservation letters*, *Ecological Applications*, *Journal of Applied Ecology* and *Environmental Modeling and Assessment*, among others, include many studies concerning quantitative conservation prioritization methods.

2.2 The Zonation meta-algorithm

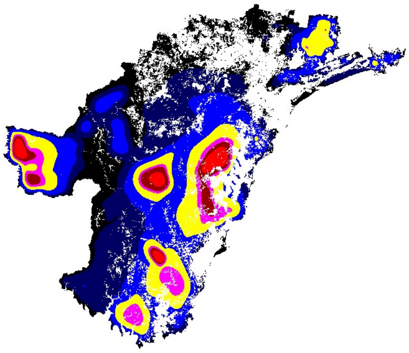
The important characteristic of the Zonation algorithm is that it can balance many considerations in a sensible manner during the prioritization. Such characteristics include local occurrence levels of many biodiversity features, multiple costs, needs of other land uses, connectivity, etc.

The Zonation algorithm (Moilanen *et al*. 2005) produces a hierarchical prioritization of the conservation value of a landscape. The algorithm is hierarchical in the sense that the most valuable 5% of the landscape is within the most valuable 10%, the top 2% is in the top 5%, etc. Generally speaking, Zonation iteratively ranks (removes) cells one by one from the landscape using minimization of marginal loss as the criterion to decide which cell is ranked (removed) next. The order of cell removal is recorded and it can be post-hoc used to select any given top fraction (e.g., the best 10%) of the landscape. Information about the decline of representation levels of biodiversity features is collected simultaneously during cell removal.

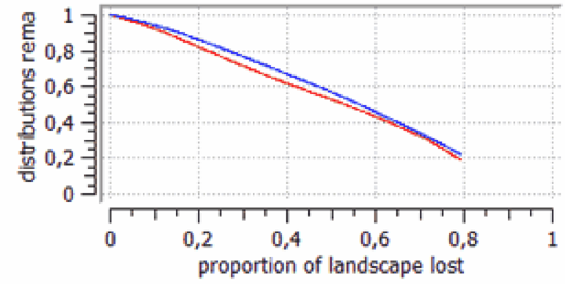
The Zonation meta-algorithm:

1. Start from the full landscape. Set rank r = 1.
2. Calculate marginal loss following from the removal of each remaining site i, δi. Complementarity, the balancing between features and other considerations, is accounted for in this step.
3. Remove the cell with smallest δi. Define removal rank of i as r. Set r = r+1, and return to step 2 if there are any cells remaining in the landscape.

Thus, sites are ranked based on occurrence levels of biodiversity features, costs, connectivity, and other relevant considerations. The least valuable cells are removed one (or more) at a time, producing a sequence of landscape structures with increasingly important biodiversity features remaining. We want to emphasize that the result of a Zonation analysis is not a single set of areas. Rather, it is (1) the nested ranking of cells (see figure below, left) and (2) a set of performance curves describing the performance of the solution at the given level of cell removal (see figure below, right).



Example figure illustrating the nested ranking produced by Zonation.



Example graph illustrating the performance curves produced by Zonation.

The Zonation meta-algorithm can, among other things, answer three questions frequently encountered in conservation biology:

1. Which parts of the landscape, totaling x% of landscape cost or area, have the highest conservation priority (ranking), or,
2. Which part of the landscape includes at least y% of the distribution of each species (proportional coverage selection)?
3. Where are the ecologically least valuable parts of the landscape? These areas may be good for impact avoidance in the context of development projects.

Whether the Zonation algorithm makes any sense at all depends on the definition of marginal loss (δ_i, step 2 in the algorithm above). This is defined by a separate cell removal rule (see section 2.3), which implements our conception of how conservation value is aggregated across the landscape and across species (or other biodiversity features). The Zonation method can thus be divided into two parts, the Zonation meta-algorithm and the cell removal rule (= definition of marginal loss), which should not be confused. The cell removal rule should be seen as a separate component with several alternatives that have different interpretations. Note that the notion of complementarity is inherent in the way the cell removal rule is defined.

The algorithm applied by Zonation can be classified as a reverse, accelerated, iterative heuristic algorithm. It is 'reverse' because it starts from the full landscape and removes cells (this is important for the treatment of connectivity). The algorithm is 'accelerated' because it has the option of removing more than one cell at a time via the warp factor (see section 3.3.2.3). There is one feature that, according to Moilanen *et al*. (2005), is a part of the Zonation algorithm but should more appropriately be regarded as a relevant detail for which there are alternatives. This is edge removal (see section 3.3.2.3), a feature that allows cells to be removed only from the edge of the remaining landscape. Edge removal promotes maintenance of structural habitat continuity in the removal process. It also makes the cell removal process much faster with large landscapes, which is the primary reason for using it.

Literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences,* 272: 1885-1891.

Moilanen, A., B.J. Anderson, F. Eigenbrod, A. Heinemeyer, D. B. Roy, S. Gillings, P. R. Armsworth, K. J. Gaston, and C.D. Thomas. 2011. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

2.3 Aggregating conservation value: the cell removal rule

This section is largely based on Moilanen (2007). See also Moilanen *et al*. (2011).

The Zonation meta-algorithm is the same for all analyses described in this manual, but the actual removal order of cells depends on the cell removal rule. The rule determines which cell leads to smallest marginal loss of biodiversity value. There are four cell removal rules that have specific characteristics that may correspond to slightly different mental models about conservation prefernces:

1. Core-area Zonation (section 2.3.1)
2. Additive benefit function (section 2.3.2)
3. Target-based planning (section 2.3.3)
4. Generalized benefit function (section 2.3.4): a two-piece power function that can assume very versatile forms, allowing flexibility in the specification of conservation value.

Core-are Zonation, the additive benefit function, and target-based planning are conceptually different from each other. The generalized benefit function is a generalization of the additive benefit function.

Sometimes it may be necessary to create networks where sites are removed in a random order. For example, in order to compare different methods and their effectiveness in prioritization, it may be beneficial to do randomized prioritizations to get a baseline representation level. This is possible in Zonation with the fifth cell removal option:

5. Random removal (section 2.3.6)

Note that core-area Zonation has the property that it can identify important, otherwise species-poor, locations where a single or a few biodiversity features have an important occurrence. The additive benefit function analysis gives more weight to locations with high biodiversity richness. Therefore, it may be useful to run both analyses and compare the results. If the top-fractions do not agree, then there are some biodiversity-rich areas but also some biodiversity-poor areas with occurrences of otherwise rare biodiversity features. Running both core-area Zonation and the additive benefit function analysis may reveal information that is interesting for conservation planning.

Literature

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Moilanen, A., B.J. Anderson, F. Eigenbrod, A. Heinemeyer, D. B. Roy, S. Gillings, P. R. Armsworth, K. J. Gaston, and C.D. Thomas. 2011a. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

2.3.1 Basic core-area Zonation

This section is largely based on Moilanen *et al*. (2005) and Moilanen (2007).

In basic core-area Zonation (commonly *CAZ*) cell removal is done in a manner that minimizes biological loss by picking cell *i* that has the smallest occurrence for the most valuable feature over all biodiversity features in the cell. In other words, the cell gets a high value if even one species has a relatively important occurrence there. The removal is done by calculating a removal index δ_i (minimum marginal loss of biological value) for each of the cells, where

$$\delta_i = \max_j \frac{q_{ij} w_j}{c_i}, \qquad (1a)$$

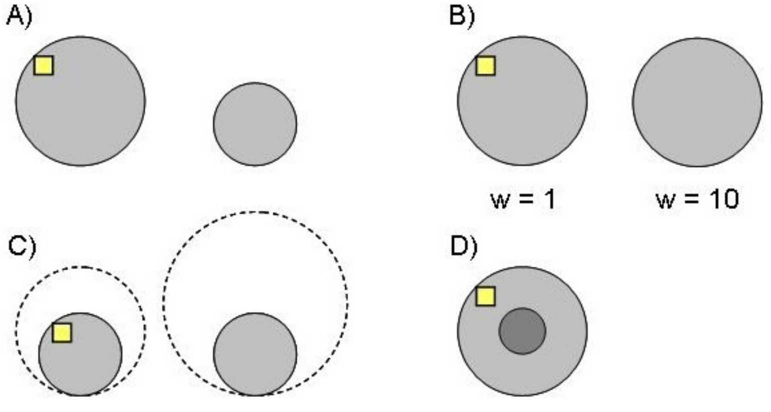
where w_j is the weight (or priority) of species *j* and c_i is the cost of adding cell *i* to the reserve network. When running the analysis, the program goes through all cells and calculates a δ_i value based on the feature that has the highest weighted proportion of distribution remaining in the specific cell (and thus represents the highest biological value to be lost if the cell is removed). The cell which has the lowest δ_i value will be removed.

Complementarity and balancing between features: The critical part of the equation is q_{ij}, the proportion of the remaining distribution of feature *j* located in cell *i* for a given set of sites (the set of cells remaining). When a part of a feature's distribution is removed, the proportion located in each remaining cell goes up. This means that Zonation tries to retain core areas of all species until the end of cell removal, even if the feature is initially widespread and common. Thus, at first only cells with occurrences of common features are removed. Gradually, the initially common features become more rare, and cells with increasingly rare feature occurrences start disappearing. The last site to remain in the landscape is the cell with the highest (weighted) richness. This is the site that would be kept last if all else was to be lost. Note that Eq. (1a) can alternatively be expressed as (Moilanen *et al*. 2005)

$$\delta_i = \frac{1}{c_i} \max_j \frac{q_{ij} w_j}{q_j}, \qquad (1b)$$

where q_{ij} is the fraction of the original full distribution of feature *j* residing in cell *i* according to data, and q_j is the fraction of the original distribution of feature *j* in the remaining set of cells.

The min-max structure of the equation also indicates a strong preference for retaining the best locations with the highest occurrence levels. Thus, the program can spare otherwise biodiversity-poor cells if they have a very high occurrence level for one rare feature. It is important to understand that core-area Zonation does not treat probabilities of occurrence as additive: ten locations with p=0.099 is not the same as one location with p=0.99. However, this is strictly true when the analysis is based on biological value only and when a landscape cost layer is not used. When cost is used, cell removal is based on local conservation value divided by cell cost (efficiency), and now a high value for a cell can be explained with either (1) a very high occurrence level for some species or (2) low cost for the cell. Thus, when cost information is used, the interpretation of a core-area becomes vague, and this should be recognized in planning. Therefore it is not recommended to use cost layers when trying to find the biologically most important areas with core-area Zonation.



The figure above illustrates the principles that core-area Zonation implements in numerical form. Essentially, the question is this: If there are two (multiple) species, and a fraction of one distribution will be lost (here one cell, marked as yellow), then where is the most preferable location to lose the cell from? (A) If there are two otherwise identical species, but one has a larger range remaining, then it is preferable to lose from the species that has the larger range. (B) If the species are otherwise equal, but one has relatively higher weight, then it is preferable to lose a cell from the distribution of the species with a lower weight. (C) There are two presently equal species with equally wide distributions. It is preferable to remove a site from the species that has a smaller historical reduction in the range (dashed line). (D) Within the distribution of a species, it is preferable to lose a cell from a location with a relatively low occurrence density (light gray).

Literature

Moilanen, A., B.J. Anderson, F. Eigenbrod, A. Heinemeyer, D. B. Roy, S. Gillings, P. R. Armsworth, K. J. Gaston, and C.D. Thomas. 2011. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences,* 272: 1885-1891.

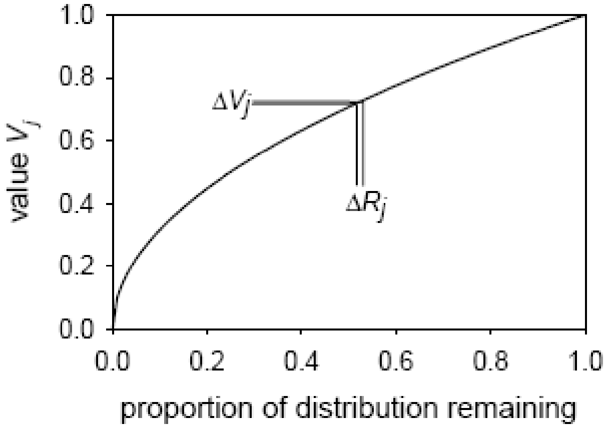
2.3.2 Additive benefit function

The additive benefit function (commonly ABF) was described by Moilanen (2007). It can be interpreted as "minimization of aggregate extinction rates via feature-specific species-area curves". This interpretation emerges because the mathematical shape of the ABF analysis is the same as that of the canonical species-area curve that originated from the literature of island biogeography in the 1960's.

Compared to core-area Zonation, the additive benefit function takes into account all (weighted) feature proportions in a given cell instead of the one feature that has the highest value. The program first calculates the loss of representation for each feature as cell *i* is removed. The δi value of the cell is simply the sum over feature-specific declines in value following the loss of cell *i*:

$$\delta_i = \frac{1}{c_i} w_j \sum_j \Delta V_j = \frac{1}{c_i} w_j \sum_j \left[V_j(q_j) - V_j(q_j - i) \right],$$

in which q_j is the representation of feature *j* in the remaining set of sites, and (q_j-*i*) indicates the set of remaining cells minus cell *i*. Here w_j is the weight of the feature *j* and c_i is the cost (or area) of planning unit *i.* As in core-area Zonation, the cell with the smallest δi value will be removed.



The figure above shows a benefit function for feature *j*. When a grid cell is removed from the landscape, the representation of each feature (ΔR_j) occurring in the removed cell goes down by a small fraction and the respective value for that feature declines by ΔV_j. The total marginal loss in value is simply a sum over feature-specific losses. Note that here the feature has a standard weight of 1.0, but as with core-area Zonation, it is possible to weight biodiversity features differently when using the additive benefit function. The effects of weighting are seen on the scale of the y-axis which will go from 0.0 to feature weight w_j instead of going from 0.0 to 1.0.

Because the additive benefit function sums value over all features, the number of features in a cell has a higher significance compared to basic core-area Zonation. For example, using the additive benefit function might lead to a situation where biodiversity-poor cells are removed even though they have a high occurrence level for one or two rare features, because the δ_i value of these cells is smaller than that for cells that have many common features with high representations. Thus, using the additive benefit function typically results in a reserve network that has a higher performance on average over all features but which retains a lower minimum proportion of original distributions for the worst-off features compared to core-area Zonation (see figure of the first three removal rules in section 2.3.5).

Literature

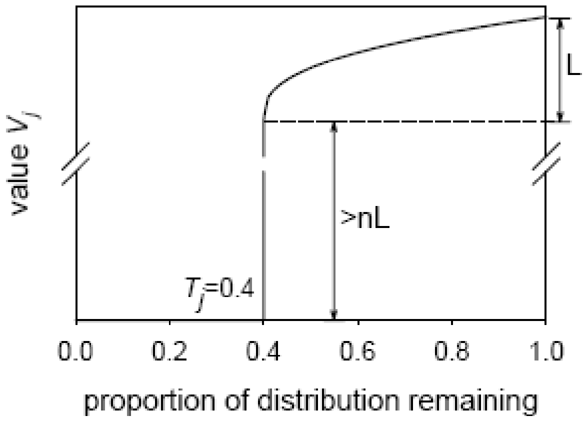
To find more information about the use of benefit functions, see the references for core-area Zonation plus additionally:

Arponen, A., Heikkinen, R., Thomas, C.D. and A. Moilanen. 2005. The value of biodiversity in reserve selection: representation, species weighting and benefit functions. *Conservation Biology,* 19: 2009-2014.

2.3.3 Target-based planning

For whatever reason, Zonation has not been used much for target-based planning even though the capability has been perfectly functional for many years now.

Target-based planning (TBF) is implemented in Zonation by using a very particular type of a benefit function (see figure below). The purpose of this special functional form is to enable the Zonation process to converge to a solution that is close to the proportional coverage minimum set solution for the data. In this function, value, Vj, is zero until representation, Rj, reaches the target, Tj. Then there is a step with the height of (n+1), where n is the number of features. When Rj increases above Tj and approaches 1, there is a convex increase in value with a difference in value of [Vj(1)-Vj(Tj)]=1. This means that the loss in value from dropping any one feature below the target is higher than any summed loss over multiple species that stay above the target.



The idea is that as cells are iteratively removed, biodiversity feature representations will approach the feature-specific targets from above, and that the convex formulation with increasing marginal losses will force features to approach targets in synchrony in terms of lost value. Thus, as one of the features approaches the target level, the program starts to avoid removing cells that contain that particular feature (at the expense of other features) in order to retain the target. At some point it will not be possible to remove any more cells without violating the target for at least one feature. After one of the features has declined below the target, the remaining distribution of that feature has no value for the reserve network. Thus removing cells where only this feature occurs does not increase the loss of biological value from network anymore.

Note, however, that the definition of how marginal value is calculated does not change from that of the additive benefit function (section 2.3.2). Feature occurrences are considered as additive in this cell removal as well, and the cell that has the lowest marginal value summed across all features will be removed next. Also, when using target-based planning the feature-specific weights have no function as the goal is to retain a given proportion of distributions for all of the features. However, it is possible to set different targets for different features. It is also recommended to avoid using very high warp factors to allow the program to find the most optimal solution near the targets.

A solution computed with the target-based planning cell removal rule needs to be interpreted with special care. For the other cell removal rules, conservation value from the perspective of one feature increases as the number of cells in the landscape decreases. With target-based planning, the conservation value for one feature in all cells goes down to zero immediately after the target has been violated. Using targets often leads to a non-optimal solution, especially at resource levels smaller than what allows achieving all targets. If you are using the target-based planning removal rule, we recommend running the same analysis with another cell removal rule to evaluate the sensitivity of the solution to using targets.

Literature

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Why other Zonation planning modes produce higher return on conservation investment than target-based planning

Laitila, J. and A. Moilanen. 2012. Use of many low-level conservation targets reduces high-level conservation performance. *Ecological Modelling,* 247: 40-47.

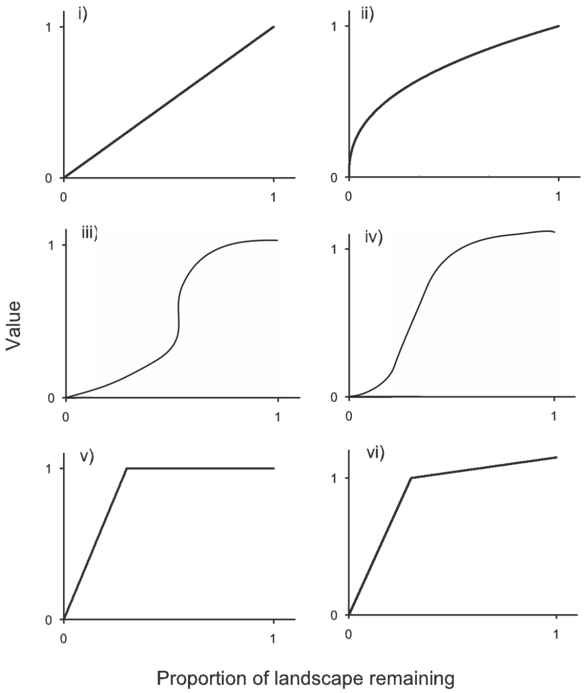
Di Minin, E. and A. Moilanen. 2012. Empirical evidence for reduced protection levels across biodiversity features from target-based conservation planning. *Biological Conservation*, 153: 187-191.

2.3.4 Generalized benefit function

The generalized benefit function cell removal rule allows very flexible function shapes. It otherwise operates like the ABF, but the function can take more flexible shapes, including sigmoids. The function is defined in parts by two power functions:

$$V_j(R_j)=\left\{\begin{array}{ll} w_1\left(\frac{R_j}{T_j}\right)^x & if \quad R_j \le T_j \\ w_1+w_2\left(\frac{R_j-T_j}{1-T_j}\right)^y & if \quad T_j < \mathrm{R}_j < 1 \end{array}\right\}$$

The generalized benefit function can take many shapes depending on the values given to the parameters. Some shapes are illustrated in the following figure. The parameters are defined in the following table.



	w_1	w_2	T_j	x	Y
i) Linear	w_j	0	1.0	1.0	NA, dummy=1.0
ii) Power function (=ABF)	w_j	0	1.0	<1 or >1	NA, dummy=1.0
iii) Mild sigmoid	w_j	same order as w_j	at inclination point	>1	<1, e.g. 1/x
iv) Steep sigmoid – step imitation	w_j	same order as w_j	at step	>>1	<<1, e.g. 1/x
v) Ramp	w_j	0	at step	1.0	NA, dummy=1.0
vi) Ramp, with linear over-representation	w_j	<< w_j	at step	1.0	1

The parameter definitions are suggestive, and the exact shape of the function can be determined easiest by plotting it. To use the generalized benefit function as a cell removal rule, the parameters of the function need to be given in the biodiversity feature list file (section 3.3.2.2).

2.3.5 General differences between cell removal rules

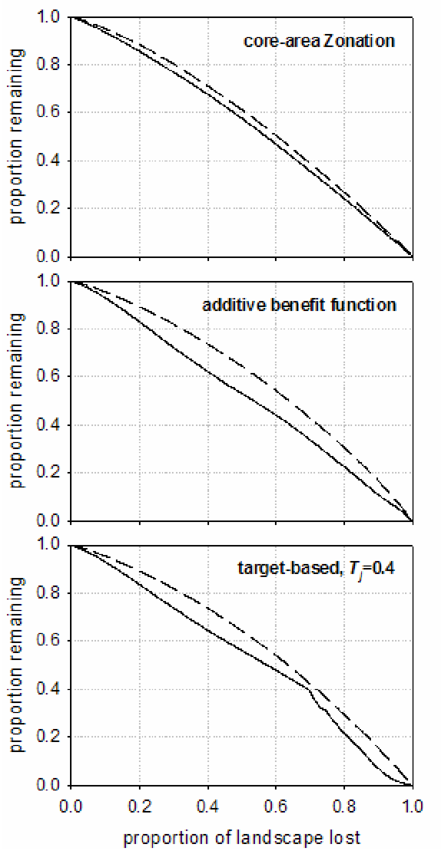
It is important to realize that there may be significant differences between different cell removal rule solutions, and the most preferable solution method depends on the goals of planning. Different cell removal rules may be conceptually better suited for different situations.

- Core-area Zonation is most appropriate when there is: (1) a definite set of biodiversity features, all of which are to be protected and tradeoffs between features are discouraged; in other words, one is confident that it is important to protect all these features (2) the hierarchy of solutions and easy weighting of features is desired; and (3) importance is given to core-areas (i.e., locations with the highest suitability for a given feature in terms of abundance or high probability of occurrence). Occurrences in cells are not additive, meaning that twenty locations with p=0.05 is not the same as one location with p=1.0.
- The **additive benefit function** formulation may be more appropriate when: (1) the features are essentially surrogates or samples from a larger regional feature pool, and tradeoffs between features are allowed, (2) a result with high return on conservation investment is desired, and (3) the hierarchy of solutions and easy weighting of features is desired.
- The **target-based** formulation is most appropriate when: (1) there is a definite set of features, all of which are to be protected; (2) there is a credible way for setting the feature-specific targets; (3) there is no need for a hierarchy of solutions; (4) occurrences are additive; and (5) easy weighting of features is not needed. In target-based planning, feature weighting is essentially done by assigning different targets for each feature.

The figure below (from Moilanen 2007) illustrates some general differences between the core-area Zonation, the additive benefit function, and the target-based benefit function. Here, the lines show what proportion of biodiversity feature distributions remain in the landscape as cells are progressively removed. Overall, the additive benefit function has the highest average proportion over all features retained (dashed line), but it simultaneously has the smallest minimum proportion retained (solid line). This is because the additive benefit function favors biodiversity-rich areas over those areas that might be significant for the existence of one or few features, but that otherwise are biodiversity-poor. Although the core-area Zonation gives a relatively low average proportion of feature distributions, it has a high minimum proportion, because it retains the most significant areas of all features (the "core areas") until the end, even though these areas might be unsuitable for all the other features. The targeting benefit function does well in terms of finding the highest level of cell removal without having any feature-specific targets violated. However, once representation targets are violated, it performs relatively poorly in terms of the mean and minimum fraction over features retained. The problem with the targeting benefit function is that it is aimed at good performance at one particular set of targets, but the hierarchy of solutions is missing in the sense that good overall performance at other levels of cell removal, especially at a level where targets have been violated, cannot be guaranteed.

There are also differences between the cell removal rules in terms how much area they require for achieving a set conservation target. To get a given minimum fraction across features, core-area Zonation requires more cells than the benefit function variants. This is because benefit function variants take occurrences as additive whereas core-area Zonation prefers the locations with the very highest occurrence levels. However, if one investigates the number of cells needed to get the target distribution for an individual feature, then core-area Zonation may require fewer cells because it prefers the higher-quality (density) cells. Thus, the benefit function variants generate landscapes with many biodiversity features occurring simultaneously at potentially low occurrence levels and with high overlap between features. Core-area Zonation produces solutions with biodiversity features occurring at higher densities, but with less overlap between features.

All these differences can logically be expected to occur with any data set, with the magnitude of differences completely depending on the nestedness of biodiversity feature distributions. Differences will be largest when there are both (1) substantial regional differences in biodiversity richness and occurrence levels and (2) a generally low overlap between biodiversity feature distributions. In this case, the core-area Zonation could catch cores of features occurring in biodiversity-poor areas, whereas the additive benefit function would concentrate the solution more towards biodiversity-rich locations where cells have a high aggregate value over features.



Core-area Zonation and presence-absence data

When biodiversity feature data is in binary presence-absence form, all cells where a feature is present receive a value of 1. In such data, it may appear that there are no core-areas particularly important for the feature, and it might seem pointless to use core-area Zonation as the cell removal rule. This is not the case for two primary reasons. First, any additional analyses such as aggregation methods or the uncertainty analysis will cause the value of the cell to be calculated based on other features (e.g., connectivity of the cell) in addition to the feature data. Differences between areas where the feature is present do, therefore, emerge. Second, because core-area Zonation does not treat feature occurrences as additive, but tries to retain occurrences of individual feature in the landscape as long as possible, there will still be a significant difference in the cell removal process between core-area Zonation and additive benefit functions. We illustrate this with an example. Assume we have a landscape where 7 different terrestrial species occur. Six of these species have overlapping distributions and one (denoted here as species A) has a distribution isolated from the other species. Because benefit functions take species occurrences as additive, the cells in sites where distributions of several species overlap receive a higher value than the cells where only one species occurs, as is the case with species A. Thus, in the cell removal process the additive benefit function would always favor cells with multiple species over the cells of species A, which would lead to unequal preservation of species (in other words species A would lose its distribution much more quickly than the other species). In contrast, core-area Zonation would retain all species distributions equally, meaning that species A would lose its distribution at the same rate as the other six species. This conclusion stays the same even when using presence-absence data.

2.3.6 Random removal

The fifth cell removal rule in Zonation is random removal. This option might be useful for quality control (e.g., to explore different methods for cell prioritization and to find the baseline representation level regardless of the efficiency of the methods used). Random removal removes cells in a random order with no consideration of occurrence levels of features in the cells. Here, the assumption with respect to the analysis outcome is that the average representation curve should be a straight diagonal line going from one to zero when the fraction of the landscape goes down from one to zero. This conclusion follows from the mathematical fact, that barring connectivity effects, removing a random fraction x of the landscape will on average remove the same fraction x of all species distributions.

2.4 Inducing reserve network aggregation

Fragmentation is an undesirable characteristic in reserve design, and it has been concluded in many studies that species persist poorly in small and isolated patches. Also, implementing a fragmented reserve network may be awkward and expensive. This section introduces different aggregation methods that can be used in the Zonation program. These methods produce relatively more compact solutions. Note, however, that aggregation always involves trade-offs. There is usually an *apparent* biological cost in more aggregated solutions because in many cases it is necessary to include lower-quality habitats in the reserve network in order to increase connectivity. In reality this apparent loss is more than offset by benefits of having a well-connected area (but these benefits may not be seen if they are not explicitly modeled). Thus, it is recommended to use aggregation methods in reserve planning since the cost of losing a minor amount of in local occurrence levels areas is usually low compared to the benefits of high connectivity. For more information on true and apparent costs related to aggregation, see Moilanen and Wintle (2006) and (2007). Starting with Zonation v.4 it is also possible to induce landscape connectivity via corridors, see section 2.13 Corridor connectivity below.

There are some distinct differences between the aggregation methods in Zonation, and choosing the right one depends on conservation targets and computational issues.

- **Boundary Length Penalty** (BLP) has been the most commonly used way to introduce aggregation to reserve planning. However, it is important to understand that BLP is a general, non-feature-specific aggregation method which does not asses the actual effects of fragmentation on species. Rather, the method uses a penalty on a structural characteristic of the reserve network (boundary length) to produce a more compact reserve network solution. The method is computationally quick and effective but might not be the most biologically realistic. Good for generating aggregation for management purposes.

- **Distribution Smoothing** is a feature (species)-specific aggregation method which retains areas that are well connected to others, thus resulting in a more compact solution. The connectivity of cells is determined with a smoothing kernel, which means that the value of a cell is "smoothed" to the surrounding area. Another way of looking at distribution smoothing is that it does a two-dimensional habitat density calculation, identifying areas of high habitat quality and density. Consequently, cells that are surrounded by many occupied cells receive a higher value than the isolated ones. The width of the smoothing kernel is feature-specific, implicitly expressing the feature's dispersal capability or scale of landscape use. This aggregation method is computationally very quick. However, it assumes that fragmentation (low connectivity) is generally bad for all features and it always favors uniform areas over patchy ones. If there is knowledge of the typical home-range size of the feature, then the radius of the kernel could be approximated as the radius of the typical home-range, if converted into a circle.

- **Boundary Quality Penalty** (BQP) is biologically the most realistic aggregation method included in Zonation. This method describes how the local occurrence level of a feature (species) in a site is influenced by the loss of surrounding habitat. The change in local occurrence is based on feature-specific responses to neighborhood habitat loss, thus local occurrence may also increase for a feature that benefits from fragmentation. The downside of this method is the required computation time, which is much longer compared to the other aggregation methods. This is because each cell removal influences the occurrence levels in all remaining neighborhood cells and needs to be accounted for in the prioritization process.

- **Directed connectivity** (Neighborhood Quality Penalty; NQP) is a generalization of BQP where the connectivity between sites is strictly directed, such as in riverine systems. This option demands the use of planning units, groups of cells (rather than single cells) that are removed as a one during the landscape ranking process. (In freshwater planning these units would typically correspond to catchments.) The occurrence level of a feature in a focal planning unit is influenced by the removal of other planning units "upstream" or "downstream" from the focal unit. Following the philosophy of BQP, the change in local occurrence is based on feature-specific responses to nearby habitat loss. The computation times are relative to the count (average size) of the planning units: many (smaller) planning units means longer computation times and vice versa.

- **Matrix connectivity** is a connectivity calculation between multiple partially similar habitat types (or other linked entities). In matrix connectivity, the local occurrence level for each focal feature is multiplied by it's connectivity to multiple other features. This can be useful for example when nominally different but ecologically partially similar habitat types contribute to each other's connectivity. For example, spruce forest and pine forest could be different features in conservation planning. However, they share a large fraction of species and it is clear that both spruce and pine forest help the connectivity of many species. Therefore, connectivity calculation for spruce forest should also account for pine forests, and vice versa. This feature can even be used to express preference for heterogeneous habitats when a mixture of certain habitat types is more desirable than a homogeneous landscape.

- Another feature that somewhat influences aggregation of the prioritization solution is the **edge removal feature**. This feature only allows cells to be removed from the edge of the remaining landscape. This option can, however, potentially create problems in the rare situation that a large area of poor habitat is completely surrounded by good habitat. In this case, the Zonation program will first have to remove some of the good habitat from the edge before reaching the poor habitat.

- **Interaction connectivity** is connectivity between a pair of features. It can be either positive, thereby emphasizing the parts of two distributions that are relatively well connected, or negative, in which case proximity between two distributions decreases priority. Interaction connectivity can occur for example between predator and a prey, or between present and future distributions of a species (or whatever feature).

- **Corridors**. Zonation v4 also includes a method for retaining corridor-type connectivity during the priority ranking.

Simultaneous use of multiple aggregation methods is possible, but may cause difficulties for interpreting results. We therefore recommend careful use of multiple aggregation methods. The BLP, interaction, and distribution smoothing aggregation features combine well with other aggregation features. Note that for distribution smoothing to be combined with other features smoothed layers must be generated and saved in advance using the "save transformed layers" method.

Use of different shapes of dispersal kernels. Note also that the connectivity methods that are based on kernels, such as distribution smoothing and matrix connectivity, can use arbitrary kernels since Zonation v.4. Traditionally the kernel based methods have used the negative exponential kernel as the default and only supported option. With newer Zonation versions you can choose different predefined simple kernel shapes, or define your own custom kernel (of arbitrary shape) in a text file. The kernel used can be chosen individually for different biodiversity features or groups of features. To do so you need to assign biodiversity features to an "arbitrary kernel group", see sections 3.3.3.12 (how to create arbitrary kernel groups), and 3.3.3.18 (how to define arbitrary kernel shapes).

2.4.1 Boundary Length Penalty (BLP)

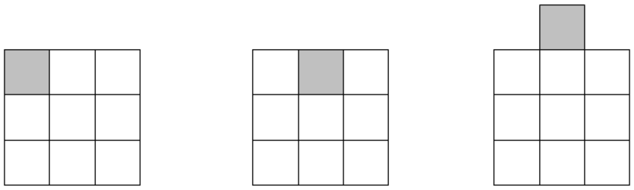
This section is largely based on Moilanen and Wintle (2007).

The boundary length penalty (BLP) has been the most common way of adding aggregation to a reserve network in the literature of spatial prioritization. It has been the only way of dealing with connectivity in several site selection formulations or software packages. This method is qualitative in the sense that the estimated conservation value of individual cells (or consequently the conservation value of the entire reserve network) is not influenced by the degree of fragmentation, but rather, aggregation is induced via a penalty given for the boundary length of the reserve. BLP does not include biodiversity feature-specific components.

When using the boundary length penalty, the hierarchy of cell removal is based on both biodiversity feature occurrence levels in cells and the increase/decrease of boundary length that results from the removal of a cell. The boundary length penalty can, in the context of core-area Zonation, be formulated as:

$$\delta_i = \max_j \frac{Q_{ij}(S)w_j}{c_i} + \beta \cdot \Delta(BL/A),$$

where $Q_{ij}(S)$ is the proportion of the remaining distribution of feature j present in cell i (calculated for the remaining cells, S), $\Delta(BL/A)$ is the change in the ratio of boundary length to area of the reserve network following removal of cell i, and β is a constant defining the strength of the boundary length penalty. [Note that a sensible value for β needs to be found by experimentation.] If cell removal decreases boundary length, $\Delta(BL/A)$ receives a negative value and the value of δ_i for cell i decreases, indicating that it is relatively advantageous to remove the cell because removing it reduces fragmentation.



$\Delta L = 0$ $\Delta L = +2$ $\Delta L = -2$

The figure above shows how different cell removal options influence the boundary length. The boundary length is calculated in the terms of cell edges. Removing the gray cell in the first example results in no changes in boundary length as two edges are removed while another two are gained. In the second example, the cell removal leads to the loss of one edge, but also to the gain of three new edges. The result is a total change in boundary length of +2. To get from ΔL *to* $\Delta (BL/A)$ one needs to account for both the change in boundary length and the decrease of the reserve area by one.

Note that like distribution smoothing, the BLP may be expected to perform poorly for features that happily occur in fragmented habitats. This is because the BLP qualitatively favors structurally connected areas, and it will therefore have a tendency to remove small habitat fragments from the solution irrespective of whether some features can actually persist in them or not.

Instructions to using BLP in Zonation can be found in section 5.1.2.

Literature

Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

2.4.2 Distribution smoothing

This section is largely based on Moilanen *et al.* (2005) and Moilanen and Wintle (2006).

With distribution smoothing, planning is based on a connectivity surface computed from the original biodiversity feature distributions that have been input in Zonation. The calculation applied to each feature distribution is identical to the calculation of a metapopulation-dynamic connectivity measure, where the connectivity value is directly proportional to the number of migrants expected at a given location in the landscape. Technically, the computation is a two-dimensional kernel smoothing using a feature-specific parameter (width of the smoothing kernel).

For practical purposes, distribution smoothing identifies areas that have on average high occupancy levels for features. The smoothing very effectively identifies important semi-continuous regions where the feature has overall high levels of occurrence, although not necessarily in every grid cell. In contrast, relatively scattered occurrences in fragmented habitat lose value. As the distribution of the feature in the landscape becomes smoother, occurrence levels in fragmented areas end up reduced compared to those in semi-continuous areas.

Distribution smoothing should be used with care if the data includes a species that lives happily as a metapopulation in a fragmented habitat. The smoothing kernel should be narrow for this species if the habitat matrix is taken as partially suitable for the species. Distribution smoothing is a convenient technique to apply because it can be run as a relatively fast pre-processing step before going on to the Zonation analysis itself. The appropriate level of smoothing for a given species would be determined based on a conception of the typical dispersal distances for that species, or from information concerning mean home range sizes for the species.

Technically, when using distribution smoothing, the value for feature *j* in a focal cell *i* is

$$O_{ij}' = \sum_{x} \sum_{y} \exp(-\alpha_j d(x-u, y-r)) O_{ij},$$

where O_{ij} is the original occurrence level of feature *j* at cell *i*. Cell *i* is located in (*u*,*r*) and *d*(*x*-*u*, *y*-*r*) is the distance between locations (*x*,*y*) and (*u*,*r*). The summation is over the landscape grid, and α_j is the parameter of the dispersal kernel for feature *j*. This is a two-dimension kernel smoothing using a radially symmetric negative exponential (dispersal) kernel.

Instructions to using distribution smoothing in Zonation are in section 5.1.3.

Literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences,* 272: 1885-1891.

Moilanen, A. and Wintle, B.A. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation*, 129: 427-434.

2.4.3 Boundary Quality Penalty (BQP)

This section is largely based on Moilanen and Wintle (2007).

The boundary quality penalty (BQP) is a quantitative feature-specific way of inducing aggregation in Zonation solutions. If using species as features, it can be seen as a way of approximating nonlinear effects of connectivity that may be present in species distribution models. There are very many different statistical species distribution modeling techniques (a.k.a. habitat models, resource selection functions, etc.). Typically, in such models the abundance of a species at a location is influenced by the local habitat quality, as well as the habitat in the neighborhood of the location. Such a neighborhood influence essentially states that the species is somehow dependent on connectivity or edge effects or both.

Ideally, reserve selection would be directly based on nonlinear species distribution models with neighborhood effects. However, this is not realistically possible, because it would make reserve selection computationally very slow. Also, implementing dozens of different modeling techniques efficiently in reserve selection software would be an enormous task. Herein enters the BQP.

The BQP is a mechanism for approximating the aggregate response of a species to edge effects, metapopulation size, and connectivity. It can be seen as a way of utilizing the connectivity response that is present for a species in a species distribution model. Essentially, the model is used for two things. It is used to predict an abundance or probability of occurrence in every cell in the landscape. This gives the standard input layer for one species. The species distribution model is also analyzed to determine the aggregate response of a species to habitat loss and fragmentation. This response is transferred to Zonation as a standardized curve that mediates the boundary quality effect in Zonation. See Moilanen and Wintle (2007) for the method to analyze a habitat model to obtain a species-specific response curve. Note also that the BQP responses for species can also be generated based on expert opinion; they do not always need to be fit statistically.

Different species can have different responses to fragmentation and habitat loss, which are entered into Zonation as two BQP components: (1) a species-specific radius and (2) a response curve (see figure below). The radius defines the distance from the focal cell that habitat loss will effect a species. The effect can be very localized (e.g., if the species is only sensitive to edge effects very close to the focal cell), or the effects can extend over a long distance. This could be the case with a timid larger animal that avoids human proximity. Habitat loss could influence such a species a long distance from the cell where it actually occurs. Note that inside the radius, only the loss of those cells that have data on the particular species (in other words cells that are not marked as missing data) can influence the value of the focal cell via connectivity.

The second component, the response curve, specifies the kind of effect neighborhood habitat loss has on a species. There are three general ways in which a species can respond to habitat fragmentation. First, habitat loss could have no effect on a species. This scenario would be modeled by a flat line. Second, the species could suffer variable degrees of loss in local population density if habitat is lost in the neighborhood. For example, you could have a relatively insensitive species, which loses half of the population density when the focal cell has lost all it neighbors from inside the species-specific buffer. Then again, the species could be very sensitive to neighborhood fragmentation, and all local value could be lost when only half of the neighboring cells have been lost. Third, a species could favor fragmentation, which would be modeled by a response curve that goes above 1 at some levels of habitat loss.

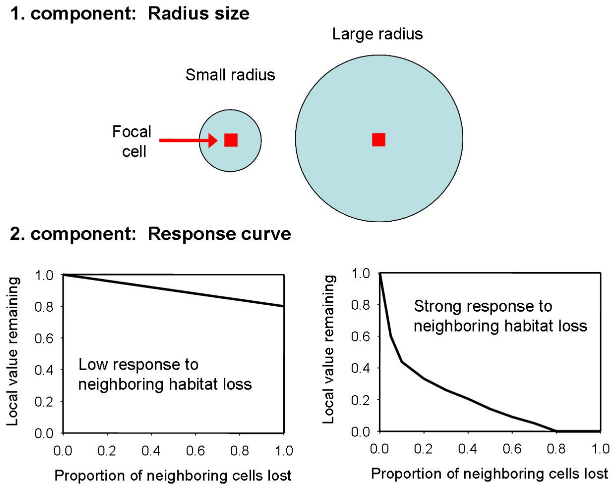


Figure clarifying the two components used when specifying the BQP for a species.

Note that there are major differences between distribution smoothing and the BQP even though both induce aggregation in a feature-specific manner. First, from a practical point of view, the BQP is much slower to run as the effect of removing a cell is not only local but extends over the neighborhood area which needs to be accounted for in computations. Second, the BQP is biologically better justified. The BQP definitions can be based directly on species responses in statistical habitat models. The difference between the methods would be most strikingly visible in fragmented areas. Distribution smoothing perceives the value of fragmented areas as relatively low. In comparison, the BQP could recognize a species that happily lives as a metapopulation in a fragmented environment. The response for that species would be such that it is recognized that the species can have high-value habitats in fragmented areas.

In the implementation of the BQP in Zonation, the value of a cell that is removed (marginal loss of conservation value) is now divided into two components: (1) local value, which is as before, and (2) loss of conservation value in the neighborhood of the focal cell, as modeled via the BQP specification. Thus, with BQP the effect of cell removal is not only the loss of the value in the cell itself but also a (potentially species-specific) reduction in quality in the neighborhood cells. The algorithm used by the BQP is:

$$\delta_i = \max_j \frac{w_j}{c_i}\left\{ H_j\left(\frac{n_{ij}}{n_{ij}^*}\right) Q_{ij}(S) + \sum_{k \in N_j(i)} Q_{kj}(S)\left[H_j\left(\frac{n_{kj}}{n_{kj}^*}\right) - H_j\left(\frac{n_{kj}-1}{n_{kj}^*}\right)\right]\right\},$$

where $N_j(i)$ indicates the cells (containing data for that species) within the species-specific radius of cell i for species j. $H_j(h)$ is the proportion of the original value of cell k remaining for species j when the focal cell has fraction h of its original neighbors remaining. The fraction of cells remaining is $h = n_{kj} / n_{kj}^*$, where n_{kj} is the number of neighbors remaining for cell k within the buffer radius of species j, and n_{kj}^* is the original number of neighbors. The loss term in the curly brackets consists of two terms, local loss and loss in the neighborhood of the focal cell. Local loss is the fraction remaining of the original value of the focal cell *i.* If many of its neighbors have already been lost, the value of $Q_{ij}(S)$ has been reduced. Loss in the neighborhood is mediated via the loss of one cell from the number of neighbors, which goes down from n_{kj} to $n_{kj} - 1$. Note that the formula above is employed as it is only for the core-area Zonation. For the additive benefit function and target-based planning the formula includes few trivial differences (see Moilanen 2007), but the concept behind BQP is the same in all cases.

The size of the neighborhood of a cell and the effects of habitat loss are defined separately for each species according to habitat models, which themselves mediate the boundary quality penalty. Because BQP ranks the cells based on the responses of species to fragmentation, species that benefit from the loss of surrounding habitats will be equally protected (compared to distribution smoothing).

Instructions to using BQP in Zonation can be found in section 5.1.4.

Literature

Moilanen, A., and B.A. Wintle. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

2.4.4 Directed connectivity (NQP)

This section is largely based on Moilanen, Leathwick, and Elith (2008).

The directed freshwater connectivity measure is a generalization of the BQP technique in which the concept of neighborhood is generalized, hence the name Neighborhood Quality Penalty (NQP). Instead of using a circular neighborhood, the NQP is defined using a tree-hierarchy of linked planning units. A focal area (planning unit) is influenced by negative action (habitat loss) potentially both “downstream” and “upstream” from the focal location, depending on the requirements of the feature.

The NQP technique was originally developed for freshwater planning in riverine systems, but the technique is suitable for other situations as well. The NQP method is based on a bidirectional linking of planning units. These linkages could correspond to hydrological flow, but they could also correspond to other kinds of biological linkages including (1) other spatially continuous connecting landscape elements, such as hedge rows; (2) spatially discontinuous, but functionally linked planning units, such as areas on migration routes of birds; or (3) they could approximate connectivity at marine areas where very strong flows generate a situation analogous to a river system. In the end, an appropriate aggregation of cells to planning units and suitable linkage and loss functions allow modeling of relatively variable situations.

Following Moilanen *et al*. (2008) closely, the present version of the NQP technique is technically specified by the following modification of the marginal loss value used in the cell removal rule:

$$\begin{aligned}
\delta_{ij} &= \delta_{ij}^{local} + \delta_{ij}^{neighborhood} = \delta_{ij}^{local} + \delta_{ij}^{upstream} + \delta^{downstream} \\
&= p_{ij} h_j^{up}\left(\frac{r_{ij}^{up}}{o_{ij}^{up}}\right) h_j^{down}\left(\frac{r_{ij}^{down}}{o_{ij}^{down}}\right) \\
&\quad + \sum_{k \in N_{ij}^{up}} p_{kj}\left[h_j^{up}\left(\frac{r_{kj}^{up}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down}}{o_{kj}^{down}}\right) - h_j^{up}\left(\frac{r_{kj}^{up}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down} - A_{ij}}{o_{kj}^{down}}\right)\right] \\
&\quad + \sum_{k \in N_{ij}^{down}} p_{kj}\left[h_j^{up}\left(\frac{r_{kj}^{up}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down}}{o_{kj}^{down}}\right) - h_j^{up}\left(\frac{r_{kj}^{up} - A_{ij}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down}}{o_{kj}^{down}}\right)\right].
\end{aligned}$$

in which p_{ij} is the occurrence level of feature *j* in cell *i*. The equation above describes the fraction lost from the original distribution of feature *j* following the removal of site *i*. The loss consists of three components: local loss, loss upriver, and loss downriver. The assumption is that everything remaining locally is lost if a cell is removed, and that loss accrued upriver and/or downriver will depend on the size of unit *i* for feature *j* (A_{ij}). Loss of a larger unit implies greater influence on connectivity nearby. The influence of connectivity on occurrence levels is mediated via functions h_j (upriver and downriver), which are response functions like those in BQP with the x-axis reversed. When the full landscape remains and nothing has been lost, h_j (1)=1. Importantly, when calculating marginal loss, the equation accounts for degradation that already has occurred. This implies that if no local value remains due to past neighborhood loss, further loss of connectivity has no local influence on the feature.

Connectivity in the equation above is modeled separately upriver and downriver. Quantities r_{ij} and o_{ij} (up and down) are the remaining and original connectivity, respectively, of unit *i* for feature *j* both upriver and downriver. Loss of planning unit *i* influences the downwards connectivity of sites upriver from it, which is the component of the equation having the summation across neighbors *k* upriver from focal site *i*, ($k \in N_{ij}^{up}$). Similarly, loss of unit *i* influences the upwards connectivity of units downriver from it, indicated by the summation across neighbors *k* downriver from the focal site *i*, ($k \in N_{ij}^{down}$). Note that the present version of Zonation uses feature and unit specific predictions of occurrence p_{ij} and feature-specific connectivity responses $h_j()$.

However, connectivity up and down river is based directly on the number of grid cells in planning units in a non-feature-specific manner, meaning that r_{ij}, o_{ij}, and A_{ij} are taken as the same for all features. This assumption might be relaxed in a later version of Zonation.

The equation above is simply the fraction of distribution loss for one feature, and it does not account for how lost representation is translated to loss in conservation value. When deciding which cell can be removed with the smallest loss of conservation value, δ_{ij} is aggregated across features according to the chosen cell removal rule.

Instructions for using directed connectivity in Zonation can be found in section 5.1.5.

Literature

Moilanen, A., Leathwick, J. and J. Elith. 2008. A method for freshwater conservation prioritization. Freshwater Biology, 53: 577-592. *Freshwater Biology*, 53: 577-592.

2.4.5 Matrix connectivity

This section follows Lehtomäki *et al*. (2009).

The rationale behind this technique is that multiple features can facilitate connectivity for each other. For example, Lehtomäki *et al*. (2009) applied matrix connectivity to account for the extent to which different types of forest habitats enhance connectivity of each other. The landscape can be divided into different forest types (e.g., spruce, pine, birch, etc.). While the forest types differ from each other, they will still help each other's connectivity when compared to non-forest habitats such as agricultural fields or water bodies.

Matrix connectivity is essentially a multi-feature extension of ecological interactions where the occurrence level of a feature in a cell is influenced by its connectivity to cells where interacting features occur weighted by the dispersal kernel of the feature (see section 2.6 for a detailed description of ecological interactions and section 2.4.2 for distribution smoothing). The connectivity-transformed occurrence level p' for feature k in cell i is calculated as

$$p'_{ik}=p_{ik}C_{ik}$$

where p_{ik} is the original occurrence level of feature k in cell i and C_{ik} is the multi-feature connectivity of cell i from the perspective of feature k. The important part here is the way C_{ik} is defined:

$$C_{ik} = \sum_{n=1}^{F} \{S_{nk} \sum_{j=1}^{J} p_{jn} \exp [\alpha_k d(i,j)]\}$$

where F is the number of all features, J is the total number of cells, p_{jn} denotes the occurrence level of feature n in cell j, $d(i,j)$ is the geographical distance between cells i and j, and α_k is the parameter giving the spatial scale for feature k. S_{nk} is a coefficient specifying how much feature n contributes to the connectivity of feature k. S_{nk} values are read in to Zonation from a connectivity matrix (section 3.3.3.4).

Instructions for running an analysis with matrix connectivity are in section 5.1.6.

Literature

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

2.4.6 Edge adjustment in connectivity

This section is largely based on Arponen *et al*. (2012).

This is an adjustment to feature-specific matrix connectivity values of cells at the edges of habitat. Connectivity decreases towards edges, but in some cases it may not make sense to consider edges to have lower priority. There are two typical cases where such an adjustment may be used:

1. Cells on the other side of the national border could be marked as non-harmful base habitat if it is known that habitat is much the same on both sides of the border. Thus, it is effectively assumed that habitat outside the border will influence connectivity as habitat inside the border.

2. At the edges of water bodies. For example, the connectivity of a forest will be reduced at the edge of a large lake. But, there may be many cases where such an edge-effect is not

desirable. If the edge effect is not desirable, mark both forest and water as suitable base habitat in the connectivity edge effect fix file. This option could also be relevant at the border of a forest and a marshland as some species will not perceive the marshland as bad for connectivity.

In practice, specific areas of non-habitat are ignored in matrix connectivity calculations. Connectivity values are corrected by the fraction of landscape around the focal cell from which connectivity is aggregated. If the neighborhood of the focal cell includes what can be considered harmless non-habitat (e.g., habitat that continues across a country border unchanged), then the connectivity value is corrected to c'=c/(1-f), where c is original connectivity and f is the fraction of harmless non-habitat. For example, consider a cell with a connectivity value of 2.0 that has been aggregated from a neighborhood that is only 1/3 within the country. If the cells outside the country have been marked as non-harmful base-habitat, the value of connectivity becomes 2.0/(1-2/3)=6.0.

It is also possible to correct for the proportion of habitat within each cell. When 50% of a cell is located outside a country border or when half of it is covered by water, its biodiversity value cannot be as high as it is for cells covered entirely by habitat. This can be done through the use of a cost layer, where cost is the proportion of habitat per cell. In practice, the size of the occurrence of a feature in a cell is divided by the proportion of habitat in the cell, transforming absolute habitat amount to habitat density. For example, if a cell's biodiversity value as such was 5 but only 50% of it actually contains habitat and the rest is covered by a lake, then this correction increases the cell's value to 10. The other way to think of it is that the cost of protecting the 50% of the cell that actually contains habitat should be 50% of land cost per cell in the region.

It is up to the user to decide when these edge related adjustments are desirable and when they are not. The latter correction considers "quality" of cells, but not actual land costs. Land cost, when known, should be included as a product with habitat proportion in a single cost layer. Please note that this feature is used only with matrix connectivity.

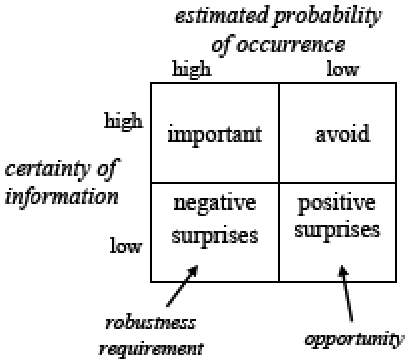
Literature

Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. and A. Moilanen. 2012. Effects of connectivity and spatial resolution of analyses on conservation prioritization across large extents. *Conservation Biology*, 26 (2): 294-304.

2.5 Uncertainty analysis

A common problem with conservation planning is the uncertainty of planning inputs. Mostly these uncertainties are due to lack of data since we simply do not have a comprehensive database with accurate information of the distribution of every biodiversity feature. Uncertainty can also arise, for example, from outdated or false observations, the use of predicted data (e.g., distribution models), or from any future factors such as the potential for anthropogenic land-use changes or climate change.

Taking into account both ecological value and uncertainty creates a prospect of four scenarios (see figure below). (1) Areas with high conservation value and high certainty of that information would be important for conservation. (2) Areas with low conservation value and high certainty (car parks, etc.) would ordinarily rank low among conservation priorities. (3) Areas with high estimated conservation value, but low certainty have potential for producing negative surprises for conservation. (4) Finally, areas with low conservation value and also low certainty have potential for producing positive surprises.



The goal of uncertainty analysis in reserve selection is to implement and evaluate trade-offs between biological quality and the certainty of that information. Ideally, one would identify a reserve network that guarantees high biological quality despite some uncertainty in input data. Uncertainty analysis could also be used for evaluating the opportunities arising from uncertainty, that is, the potential for positive surprises.

Here we introduce two methods of uncertainty analysis that can be used in Zonation. The first method, distribution discounting, enables the ranking of the landscape using species distribution data that includes uncertainties. The second, opportunity analysis, differs from the previous by giving high value to low uncertainty with the aim of maximizing positive surprises.

The simplest method to account for uncertainty is to influence feature weights

Before going on to more complicated methods, please do note this simple option. If information for a feature (species, other biodiversity feature, cost) is uncertain, then reduce the weight for the layer. This technique has been utilized in the following two publications with the idea that the future is less certain than the present and connectivity is less certain than distribution of habitat suitability.

Carroll, C., Moilanen, A., and J. Dunk. 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. *Global Change Biology*, 16: 891-904.

Kujala, H, Moilanen, A., Araújo, M.B., and M. Cabeza. 2013. Conservation planning with uncertain climate change projections. *PLoS One*, 8: e53315, doi:10.1371/journal.pone.0053315.

Literature

For more information about the aims and methods of uncertainty analyses in reserve selection, see:

Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Y. Ben-Haim. 2006. Planning for robust reserve networks using uncertainty analysis. *Ecological Modelling*, 199: 115-124.

Moilanen, A., B. A. Wintle, J. Elith and M. Burgman. 2006. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology*, 20: 1688-1697.

See also the following study for an analysis about uncertain effects of fragmentation. It proposes that following an uncertainty argument, some additional aggregation in the reserve network is probably an ecological net positive.

Moilanen, A. and B.A. Wintle. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation*, 129: 427-434.

2.5.1 Uncertainty in species distributions, distribution discounting

This section is largely based on Moilanen *et al*. (2006).

Distribution discounting is a method for including uncertainty analysis in a Zonation conservation prioritization. This method helps find the most robust solutions (i.e, those that most likely achieve a conservation goal given a level of uncertainty in species distributions). This analysis utilizes both the estimated occurrence levels of a feature in a cell and the certainty of that information.

Looking for robust reserve networks

In the framework of uncertainty analysis, one goal for reserve selection would be to find those network candidates that would achieve the given conservation targets despite uncertainty in input data. Thus, cells need to be ranked so that the highest priority is given to cells that have both relatively high conservation value and high certainty of information. In Zonation, uncertainty analysis has been implemented according to a convenient formulation that uses information-gap decision theory (see Ben-Haim 2006).

Conceptually, relevant components of the info-gap theory are:

1. The **nominal model**. This is the best set of predictions for the feature.

2. The **uncertainty model**. This states that even though with a nominal estimate, the true probabilities of occurrence are certain to deviate from the nominal model. The uncertainty model specifies a set of bounds that expand around the nominal estimate as a function of an uncertainty parameter, α.

3. A **performance function**. This is a function measuring the quality of an analysis (e.g., what proportion of feature distributions are covered by a given set of areas).

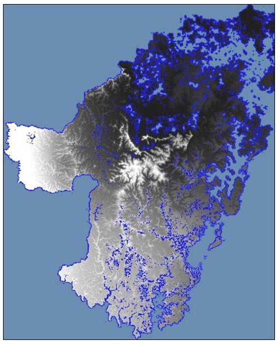
4. **Robustness function**. This function measures how much the uncertainty parameter α can deviate from the best available estimates (degree of uncertainty) so that conservation goals are still met even if the most adverse choice of occurrence levels from inside the uncertainty bounds are taken. A good reserve candidate is such that it achieves goals while allowing for high uncertainty (α). The *robust optimal reserve candidate* is the one that achieves conservation goals while allowing for the highest uncertainty.

At simplest, distribution discounting is implemented as follows:

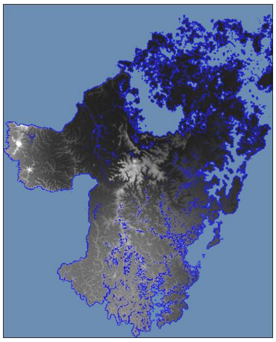
1. Take the nominal estimates (the normal input distributions for biodiversity features).

2. Develop a respective uncertainty map for each feature. The uncertainty layer could, for example, represent the standard deviation (SD) of the nominal estimate or the length of the lower half of the 95% confidence interval. This map could also be based on non-statistical uncertainty (e.g., an expert-based estimation on the likelihood of an area becoming damaged in the near future).

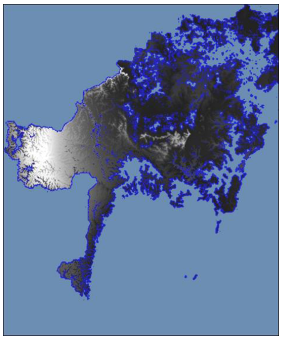
3. Specify that the degree of uncertainty, α. (i) If $\alpha = 0.0$, then analysis proceeds as if uncertainty analysis is not used. (ii) If $\alpha > 0$, robustness analysis is executed, with the aim of finding robust solution. If $\alpha = 0.5$ and the uncertainty model is 1xSD, then essentially one half SD of the nominal estimates would be subtracted, thus emphasizing locations with relatively certain predictions. Frequently values like 0.5 or 1 have been used for info-gap alpha, but the number could as well be something else. (iii) if $\alpha<0$, an opportunity analysis is done, favoring locations that could offer a better than expected outcome.



A. Distribution Model



B. Error Surface



C. Discounted Distribution

The pictures above demonstrate the concept of distribution discounting. Here picture A shows a modeled map of a species distribution with white areas representing a high probability of occurrence. The picture B displays an error surface (eg., standard deviation of the modeled values), with white color indicating large deviation and therefore high uncertainty. The α-value (horizon of uncertainty) can be used to either enhance or diminish the strength of the error surface (e.g., $\alpha = 2$ would double all error values in the map). Picture C is a species discounted distribution, where the error surface has been subtracted from the modeled distribution map. This is the map that Zonation finally uses to run the analysis.

An expanded explanation for distribution discounting.

The occurrence of feature j in cell i (here indicated as p^*_{ij}) is not certain, but merely the nominal "best guess" probability. Thus the true probability $p_{ij} \in [0,1]$ could be within an interval given by:

$$\left(p^*_{ij} - \alpha w_{ij}\right) \leq p_{ij} \leq \left(p^*_{ij} + \alpha w_{ij}\right) \qquad (1)$$

where α is the degree of uncertainty and w_{ij} is any error measure related to the accuracy of p^*_{ij} (for feature j in cell i).

Thus the true probability p_{ij} could be either higher or lower than the estimate p^*_{ij}, with bounds for p_{ij} determined by α and the relative error measure w_{ij}, which could be, for example, related to the accuracy of statistical prediction. The model of Eq. (1) is called a uniform bound model in info-gap terminology.

When using predictions based on logistic regression habitat models, a plausible model for uncertainty is to define the uncertainty interval in logit space, where w_{ij} is the standard error for the linear predictor of a logistic regression:

$$\left|\ \text{logit}\ \left(p_{ij}\right) - \text{logit}\ \left(p^*_{ij}\right)\right| < \alpha w_{ij} \qquad (2)$$

According to info-gap theory, one should favor reserve structures that achieve given conservation targets even with the most adverse choice of probabilities (in other words, in the worst case scenario). Given the present definitions, the most adverse choice of probabilities occurs when all probabilities are at their lower bounds (this is when the lowest expected number of populations is obtained). Assuming the analysis in logit space:

$$\text{logit}\ \left(p_{ij}\right) = \text{logit}\ \left(p^*_{ij}\right) - \alpha w_{ij} \qquad (3)$$

Thus, the program calculates the discounted biological value of a cell by reducing (discounting) the value of the logit of probability (p^*_{ij}) by a multiple of the error (αw_{ij}). In the distribution discounting technique, the original estimated occurrence data is simply replaced by the discounted data before proceeding to do the Zonation run. Thus, one Zonation run with discounted data is needed for each value of the degree of uncertainty, α. Note that Zonation does not care how the nominal estimates and associated error measures are obtained; any statistical method or expert evaluation can be used as a basis for developing those quantities.

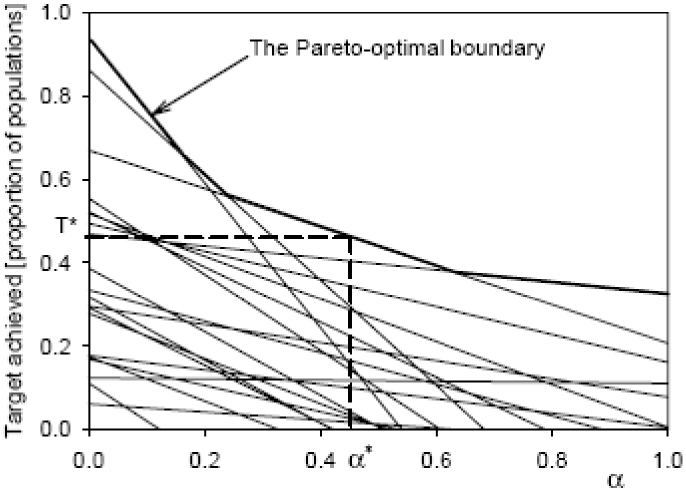
Determination of the degree of uncertainty, α

As mentioned above, the relative error w_{ij} can be any error measure related to the predicted species distribution. It can, for example, be a statistical error (e.g., the length of the lower half of the 95% confidence interval), a probability of future anthropogenic threat, or both. The value of α, on the other hand, is unknown and has no correct value. Several levels of α should be tested before selecting a reserve network. If an area is always selected irrespective of the value of α, then the area is definitely important for conservation. If the area is selected with low α but not with high α, then the area is selected because of the occurrence of an uncertain biological feature. If the area is not selected with low α but is selected with high α, then something of interest occurs in the area with relatively low density but high certainty. Based on such uncertainty analysis, one can split the landscape in three categories of areas: (1) areas that are good for sure, (2) areas that are never selected and thus have low priority and (3) areas that are selected with some levels of uncertainty. The areas in the third category may need further investigation before a choice of conservation priority can be fixed.

Distribution discounting uses the following algorithm for finding robust-optimal reserve designs:

1. Specify robustness requirement (α).
2. Distribution discounting. Read in biodiversity feature information. For every feature and cell, apply Eq. (3) or its analogue, depending on type of data/uncertainty model.
3. Use any reserve selection algorithm (here Zonation) to search over spatial patterns. The robust optimal design **X*** (at the given level of α) is the one achieving the highest possible conservation value.

The advantage of this approach is that the worst-case probability set (Eq. 3) has to be calculated only once (item 2), and thereafter the contributions of cell to representation levels, p_{ij}, do not change in the reserve selection process. Testing several α-values outlines how different reserve structures behave in increasing uncertainty. Some designs are always bad while some are good according to nominal habitat model predictions but bad if uncertainty is incorporated into the models. Others have intermediate nominal performance, but have a good robustness to uncertainty. The robust-optimal designs are always at the **Pareto-optimal boundary** with respect to the target, as demonstrated below:



In the figure, each thin line represents one solution (spatial reserve structure). An increasing robustness requirement (*α*) implies that a decreasing biological value can be achieved reliably. The thick line is the Pareto-optimal boundary representing solutions that are optimal in the sense that increased biological value can only be obtained with the cost of lowered robustness and vice versa. When doing reserve selection on a large grid, there is a huge number of potential reserve structures, but only one or few of them would correspond to the Pareto-optimal boundary at any given *α*-level and resource (fraction of landscape). Any solution not at the Pareto-optimal boundary is inferior in the sense that another solution exists with either higher biological value or higher robustness or both. The distribution discounting technique used inside Zonation automatically identifies the robust-optimal nested Zonation set of solutions for the given level of α.

Instructions for using distribution discounting in Zonation can be found in section 5.1.8.

Literature

For more information about distribution discounting, see:

Moilanen, A., B. A. Wintle, J. Elith, and M. Burgman. 2006. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology*, 20: 1688-1697.

For more information about information-gap theory in general, see:

Ben-Haim, Y. 2006. *Info-gap decision theory: Decisions under severe uncertainty*, 2nd edition. Elsevier Academic Press, London.

2.5.2 Positive uncertainty in species distributions - opportunity analysis

The distribution discounting uncertainty analysis reduces feature-specific value at locations according to the level of uncertainty there. The approach applies information-gap theory. The level of discounting is specified by the parameter α, which in Zonation v.2.0 was constrained to be positive. As a multiple of α is subtracted from the feature-specific occurrence levels, positive values of α mean that local occurrence levels are always reduced with growing uncertainty.

In Zonation v.3.1 and higher, α can also be assigned negative values. Negative α values can be utilized for an opportunity analysis (i.e., a search for areas where a better-than-expected outcome could be obtained). The opportunity variant of uncertainty analysis prefers an area with higher uncertainty to an area with low uncertainty if the nominal predicted level for the feature (probability of species occurrence) is the same. In other words, increased uncertainty would in this case inflate the distribution instead of reducing (discounting) it. The technical implementation of an opportunity analysis is identical to that of distribution discounting, the only operational difference being that negative values are assigned to α.

Opportunity analysis might be appropriate when one wishes to gain the most from very limited resources. For example, if an area has low feature occurrence estimates because it is poorly surveyed, then opportunity analysis might reveal that the area could have more value than what is evident based on the limited data. See Ben-Haim (2006) for examples of opportunity analysis.

Literature

Ben-Haim, Y. 2006. *Info-gap decision theory: Decisions under severe uncertainty*, 2nd edition. Elsevier Academic Press, London.

2.6 Ecological interactions

This section utilizes methods first described in Rayfield *et al*. (2009), but the method has since been applied in various ways. See, for example, Lehtomäki *et al*. (2009) and Carroll *et al*. (2010).

In many cases, conservation planning is done purely based on occurrence data, either of species or other biological features such as vegetation types. However, there are cases in which it is desirable to include interactions between distributions of features in the analysis. With Zonation, this type of specific planning is possible as the software includes a facility for modeling a variety of interactions. In this section we describe the method and philosophy behind this facility and provide some examples of cases where this type of analysis would be useful.

Possible analyses with pairwise species (or feature) interactions

Analysis variants that can be done doing using the interactions facility include:

1. Modeling of predator-prey, resource-consumer, and host-parasitoid interactions. In these analyses, the objective would be to ensure protection of those parts of the resource distribution that are “close enough” to be utilized by the consumer. This corresponds to interaction variant 1 (below), in which the connectivity of the resource to consumer is included in analysis. See Rayfield *et al*. 2009.

2. Emphasizing proximity to existing reserve areas when planning expansion of a reserve network. See Lehtomäki *et al*. 2009.

3. Application of Zonation to alleviating conservation consequences of climate change. In this analysis, the connectivity of the predicted future distribution of the species to the present distribution is used. At least three maps would be entered for each species: present distribution, future predicted distribution, and connectivity from future to present (~stepping stones). Could also include connectivity from present to future (dispersal sources). See Carroll *et al*. 2010.

4. Avoidance of invading species or sources of pollution. Interaction variant 2 (below) can be used to discount a distribution in the proximity to a known or potential source of an invading species. In effect, occurrence levels of the target species will be reduced at locations that are close to (well connected to) the distribution of the invasive species (or source of pollution).

5. Modeling of food chains or food webs. The interactions between multiple species can be calculated, multiple resources can be connected to one consumer, etc. By appropriately chaining connectivity effects between distributions, it should be possible to model more complicated relationships than just simple pair-wise interactions.

(1) Interaction type 1: Resource-consumer interactions

Interaction type 1 refers to positive interactions. This type of interaction can be used to account, e.g., for plant-herbivore, predator-prey, host-parasitoid, or present to future interactions.

The general idea is that one wishes to (1) protect the resource independently, (2) protect a part of the resource distribution which is available to the consumer, and (3) emphasize protection of the consumer at areas that are within foraging distances from the resource. Item (1) can be achieved simply by entering the resource as an independent layer in the Zonation analysis.

More mathematically, the local abundances of resource *j* and consumer *k* are indicated by r_{ij} (in grid cell *i*) and c_{nk} (in grid cell *n*), respectively. Let β_k be the parameter modeling the spatial scale of foraging for consumer *k*. β_k is the parameter of a negative exponential function. We specify that the resource use intensity of resource *j* at cell *i* by consumer *k* is R_{ij}:

$$R_{ij} = r_{ij}\max\left\{1.0, \frac{S_{ijk}}{\gamma_j \max_i S_{ijk}}\right\} = r_{ij}\max\left\{1.0, \frac{\sum_{n=1}^{N}\exp(-\beta_k d_{in})c_{nk}}{\gamma_j \max_i \sum_{n=1}^{N}\exp(-\beta_k d_{in})c_{nk}}\right\}, \qquad (1)$$

which is the local resource density multiplied by the connectivity of the cell to the consumer population, S_{ijk}, using parameter β_k to model the foraging distances of the consumer (d_{in} is the distance between cells *i* and *n*). Thus locations with high R_{ij} have both an abundance of a resource, and that resource is within the foraging distance to a relatively high number of consumers. Eq. (1) is the connectivity of the resource to the distribution of the consumer. $\gamma_j \in [0, 1]$ is a feature-specific (often species-specific) parameter describing how fast resource use is saturated. If γ_j=1, R_{ij} scales linearly with connectivity between distributions, S_{ijk}. If, for example, γ_j=0.1, R_{ij} achieves its maximum value when connectivity is 10% of the maximum it gets anywhere in the landscape, and so on.

Concerning parameterization, in the two-dimensional case half of the foraging would be contained within a distance of $2/\beta_k$ from the focal cell. Knowledge of foraging distances thus allows setting a reasonable estimate for β_k. Note that the size and unit of the grid cell needs to be accounted for when calculating β_k, see section 3.3.3.8. Essentially, β_k is calculated identical to the scale of landscape use for the distribution smoothing technique.

(2) Interaction type 2: Negative interactions

Interaction type 2 refers to interactions that have negative consequences for the target layer of the connectivity interaction. The general idea is that one wishes to de-emphasize those parts of the distribution of feature (species) A, that are close (well connected to) the distribution of B. Feature B could be, for example, a competitor, a potential source of an invading species, or a source of pollution (or any other pressure or threat) that may cause future degradation of habitat quality and consequent reductions in the population sizes of biodiversity features (species) of conservation interest.

Using the notation above, we now specify that the discounted value of feature *j* at cell *i* is R_{ij}:

$$R_{ij} = r_{ij}\left\{1.0 - \max\left[1.0, \frac{S_{iju}}{\gamma_j \max\limits_i S_{iju}}\right]\right\} = r_{ij}\left\{1.0 - \max\left[1.0, \frac{\sum\limits_{n=1}^{N}\exp(-\beta_k d_{in})u_{nk}}{\gamma_j \max\limits_i \sum\limits_{n=1}^{N}\exp(-\beta_k d_{in})u_{nk}}\right]\right\}, \qquad (3)$$

which is the local density of the focal feature r_{ij} discounted by connectivity to the undesirable feature to be avoided *u*, using parameter β_k to model the distances to which the undesirable influence spreads. Effectively, R_{ij} is the distribution of feature *j* reduced by connectivity *to* the distribution of *u*. In equation (3), the nominator inside the brackets is the connectivity of the focal cell to the distribution *u*. The denominator is the maximum connectivity any cell has to the distribution of *u*.

Thus, assuming γ_j=1, the fraction term scales from zero (for unconnected locations) to one (for a maximally connected location). If γ_j<1, then the (here negative) effects of connectivity saturate with a lower level of connectivity, as in Eq (1).

When Zonation is started, the biodiversity feature and uncertainty layers are read in one by one. For each layer distribution discounting is done first and then distribution smoothing (if used). Interactions are implemented after all biodiversity features layers have been read and discounted/smoothed. The interactions are read in one by one and performed between the layers specified in the interaction definitions file. Each interaction transforms one of the feature layers that were read in. This means that if, for example, a habitat quality layer, a connectivity layer, and an interaction layer are to be calculated and used based on the distribution map of one biodiversity feature, then the distribution map needs to be entered in Zonation three separate times: once as a plain unsmoothed habitat quality layer, a second time smoothed to implement the connectivity calculation, and a third time to be used in the interaction as a focal layer. Note also that a feature map may be transformed by more than one interaction. If the same layer is the focal layer of an interaction multiple times, then the layer models simultaneous connectivity to multiple (different) sources.

After interactions have been implemented, Zonation proceeds as before. Connectivity methods BQP and BLP operate as before and on interaction layers also. Instructions for including ecological interactions in Zonation can be found in section 5.1.9.

Literature

The method to include considerations of ecological interactions was first described by:

Rayfield, B., Moilanen, A. and M.-J. Fortin. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

See also:

Carroll, C., Moilanen, A., and J. Dunk. 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. *Global Change Biology*, 16: 891-904.

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

2.7 Replacement cost analysis

Very seldom can reserve selection start with a planning region that has no existing reserves nor any restrictions for conservation (such as areas ear-marked for residential building). More commonly, due to logistic or social constraints, certain sites need to be included or excluded in the final solution. In most cases this leads to a suboptimal network either in terms of conservation value or in terms of the economic cost of achieving a given conservation goal (compared to the one that could have been obtained without land-use restrictions).

It is useful to be able to asses the degree of suboptimality of solutions compared to the optimal ones. A method called *replacement cost analysis* can be used to evaluate the effects of forced site inclusion/exclusion, or in a more general case, some other constraint on the solution. *Replacement cost* refers to the **loss in the value of the solution** given that the optimal cost-efficient solution cannot be had and that alternative solutions, with additional constraints (e.g., with particular sites forcibly included or excluded), must be accepted. It tells us the cost (biological or economic) of including or excluding a site from the reserve network. If budget is constant, the *exclusion cost* of a site (or set of sites) is the loss in the network's conservation value that follows when sites that would belong to the optimal solution have to be excluded from the reserve network. The *inclusion cost* of a site is the loss in conservation value that must be accepted if a suboptimal site is forced into the reserve network. On the other hand, when the conservation budget is not fixed, replacement cost can also be defined in terms of the extra funding required to maintain conservation value that is equal to the value of the optimal solution. In Zonation, replacement cost can be determined from the difference between performance curves for the unconstrained and constrained solution. The four variants of replacement cost can be defined as:

1. **Biological exclusion cost:** Decrease in conservation value following forcible exclusion of a given (group of) site(s).

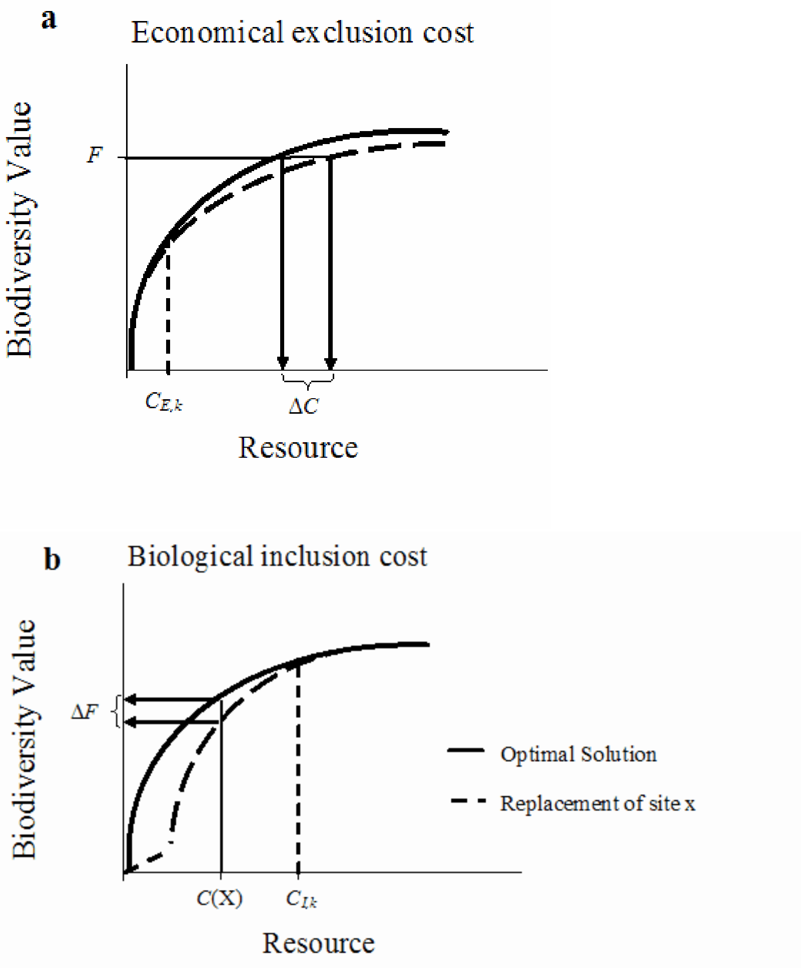
2. **Biological inclusion cost:** Decrease in conservation value following forcible (suboptimal) inclusion of a given (group of) site(s).

3. **Economic exclusion cost:** Increase in solution cost required to keep the same total conservation value following the forcible exclusion of a given (group of) site(s).

4. **Economic inclusion cost:** Increase in solution cost required to maintain conservation value following the forcible inclusion of a given (group of) site(s).

For practical purposes, the replacement cost is calculated in the following manner. First, find optimal reserve selection (**X***) which has highest possible value ($F(\mathbf{X^*})$) obtainable with the available resource (C_{max}). Then run the analysis again with some areas forced in or out. Replacement cost is the difference between the value of the optimal solution and the value of the new solution. We emphasize that **this does not mean keeping the optimal set of sites plus/minus a particular site**, **but finding a completely new solution given that the particular site(s) are forcibly included/excluded.**

A replacement cost of zero tells us that there exists an alternative solution with the same value as the current (best) solution (i.e., the same cost and same conservation value although obtained via a different selection of areas compared to the original optimal selection). A replacement cost larger than zero means that any alternative solution including/excluding the focal site(s) will have either a lower conservation value or a higher economic cost than the optimal one.



Above is a conceptual illustration of the *replacement cost* of a hypothetical site (a) in terms of increase in resources required to maintain value (ΔC) and (b) in terms of loss of biodiversity value (ΔF) (from Cabeza and Moilanen 2006). (a) *Exclusion cost*: the dashed line indicates the value of the best solution when site *x* is forcibly excluded. Up to a certain resource level, $C_{E,k}$, site *x* does not belong to the optimal solution and thus exclusion cost is zero. Even with $C > C_{E,k}$, exclusion cost can be zero if the site is fully exchangeable with another site or a combination of other sites. (b) *Inclusion cost*: the dashed line indicates the value of the best solution when a site is forced to be included in the solution. Inclusion cost is likely to be highest when resource availability is low and the forced inclusion of the unwanted site prevents the acquisition of other biologically much more valuable locations. At $C > C_{I,k}$, the focal site becomes included in the optimal set and inclusion cost becomes zero. (Note that the resource here should be understood as the proportion of landscape retained in the Zonation analysis.)

The figure also shows an example of how exclusion and inclusion can be expected to behave qualitatively. With a small resource ($< C_{E,k}$), the exclusion cost (a) of a site is likely to be zero, because the site would not be in the optimal set in any case. At level $C_{E,k}$ the site becomes part of the optimal solution. With a resource slightly higher than $C_{E,k}$, it is likely that the exclusion of the site can be compensated with small cost, at least if there are many selection units. However, when the available resource is large, sites of less importance are included in the solution and the exclusion of a high-quality focal site has a clearly positive cost.

Inclusion cost (b) behaves differently. When a site that would not belong to the optimal solution is included in the network it generates an increase of cost even when the resource available is small because the resource is spent on suboptimal areas. With increasing resource availability, the inclusion cost gradually decreases. At a level $C_{I,k}$, the site would already become part of the optimal solution and inclusion cost thus becomes zero.

Instructions for including and excluding areas in a Zonation solution for analyzing replacement cost can be found in section 5.2.4.

Literature

Replacement cost analysis is described in:

Cabeza, M. and A. Moilanen. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation*, 132: 336-342.

See also:

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

Moilanen, A., Arponen, A., Stockland, J.N. & M. Cabeza. 2009. Assessing replacement cost of conservation areas: How does habitat loss influence priorities? *Biological Conservation*, 142: 575-585.

2.8 Community level analysis

This section follows the manuscript in which community-level analysis in Zonation was first described:

Leathwick, J.R., Moilanen, A., Ferrier, S., and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

As legislation and politically set conservation targets often consider habitat or community types, it is often reasonable to consider these when selecting conservation priorities. Considering representation of community types instead of individual biodiversity features can also be suitable in situations where information about individual feature distributions is too sparse for fitting reliable habitat suitability models. The word community is used here, but for practical purposes the question is of any multi-feature ecological entities that are partially overlapping. This includes features such as *ecosystems*, *habitat types*, *vegetation types or classes*, etc.

Zonation utilizes three components during community level analysis: (1) raster layers that define distributions of community types , (2) a similarity matrix that describes pair-wise similarities in biodiversity feature composition, and (3) information about the feature richness of each community class. Feature richness is included in the weight given to the community class (or habitat type); see Leathwick *et al*. (2010) for details. Typically, this analysis would utilize a (statistically-based) community type classification, in which the effective number of community classes is reduced to tens or hundreds.

The pair-wise similarities can be obtained in a number of ways. The pairwise similarities can be determined by expert opinion. Alternatively, an ecologically reasonable method for quantifying similarities in feature composition is generalized dissimilarity modeling (GDM; Ferrier 2002). A GDM models turnover in biodiversity feature composition as a function of environmental variables. In a way, it calibrates environmental gradients to match the turnover rate in feature composition at different segments of the gradient. A GDM can then be used to predict compositional dissimilarities between two sites, for which the relevant environmental variables are known (see Ferrier *et al*. 2007).

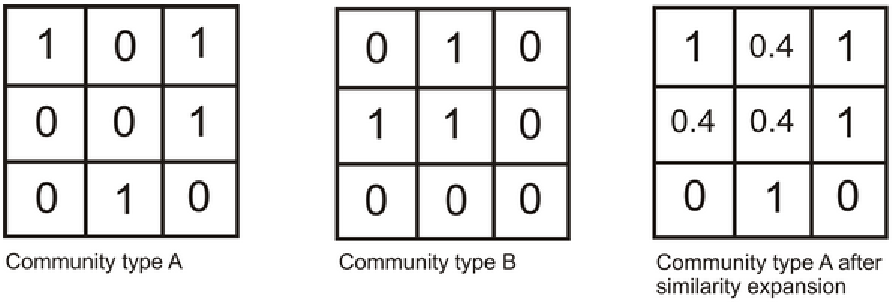
Zonation uses the similarity matrix either in (i) connectivity calculation (preferred) or (ii) to expand the primary habitat type classifications across the landscape to effective occurrences, which also account for similarity between habitat types. The fraction of a community type j protected under management actions a (in this context a would most often be reserving areas for conservation) is defined here as:

$$f_j(a) = \frac{\text{protected}_{\text{j}}}{\text{total}_{\text{j}}} = \frac{\sum_{\text{k=1…n}} \sum_{\text{i=1…m}} L_{ik} S_{ij} C_{ik}(a_k)}{\sum_{\text{k=1…n}} \sum_{\text{i=1…m}} L_{ik} S_{ij}}$$

where L_{ik} describes the extent of classification group *i* in cell *k*, S_{ij} describes the modeled biological similarity between groups *i* and *j*, and C_{ik} describes the landscape condition (see section 2.10) for group *i* in cell *k*. The effective occurrence level for group *i* is the sum of the occurrence level x similarity of all other groups present in cell *k*.

If, for example, community type A has 40% of its features in common with community type B, cells where type B is present (with occurrence level of 1) is considered to represent community type A with an occurrence level of 0.4, as illustrated in the figure below.

The original community type data do not need to be binary presence/absence, but they can reflect the proportional cover of the community type in the grid cell.



1	0	1
0	0	1
0	1	0

Community type A

0	1	0
1	1	0
0	0	0

Community type B

1	0.4	1
0.4	0.4	1
0	1	0

Community type A after similarity expansion

A figure illustrating the similarity expansion. Here, 40% of species characteristic to community type A are also present in community type B.

As Zonation prioritization takes place only after the similarity expansion, it is important to understand that representation of community types in the Zonation solution does not exactly correspond to the original community classes, but the expanded classes. A representation level of 0.4 of community type A in a given top fraction does not mean that 40% of the original community type A is covered in the solution, but rather that a fraction, 0.4, of effective occurrences are covered. It is, however, possible to return to the original unexpanded classifications and check their coverage in any selected area of a solution. It is also possible to enter the communities twice into the same analysis, once with the similarity expansion and another time without.

Note that community level analysis would typically be combined with at least condition information and possibly retention. See the also matrix connectivity computations for connectivity that is applicable to a set of partially overlapping habitat types (section 2.4.5; Lehtomäki *et al*. 2009).

Instructions for using community level biodiversity features are in section 5.3.1.

Literature

Community level analysis has been applied by:

Leathwick, J.R., Moilanen, A., Ferrier, S., and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

Note: Leathwick *et al*. did not use similarity expansion of Zonation v.3.1, but pre-processed data using GIS and R. Following their protocol, it is possible to use the technique also with Zonation v.2.0.

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

For more information about generalized dissimilarity modeling, see:

Ferrier, S. 2002. Mapping spatial pattern in regional conservation planning: where to go from here? *Systematic Biology,* 51: 331-363.

Ferrier, S., Manion, G., Elith, J. and Richardson, K. 2007. Using generalized dissimilarity modeling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions,* 13: 252-264.

2.9 Alternative land uses - multi-criterion analysis

The primary function of Zonation is to identify priority areas for conservation. However, conservation is not independent of other land uses, and it is useful to take multiple land use criteria into account during the conservation planning process. The ability to balance between biodiversity and competing land uses is relevant when considerations of cost must influence decisions or when compromises between conflicting interests need to be sought. Previous versions of Zonation could be adjusted to mask out areas that cannot be considered for conservation, such as heavily constructed urban areas, power plants, industrial harbors, etc. It was also possible to incorporate variation in land cost in the analysis. Zonation v.3.1 and higher goes a step further and allows balancing of multiple, possible antagonistic, land use criteria in prioritization. The idea is to find a spatial separation between positive biodiversity features and negative alternative land uses. Biodiversity should be retained in the top fraction of the priority ranking whereas areas suitable for alternative (negatively weighted) land uses should receive low priority ranks.

Technically, any analysis with more than one feature optimized is multi-criterion analysis, even when there are only two biodiversity features. Here, however, the term multi-criterion analysis is reserved for the case that there are biodiversity considerations and opposing land uses, opportunity costs, etc.

Including competing land uses in Zonation is remarkably simple. Competing land uses are fed into Zonation in a similar way as distribution layers of species or other biodiversity features. The difference is that the competing land use features are assigned negative weights (sections 3.3.2.2 and 5.1.1) instead of positive ones. This will enable Zonation to give high conservation priorities to sites that have high conservation value but are not very favorable for alternative land uses. Using this method ensures that areas hosting features with negative weights are removed early on in the prioritization while biodiversity is retained to the top ranks of the landscape.

The structure of conservation value in Zonation now becomes:
(biodiversity - competing land uses) / cost, where cost should be taken as the direct cost to the conservation agency.

There is no rule as to how large negative weights should be relative to the positive weights for beneficial biodiversity features. If it is critical that almost all of the negative features remain outside conservation areas, then the negative weights should be relatively large. A small negative weight is appropriate for mildly detrimental features. Experimenting with the relative weights will help to find the best balance between positive and negative features that answers planning needs.

The difference between forcing cells to be removed first with a removal mask (section 3.3.3.9) and multi-criterion analysis is that the removal mask will discard the areas completely, whereas negative weighting of competing land use features allows the sites to still be taken into account during the planning process. The multi-criterion approach is more appropriate when suitability of sites for competing land uses is varied and identifying solutions to account for several purposes are needed. Excluding areas completely at the starting point with the removal mask is more appropriate if it is already known that some sites will be used in ways that do not support biodiversity.

Please note: While it is useful to take multiple criteria into consideration in the conservation planning process, the primary focus in conservation prioritization should be maximizing biodiversity value. If too much attention is given to cost and other secondary matters, then they may end up being the main drivers of conservation decisions. This would reset systematic conservation planning back to its starting point, in which protected area networks were not representative precisely because unexploitable or otherwise least costly areas were designated for conservation (for discussion see Arponen *et al*. 2010).

Instructions for including considerations of multiple land uses are in section 5.3.4.

Literature

Multi-criterion analysis using Zonation has first been applied in:

Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J. and C.D. Thomas. Balancing alternative land uses with the Zonation conservation prioritization approach. *Ecological Applications*, 21: 1419-1426.

A recent application:
Di Minin, E., MacMillan, D.C., Goodman, P.S., Escott, B., Slotow, R., and A. Moilanen. 2013. Businesses and conservation planning in a biological diversity hotspot. *Conservation Biology*, 27: 808-820.

For discussion about cost considerations in conservation planning, see:
Arponen, A., Cabeza, M., Eklund, J., Kujala, H. and J. Lehtomäki. 2010. Costs of integrating economics and conservation planning. *Conservation Biology,* 24: 1198–1204.

2.10 Landscape condition and retention

Landscape condition and retention are integral features of ecosystem and community level analysis (section 2.8).

Each biodiversity feature can be linked to a condition and retention group. This means that certain types of habitats and groups of species have specific responses to prior and expected habitat loss. These responses need not be, and generally would not be, feature-specific. Rather, the responses would be specific to (user-chosen) groups of features, such as community types or guilds of species.

(1) Condition

In Zonation, condition represents information about local habitat deterioration and its influence on biodiversity features or groups of features. Condition is feature- (group) and grid cell-specific. Typically, there are several condition map layers that can be linked to different groups of features.

The value of landscape condition can vary between 0 (all local conservation value has been lost, for example in an urbanized area) and 1 (all habitat remains locally in a pristine state when compared to some historical baseline). In analysis, condition normalizes present landscape condition to be measured against a historical baseline. The idea is that remaining occurrences for features that have lost a lot are relatively more valuable than occurrences of features that have lost little. For example, habitats (and their species) occurring on fertile low-land soils may have suffered from significant losses. Low-productivity habitats occurring at high elevations and rough terrains would have suffered little loss. Use of condition allows one to model this past loss in a feature-specific manner.

(2) Retention

This section closely follows the manuscript in which the use of combined condition and retention in spatial prioritization has first been described:

Moilanen, A., Leathwick, J.R. and J.M. Quinn. 2011. Spatial prioritization of conservation management. *Conservation Letters*, 4: 383-393.

The conventional systematic conservation planning process aims at maximizing representation of biodiversity features in a network of protected areas. It carries the implicit assumption that areas outside the network contribute nothing to the overall biodiversity value of the region. This may be a realistic assumption if pressure for land conversion is high. However, conditions are often such that habitat is retained in the landscape even when some areas are not protected. This may be the situation with areas that are remote and difficult to access or with areas that are not economically exploitable. In such situations, the non-protected areas will also support biodiversity, and this could (and should) be accounted for in the conservation planning process. Zonation v.3.1 and higher can account for the occurrences of features that are retained across the whole landscape. There are two ways in which features can be retained: (1) retention implies that not all occurrences are lost from a location even without conservation action, or (2) management intervention increases habitat quality (occurrence levels of features) at retention sites.

Most often one would like to account for both representation in the protected area network and retention across the entire landscape. Use of only representation can hamper effective conservation, as some biodiversity features are likely to be retained also in non-protected areas. Then again, relying only on retention is risky, as high uncertainty is inherent in assumptions about threat and loss of habitat. To consider both representation and retention, Zonation needs two sets of biodiversity feature layers (in practice, this means that the set of biodiversity features used for the analysis must be listed twice). For one set, representation, it is assumed that all conservation value is lost if the site is not protected. The other set of layers contain the difference between what is expected to remain if the site is protected and what will remain even in the absence of conservation action. The relative weighting of these sets of layers will determine how much emphasis is placed on representation and how much on retention. A high weight on retention layers means that emphasis is placed on locations where conservation action can make a difference. A high weight on representation means emphasis is on locations where rarity and richness of features is highest based on known data.

Operation of Zonation when using condition and retention

When Zonation reads in raster files, the following steps occur:

1. Absolute occurrence levels for features *j* in cells *i* are read in. Call these *Oij*.

2. *Oij* is normalized to unity (1.0) by division with sum of *Oij* across cells *i*. Call this normalized occurrence *Nij*; with all the normalized occurrencies summing up to 1, *ΣNij*=1.0. Also at this stage, connectivity transforms such as distribution smoothing and matrix connectivity are done for selected layers, after which the distributions are re-normalized.

3. Apply condition to each cell for each relevant feature: *Nij*(condition)= *Nij* * *Cij*, where *Cij* is the fractional quality read in from the respective condition group.

This is a condition transformed distribution for the feature, and it no longer sums to unity. So far everything is straightforward, *but in the following steps it is relevant to pay attention to the meaning of retention matrices at input*.

4. Retention is applied on top of condition. It is applied to calculate the difference made with and without conservation, and retention ranking is based on the difference between what happens with and without conservation. At input, retention for a cell, *Rij*, is first transformed to fractional DIFFERENCE MADE with respect to *Nij*(condition). We state that the fractional difference *Dij*=abs(1- *Rij*), where the absolute value returns a positive difference independent of whether we use mode 1 (*Rij*<1) or mode 2 (*Rij*>1). A retention value of 1.0 implies no change in condition even in the absence of conservation action.

5. For stop loss (retention mode 1), the assumption is that condition goes *down* by the difference in the absence of action.

6. For management gain (retention mode 2), the assumption is that condition goes *up* by the difference in the absence of action.

7. The difference between stop loss and management gain manifests in that the weight for the retention transformed layer is calculated in a different manner, as described next.

8. Computation of weights for retention-transformed features:

The set of features used for representation can already have variable weights arising from following a feature weighting scheme or because different communities inherently differ in feature richness. These weights are denoted by w_j for community type *j*. The weights given for retention-transformed layers are denoted with w^r_j. Do not assume that $w^r_j = w_j$, as different community classes will likely lose different fractions of their remaining distribution. A high weight should go to a feature that would lose a large part of its remaining distribution in the absence of conservation action. When the expected relative loss for the feature would be small, the weight for the respective retention-transformed layer should also be small. Values for w^r_j come out different for the two analysis modes:

8a. Feature weight calculation for retention Mode 1, Stop Loss:

The weight calculation is linked to feature-specific occurrence levels, condition, and retention for each grid cell in the landscape. To correct for variable expected loss in the absence of conservation, we set:

$$w^r_j = \beta \frac{difference_j}{remains_j} w_j,$$

where parameter β tunes the balance between representation and retention in the analysis. The larger the value of β, the more emphasis is given to retention. Both $difference_j$ and $remains_j$ are measured for feature *j* across the full landscape. $Remains_j$ is a fraction of occurrences remaining for the feature after application of condition:

$$remains_j = \sum_i c_{ij} N_{ij},$$

in which N_{ij} is normalized pristine state (or other past baseline) occurrence level of feature *j* at grid cell *i*, and c_{ij} is condition of feature *j* in grid cell *i*. $Difference_j$ is the fraction of $remains_j$ that would be lost in the absence of conservation:

$$difference_j = \sum_i abs(1 - R_{ij}) c_{ij} N_{ij},$$

where R_{ij} is the respective retention value.

8b. Feature weight calculation for retention Mode 2, Management Gain:

Using the notation above, we now set:

$$w_j^r = \beta \frac{difference_j}{difference_j + remains_j} w_j,$$

where the idea is that with application of management intervention there is a gain, and the relative loss is from better than present condition to present condition.

Instructions for using landscape condition and retention in your analysis are in sections 3.3.3.14, 3.3.3.15 and 5.3.5.

Literature

Landscape condition in Zonation has been applied by:

Leathwick, J.R., Moilanen, A., Ferrier, S. and K. Julian. 2010. Community-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143: 984-991.

For discussion on retention, see:

Moilanen, A. and M. Cabeza. 2007. Accounting for habitat loss rates in sequential reserve selection: simple methods for large problems. *Biological Conservation*, 136: 470-482.

Pressey, R.L., Watts, M.E. and T.W. Barrett. 2004. Is maximizing protection the same as minimizing loss? Efficiency and retention as alternative measures of the effectiveness of proposed reserves. *Ecology Letters,* 7: 1035-1046.

2.11 Landscape dynamics

Zonation can emulate landscape dynamics when habitat suitability maps for biodiversity features are entered for the present and for one or several time steps in the future. This analysis produces solutions that are balanced across all time steps. At least with large landscapes, this approach to forcing landscape dynamics is really only convenient with the increased memory capacity of Zonation v.3.1. Connectivity between time steps in this analysis could be modeled by some connectivity method which accounts for multiple distributions simultaneously. In Zonation, this would be interaction connectivity (section 2.6) or matrix connectivity (section 2.4.5).

2.12 Administrative units

This section closely follows the manuscript in which the method was first described:

Moilanen, A. and A. Arponen. 2011. Administrative regions in conservation: Balancing local priorities with regional to global preferences in spatial planning. *Biological Conservation*, 144: 1719-1725.

See also application in the following work, which clearly demonstrates how and why global solutions can be expected to be much more cost-effective than local solutions.

Moilanen, A. Anderson, B.J., Arponen, A., Pouzols, F.M., Thomas, C.D. 2012. Edge artefacts and lost performance in national versus continental conservation priority area. *Diversity and Distributions*, 19: 171-183.

In reality, conservation decisions are usually taken at national or regional levels, but they can also be at the scale of individual land parcels. Distributions of species and other biodiversity features are spread over multiple such administrative units. Different administrative units may, for whatever reason, have different priorities for conservation of biodiversity and its components (Arponen *et al*. 2005). As population dynamics and connectivity effects do not respect such administrative borders but extend across them, it is reasonable that the global conservation status of a biodiversity feature should influence its conservation locally. How, then, should the global and local conservation needs of a biodiversity feature be balanced to account for connectivity across borders?

It is possible to account for variable local and global priorities in the Zonation analysis via locally variable weights that are assigned to species or other biodiversity features. Below we describe two alternative methods for doing so. The first relies on what we call weak local representation. The emphasis with this approach is in global representation with locally varying weights, and it allows a degree of flexibility between what features are represented in which administrative regions. The second option, strong local representation, requires all features to be represented separately in each administrative region, when at all possible. This is irrespective of how the local abundance of the feature compares to global abundance.

Weak local administrative priorities (ADMU mode 1)

In this analysis variant, it is assumed that priorities (feature weights) can vary between administrative regions. It is also assumed that analysis is primarily global and that a degree of substitution of representation between regions is allowed. For example, if feature *j* is rare in area *A* but common elsewhere, representation of *j* within *A* would not be enforced. Representation for feature *j* is preferably obtained from a location where both the priority of *j* is high and the feature occurs at high local occurrence levels. Thus, there is no explicit guarantee that all features would become "protected" in all regions.

Technically, the implementation of this analysis variant in Zonation is simple. The equations that define the marginal loss of conservation value following the removal of a focal cell from the conservation solution include feature-specific weights, w_j (Moilanen *et al*. 2005, 2008; Moilanen 2007). Here, in addition to the global feature weights, w^G_j, the features can be assigned different weights for each administrative region separately, w^L_{jA}. In other words, one uses weight w^L_{jA} for feature *j* if the focal grid cell is part of administrative region *A*. Zonation calculates effective feature- and area-specific weights w^{eff}_{jA} by combining global and local weight given to features.

The global loss of benefit is defined as:

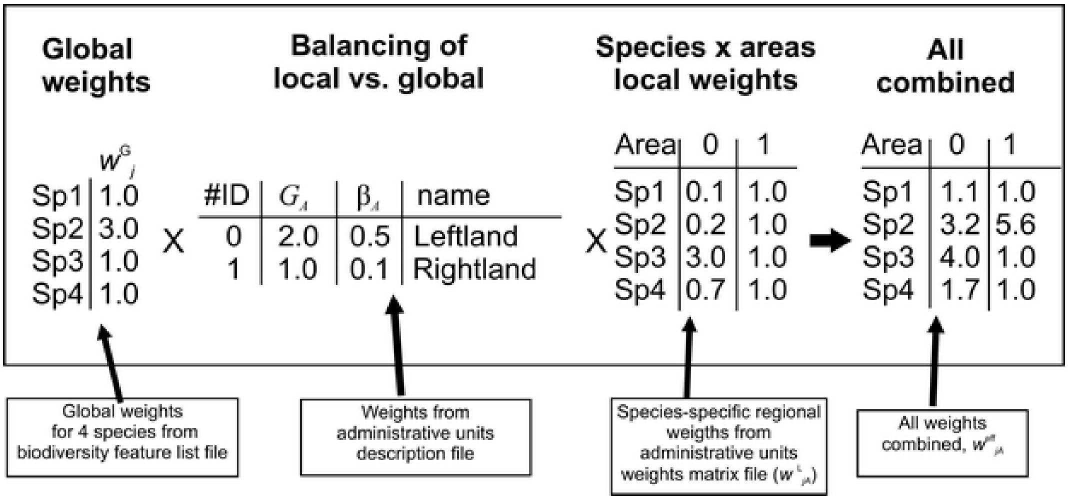
$$\Delta^{weak}(\mathbf{S},i) = \sum_j w^{eff}_{j,A_i} \left\{ f_j\left[R_j(\mathbf{S})\right] - f_j\left[R_j(\mathbf{S}-\{i\})\right] \right\}.$$

where the value is converted via the feature-specific benefit function f_j and multiplied by the effective local weight of the feature. Thus, in this analysis, representation is global while effective weights are local.

The effective local weights are constructed as a composite of global and local considerations in the following manner:

$$w^{eff}_{jA} = \left[\beta_A w^G_j + (1-\beta_A) w^L_{jA}\right] G_A,$$

in which w^{eff}_{jA} is the effective local weight, w^G_j is the global weight of feature *j*, w^L_{jA} is the locally given weight of feature *j* in area *A*, and β_A is a parameter tuning how much weight administrative area *A* places on global considerations. An administrative region can decide to focus only on global priorities (β_A=1) or completely ignore global priorities (β_A=0), or anything between ($0 < \beta_A < 1$). G_A is a region-specific weight which is used to multiply the species-specific weighting within each region. The figure below describes an example of how the effective local feature weights are constructed.



	w^G_j
Sp1	1.0
Sp2	3.0
Sp3	1.0
Sp4	1.0

#ID	G_A	β_A	name
0	2.0	0.5	Leftland
1	1.0	0.1	Rightland

Area	0	1
Sp1	0.1	1.0
Sp2	0.2	1.0
Sp3	3.0	1.0
Sp4	0.7	1.0

Area	0	1
Sp1	1.1	1.0
Sp2	3.2	5.6
Sp3	4.0	1.0
Sp4	1.7	1.0

Strong local administrative priorities (ADMU mode 2)

Heuristically, this analysis variant proceeds from the assumption that all locally occurring biodiversity features must be represented locally, but that global considerations should nevertheless influence local priorities. This differs from weak local priorities in that (instead of only weights) weights, representation, and conservation value are all considered both globally and locally. Conservation value is aggregated from a global component together with a local component for each administrative area *A*. Representation of each feature is considered separately globally and in each administrative region. Effectively, the analysis joins multiple conservation prioritization analyses: one global and one for each administrative region.

Conservation value is aggregated in the case of strong local representation in the following

manner:

$$V^{strong}(\mathbf{S}) = p\sum_{j} V_j^{global}(\mathbf{S}) + (1-p)\sum_{A}\sum_{j} V_{jA}^{local}(\mathbf{S})$$

$$= p\sum_{j} w_j^G f_j^G\left[R_j(\mathbf{S})\right] + (1-p)\sum_{A} G_A \sum_{j} w_{jA}^L f_{jA}\left[R_{jA}(\mathbf{S})\right].$$

The analysis is performed by a global actor who has the possibility to decide how much weight is given to priorities of local administrations and how much to global considerations by adjusting the balancing parameter *p*. When *p* = 0, the analysis is entirely based on priorities within each local administration (right hand side of the equation), and when *p* = 1 the analysis corresponds to a purely global analysis without any consideration of priorities of the local administrations. The priorities of each local administration are on the right hand side of the equation. G_A is an additional region-specific weight, which determines the region's value with respect to other regions. If a region's G_A is 10, then all value within the region, derived from the combined global and local feature weights and respective benefit functions, is multiplied by ten. Also, conceptually, one could use different benefit functions for a feature globally and in each administrative region, as allowed by different functions f_j^G and f_j^A, but this is currently not implemented in Zonation. Representation also has to be treated both globally, $R_j(\mathbf{S})$, and locally $R_jA(\mathbf{S})$, where the latter would be the fraction of the local occurrence of j covered in area A under some conservation solution **S**. Parameter β_A from the ADMU descriptions file is not used in the strong priorities variant.

The instructions for running a Zonation analysis accounting for administrative regions are in section 5.3.8.

2.13 Corridor connectivity

Zonation 4 introduces a novel approach to building corridors based on a penalty mechanism that is embedded in the prioritization (or ranking) process. In principle it is possible to incorporate corridors in a Zonation prioritization analysis by using other methods and/or software specifically intended to design corridors. With these other software you can generate input feature maps and/or masks that can be used in Zonation setups. Diverse methods and software tools are available that can design corridors. While the outputs of these tools can be used as inputs in Zonation or other spatial conservation planning tools, there has been a general lack of integration of corridor building methods with systematic conservation planning and spatial conservation prioritization methods.

The corridor building method implemented in Zonation is called corridor loss penalty (CLP) and can help in bridging the gap between computational methods for building corridors and methods for more general spatial conservation planning. The CLP maintains structural connections guided by species or feature specific spatial patterns. The essential idea is to extend the well-known boundary penalty (BLP) for reserve network aggregation. The BLP is as a penalty for fragmentation that is formulated as an additional term in the objective function of reserve selection methods, (see section 2.4.1). In a similar way, the CLP is an additional penalty term introduced to prevent loss of structural connections in networks of patches.

The distinguishing feature of the CLP as compared to other corridor building approaches is that corridors are identified within a general conservation prioritization process that accounts for a host of other relevant considerations simultaneously. By modifying a corridor penalty strength parameter it is possible to regulate the trade-off between increased structural connectivity via corridors and an array of other considerations relevant to spatial conservation prioritization.

In conservation, corridor building faces many challenges. What is the best way to build corridors in irregular networks made of a large number of patches of diverse morphology? How should corridors be placed in broad-scale planning? How does one determine an appropriate corridor width, and account for the diverse requirements of many species and other relevant factors simultaneously? The corridor building mechanism of Zonation has been parameterized so that it is flexible and can be used to address these questions, but it is still relatively simple to use in simple applications.

The CLP has several distinctive characteristics. First, habitat patches need not be specified a priori, rather they are dynamically identified throughout the prioritization process. End points for corridors need not be specified either; corridors between patches or networks of patches will emerge throughout the ranking. Also, there is no need to specify distance thresholds or thresholds on the relevance of patches for connectivity. Instead, in the Zonation ranking different patches and networks are connected at different priority levels, proportionally to their aggregated values and proximities. Finally, and in general in Zonation, there is no need to specify low-level targets for species or other features, even though this is also possible.

In practice, using the CLP requires at least defining or adjusting two key parameters:

- *Penalty strength*, which is defined as a real number that regulates the trade-off between increased connectivity via corridors and any other considerations relevant to conservation (defined in the marginal loss mathematical expression of Zonation).
- *Corridor width*, which defines the minimum width of corridors (in practice specified as a number of cells). The method will try to keep patches connected by corridors of at least this width. Connections narrower than this parameter are not considered as corridors.

In addition, an important aspect to consider is where in the landscape corridors should be generated. By default, Zonation will produce corridors in any location where there is a high potential to increase connectivity (or a high corridor loss penalty). This may not be sensible in many cases where it may be desirable to exclude certain areas of the landscape from the corridor building mechanism, or, equivalently, to define a set or network of areas within which corridors should be build. This is possible by using so-called corridor domain layers. Where some blocks of the landscape are in (1 in a raster map) and other areas are masked out (missing or no-data in a raster map).

It should be noted that there are no known “best” settings for the CLP method. The effects of different parameters will be case-specific in response to case-specific objectives, and would depend on the specific patterns of the distribution of a feature. However, sensible results may be obtained only for a potentially narrow range of values for some parameters, especially the penalty strength (in a way that resembles the boundary length penalty, BLP). As with any spatial priority ranking careful interpretation of results is required. The effects of the corridor penalty strength parameter can be easily understood by analogy with the boundary length penalty (BLP). With a too high corridor penalty strength the CLP will tend to generate connections between poor areas or even single end elements, in a way similar that a too high BLP will generate undesired and artificial geometric boundaries for reserves. As with the BLP, the CLP method typically requires exploring a range of strength values until a suitable balance between structural connectivity via corridors and all other factors is found.

There are several additional parameters to adapt and fine-tune the way the CLP works, including how to define spatial patterns of preference for corridors, how to allow or disallow redundancy in connections between patches, and options related to outputs specific to the analysis of corridors. See the following chapters and references for full details.

Finally, we briefly outline how the CLP is defined formally. The CLP is integrated in the Zonation prioritization met-algorithm as an additional term that is added to the mathematical expression that defines the marginal loss for every landscape element (cells or raster maps or planning units). When using the CLP the marginal loss for a focal cell i is:

$$\delta_i = RR(i) + CLP(i)$$

where RR(i) is the marginal loss of the focal cell, calculated according to a particular removal rule (and excluding penalties), and CLP(i) is the corridor loss penalty. If using the CLP together with other mechanisms such as the boundary length penalty (BLP) or the boundary quality penalty (BQP), there would be an additional penalty term (see sections 2.4.1 and 2.4.3, respectively). More specifically, the CLP for a focal cell i is defined as:

$$CLP(i) = k\sigma(i) m_i CLPR(i)$$

with κ the corridor penalty strength parameter, σ(i) the split function (taking value 1 if removing cell i causes a network of patches to be split, or 0 otherwise), and CLPR(i) the corridor loss penalty rule. The corridor loss penalty rule selects between different criteria to evaluate the relevance of splitting networks. This rule can be based on the areas or normalized richness of the network being split and those of the resulting sub-networks. The m are weights given in a corridor mask or corridor domain layer, and can be used to define higher or lower preference for corridors in different parts of the landscape. If no corridor domain layers are specified, the m are all 1. See references for full details on the definition of the CLP penalty term.

Literature

The corridor building method implemented in Zonation is described in:

Pouzols, F.M. and A. Moilanen. 2014. A method for building corridors in spatial conservation prioritization. *Landscape Ecology*, 29: 789-801. doi: 10.1007/s10980-014-0031-1

2.14 Assumptions & limitations

This section lists some known assumptions and main limitations of the presently available Zonation implementation.

- The Zonation software presently only accepts data in grids and point observation lists; it does not accept vector-based planning units. This limitation is practical, not conceptual, but it is unlikely to be removed in later Zonation versions. Vector-based data simply needs to be sampled into grids using a GIS tool.
- Zonation operates on biodiversity patterns. In particular, it does not include any explicit mechanism for handling multi-year planning with considerations of stochastic site availability and the possibility of site loss. Note, however, that the retention

analysis does account for the first-order effects of expected habitat loss, for example the "minimize loss" strategy can be implemented via the retention feature.

- At present Zonation only has a single option per selection unit, that is, protect or not, restore or not, maintain or not, etc. In a more advanced analysis one could envision multiple alternative outcomes for each location (for example, different levels of protection or restoration). However, one can do analysis where alternative conservation options are explored. This involves developing sensible restoration scenarios and then running Zonation prioritization for each. (See Thomson *et al*. 2009)
- Zonation is not, and will not be, software for statistical habitat modelling.
- Zonation is not, and will not be, software for the simulation of spatial population dynamics.
- The Zonation GUI does some GIS-like operations that allow you to visualize results and perform calculations on raster maps, but it is not meant to replace your favorite GIS program.
- Except for the corridor building function, Zonation does not directly support design of linear elements such as road networks.

Major limitations removed since Zonation v.2.0 (available with v 3.1):

- Multi-criterion analysis and balancing of alternative land uses is now possible.
- Ability to work with environmental data and community classes.
- The 4GB 32 bit Windows memory limit has been removed.
- Effects of conservation action can now partially be accounted for via retention analysis.
- Use of multiple types of GIS grids has become possible.

Limitations removed since Zonation v.3.1 (available with v.4):

- Zonation 4 comes with a new corridor building method which adds the capability to operate on path-like connectivity.
- Different kernel shapes or functions of arbitrary form can be used, including predefined simple functions and custom (user defined) functions.

- Size limits and other limitations related to planning units and administrative units have been removed, and in principle it is possible to use as many planning or administrative units as memory allows.
- Zonation 4 can be orders of magnitude more efficient in using memory. It is now possible to process of the order of thousands or tens of thousands of features in landscapes with tens of millions or hundreds of millions of sites/cells, with RAM memory in the range from ~10 GB to ~100 GB. [exact number depend on how much of the layers is data and how much missing data]
- A Linux version of Zonation has also been released.

Part

III

3 ZIG - The Zonation software

3.1 Important general information about file operation

The following sections describe how the Zonation software is used from the command line, what kind of input files it requires, what kind of output it produces, and which analyses can be done with the program. For concepts and analyses implemented in the software, see section 2. Zonation v.3.1 and higher contains a Graphical User Interface (GUI), a separate piece of software which calls zig4.exe executable to execute analyses. The use of the GUI is described in section 4.

File names

The length of file names in Zonation is not limited and they may contain numbers and any unicode characters. File names containing special characters such as spaces can be read if they are in quotation marks (e.g., "my zonation run.bat"). Directory separators can be entered as slashes or backslashes. File names can be enclosed in double quotation marks. In this case, quotation marks in paths can be entered in escaped form. For example, the file path c:\my “great” zonation run.bat can be entered as “c:/my \”great\” zonation run.bat” in quoted form.

Specifying file paths

When entering file names, please note that **the path to the file must also be specified**. In the command prompt, the paths in text files are relative to the current working directory. In the GUI they are relative to the project batch file. If all files are located in the same folder, the file name will suffice:

```
my_file.asc
```

However, most often it is convenient to have subfolders for different analyses. If the subfolders partition from the folder where the batch files are located, please specify the path from that folder:

```
subfolder/my_file.asc
```

or with the full path:

```
D:/Data/subfolder/my_file.asc
```

This applies to all files that are referred to at any stage of the analysis setup. The file paths have been left out of the examples for the sake of simplicity.

Supported file formats

Zonation uses a modified version of GDAL 1.8.0 (Geospatial Data Abstraction Library) to read and write raster files. Any standard GIS format supported by the GDAL library can be used for raster map inputs. GeoTIFF (.tif), Erdas Imagine (.img), Band Interleaved by Line (.bil), Portable Networks Graphic (.png), JPEG-2000 (.jpg or .jpeg), and Arc/Info ASCII Grid (.asc) are well tested. The complete list of supported formats can be found at http://www.gdal.org/formats_list.html.

Any supported file format can be loaded anywhere in Zonation where a raster file is needed. Different formats can be used simultaneously. For example, a part of the species files can be in ASCII format while the rest are in GeoTIFF. If a multiband dataset is used in Zonation, only the first band is read and the rest are ignored.

We recommend the use of compressed raster formats with Zonation. Very commonly the used datasets are masked or mostly zero, which makes the compression very effective. It is not uncommon to have an ASCII based dataset squeezed to less than a hundreth of the size of the original when transformed into a deflated GeoTIFF. It is also much faster to read and write smaller files. With dataset sizes in tens or hundreds of gigabytes this can mean big savings.

All text files are expected to be encoded in utf-8.

3.2 Running Zonation

The Zonation numerical core operates through a command line application. There are two options for managing Zonation operations: Zonation can be called directly from the command line (section 3.2.1) or the Zonation GUI (section 4) can be used to manage and monitor runs. The GUI is useful for setting up and managing multiple analyses for a single set of input data, and it can also attach to runs that have been started from a batch command and display their progress.

With Zonation it is possible to (1) make new analyses and (2) load old solutions (section 3.2.3). The solution loading option can be used either to review old solutions or to investigate how old solutions would perform under new (different) assumptions (section 3.5.2).

After running Zonation, please check the .run_info.txt file (section 3.4.1) for verification about data characteristics and options used. Also check any warnings or errors that occurred during the run as the present Zonation version does not stop execution at every suspicious input.

3.2.1 Command line

A practical way to run Zonation is calling it with a batch file. A batch file is a simple Windows/DOS command line file that can be used to give commands to Windows. Batch files can be created with notepad by writing the program call in a new notepad document and saving it with the batch file extension (.bat). The file name extension of a batch file has to be .bat (e.g., do_zig.bat). If there is anything else after the .bat suffix, Windows cannot identify the file as a command file. Batch files are practical because they make it possible to run multiple analyses at one time, and they save information from the analysis (i.e., which input files and settings have been used) for later review.

Note

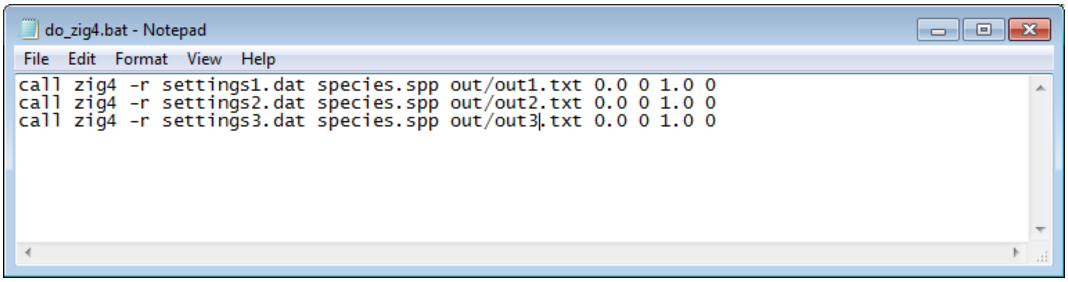
Windows/DOS batch files do not normally work in the command line of GNU/Linux systems (or other operating systems). However the parameters to the zig4 command line tool are the same (you can use the same or similar commands). Also, the Zonation graphical interface supports batch files as project files regardless of the operating system that you use.

To run Zonation from the command line, call the program by typing "call zig4" in the batch file, followed by:

- -r, if a new solution will be calculated, or -lfilename if an existing solution is loaded (the name of the rank.asc file, for example, -loutput.rank.asc)
- Name of the run settings file (.dat; section 3.3.2.3)
- Name of the biodiversity feature list file (.spp; section 3.3.2.2)
- A name for the output files (`subfolder\my_output_name.txt`).
- Value of the uncertainty parameter α for the uncertainty analysis (UCA; section 2.5). If uncertainty is not included in the computation, this should be set to zero.
- Value to determine whether distribution smoothing (section 2.4.2) is used in the analysis (parameter = 1) or not (parameter = 0).
- Factor for multiplying feature-specific widths of the dispersal kernel (a -values, the second column in the biodiversity feature list file). This parameter makes it possible to produce multiple solutions with different feature-specific scales of landscape use without needing to change the individual dispersal kernel widths manually after each run. All kernel widths will be multiplied with this factor.
- Value to indicate whether the program window will be left open (parameter = 0) or closed (parameter = 1). Closing the window at the end will allow the program to move on to the next run. If a batch file is composed for performing multiple runs, it is important to write 1 at the end of each line or the program will not proceed to the next Zonation run.

A batch file is run simply by double-clicking the file icon from Windows. A batch file can also be used to run Zonation analyses through the graphical user interface (see section 4).

A batch file is useful to experiment with different levels of distribution smoothing, uncertainty analysis, BQP, etc. (several settings files) or to experiment with alternative biodiversity feature weighting schemes (several biodiversity feature list files). Batch files can also be used to run the most complicated analyses (requiring long computation times) overnight or over the weekend. To run multiple analyses, the simplest batch file consists of several command lines, each calling the program with different parameters, as shown in the figure below. Note that for compatibility with Zonation 3.1, it is fine to use both zig3 or zig4 in batch files.



```
call zig4 -r settings1.dat species.spp out/out1.txt 0.0 0 1.0 0
call zig4 -r settings2.dat species.spp out/out2.txt 0.0 0 1.0 0
call zig4 -r settings3.dat species.spp out/out3.txt 0.0 0 1.0 0
```

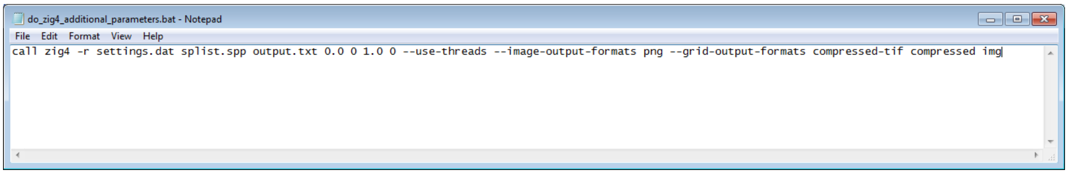
Here, the same biodiversity feature list file is run with three different settings. **Remember to identify the output files separately (give them different names) or the program will overwrite old results after each run**. By default Zonation produces only the standard output automatically (see section 3.4.1 for file output). Remember to develop the settings for automated post-processing analyses if special processing of outputs are needed.

Optional command line parameters

In addition to mandatory regular command line parameters, there are optional parameters that do not affect the simulation results. These parameters can be in any order but they must always reside after the regular parameters.

`--use-threads N`	Use N simultaneous threads, where N is optional. If N is not specified, the program tries to use the number of hardware threads (e.g., number of CPUs or cores or hyperthreading units) available on the system. By default, Zonation uses a single thread. If neighborhood quality penalty (NQP) is used, a single thread is used regardless of this option.
`--removal-rule RULENUM`	Where RULENUM is the removal rule to use (from 1 to 5), as in the "removal rule" option of the settings file, see section 3.3.2.3 "Run settings file." Using this options saves you from having to make two or more variants of a settings file where only the removal rule would change.
`--image-output-formats LIST_OF_IMAGE_FORMATS`	Where the formats given in the list can be any of the following: png, bmp, jpg, emf. Formats should be separated by spaces. For example: `--image-output-formats png jpg`. This parameter generates output color maps in the specified formats. Format specifiers can be in any order. No files are created if none of the formats are specified. By default, Zonation creates the color map in JPEG format. The following files are created with each specifier ("output" is the specified output file without the filename extension): `png`: output.png `bmp`: output.bmp `jpg`: output.jpg `emf`: output.emf

	For backwards compatibility, EMF format produces exactly the same file as BMP. We do not recommend its use. Note that since Zonation 4, the default format is png.
`--grid-output-formats LIST_OF_RASTER_FORMATS`	Where the raster formats should be separated by spaces. The following formats are supported: asc, tif, img, compressed-tif, compressed-img. For example: `--grid-output-formats asc compressed-img`. This parameter generates the Zonation output grids in the specified formats. This step will save time if a GIS program will be used to conduct further analysis on the map. Format specifiers can be in any order. No files are created if none of the formats are specified. By default, Zonation creates the output grids in compressed-tif format. The following files are created with each specifier ("output" is the specified output file without the filename extension): `compressed-tif`: output.rank.compressed.tif output.prop.compressed.tif output.wrscr.compressed.tif `compressed-img`: output.rank.compressed.img output.prop.compressed.img output.wrscr.compressed.img `tif`: output.rank.tif output.prop.tif output.wrscr.tif `img`: output.rank.img output.prop.img output.wrscr.img `asc`: output.rank.asc output.prop.asc output.wrscr.asc Note that since Zonation v4, the default format is compressed-tif. All output formats except ASCII have 32 bit float element type. Tif produces a GeoTIFF file and img an Erdas Imagine file. Compressed GeoTIFF uses DEFLATE compression. Note that the same GIS formats are used to generate the weighted range size corrected richness map and the proportional loss map (see section 3.4.1 for details on the maps that Zonation outputs by default). Also note that since Zonation v.4 the proportional loss map is not generated by default (this can be changed in the settings file, see section 3.3.2.3).
`--warp-factor NUM`	Where NUM is the warp factor given as an integer, as described in section 3.3.2.3 "Run settings file." If this option is given in the command line the value specified will override any other warp factor value specified in the settings file.



```
call zig4 -r settings.dat splist.spp output.txt 0.0 0 1.0 0 --use-threads --image-output-formats png --grid-output-formats compressed-tif compressed img
```

Additional command line parameters

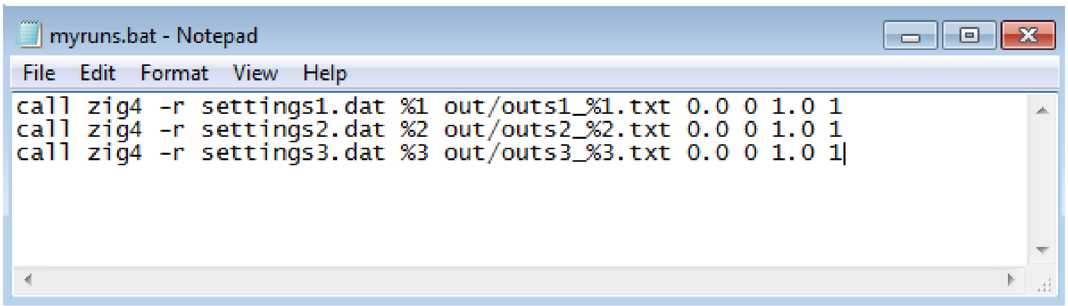
`zig4 -h` `zig4 --help`	These commands show a list of parameters and options.
`zig4 -v` `zig4 --version`	These commands display version information as well as copyright conditions of use of the Zonation computation core.
`-h, --help`	Print help.

By running the command `zig4 --help` or `zig4 -h` you can get the list of supported options. The command `zig4 --help` should produce an output similar to this:

```
Usage: zig4 run_mode settingsfile spp_file output_files_prefix IGa useSmooth SmoothMult
Run Zonation core, Zonation 4.0.0b24_compact, build: Feb 25 2014 19:16:22
Parameters:
 run_mode -r,-l               -r (run mode, calculate new solution),
                              -l FILE (reload mode, load solution in rank FILE).
 settingsfile                 settings (.dat) file as specified in the manual.
 spp_file                     biodiversity features list file.
 output_files_prefix          name (prefix) of output files
 IGa                          value of uncertainty parameter (alpha).
 useSmooth                    use distribution smoothing (1) or not (0).
 SmoothMult                   multiplying factor for species-specific dispersal
                              kernel widths.
 autoclose0_or_1              close window at the end (1) or not (0).

Supported options:
 -h,--help                    print help and exit.
 -v,--version                 print version information and exit.

 --grid-output-formats LIST   LIST is a list of formats separated by spaces.
                              Supported formats: asc tif img compressed-tif.
 --image-output-formats LIST  LIST is a list of formats separated by spaces.
                              Supported formats: png bmp jpg emf.
 --removal-rule N             set removal rule to N (overriding settings
                              file), 0 - CAZ, 1 - ABF, etc.
 --warp-factor N              set warp factor to N (integer value).
 --use-threads N              use N hardware threads (best option
                              automatically chosen if N not given).
```

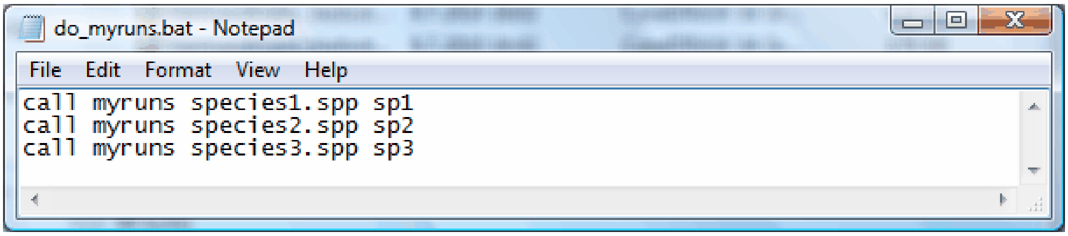
Advanced use of nested batch files

The description so far has covered the basic use of batch files. Multiple nested batch files can be used to execute several runs simultaneously. For example, a nested batch file can be used to run the batch file above (do_zig.bat) with different feature weights defined in separate biodiversity feature list files. To do this, adjust the settings as follows and save as a new batch file:



```
call zig4 -r settings1.dat %1 out/outs1_%1.txt 0.0 0 1.0 1
call zig4 -r settings2.dat %2 out/outs2_%2.txt 0.0 0 1.0 1
call zig4 -r settings3.dat %3 out/outs3_%3.txt 0.0 0 1.0 1
```

Then create another batch file, which in turn calls the "myruns" batch.



```
call myruns species1.spp sp1
call myruns species2.spp sp2
call myruns species3.spp sp3
```

Here the first parameter after "myruns" defines which biodiversity feature list file is used when running the "myruns" batch and the second parameter (sp1, sp2, sp3) defines a part of the output file name to distinguish which biodiversity feature file has been used in each run. In the "myruns" batch file these parameters will are referred to as %1, %2, and %3. When running the "do_myruns" batch, the program first calculates solutions using the "species1.spp" file with three different settings and gives each of the output files an ending "sp1" (outS1_sp1, outS2_sp1 and outS3_sp1). Then the program repeats the procedure with the other two biodiversity feature list files, "species2.spp" and "species3.spp". Thus, running the batch produces nine solutions with different settings and feature weights composition.

Using nested batch files is extremely useful when running many solutions using combinations of settings. For example, if we would use only one batch file to run the solutions described above, we would need to write a separate call for each of the nine solutions, and with many calls it is relatively easy to introduce errors into some of them.

Run simultaneous instances: zig4run

The **zig4run** utility can be used to run simultaneous Zonation instances from the command prompt. It is often more effective to use simultaneous processes than to use multiple threads in Zonation. On the other hand, multiple simultaneous single-threaded Zonation processes use more memory than a single process with multiple threads. The rule of thumb is that the more biodiversity feature layers used in the analysis, the more effective a multi-threaded Zonation becomes. With BQP set on, it is very efficient to use multiple threads.

Zig4run parses Windows batch files and tries to recognize calls to Zonation. It queues recognized calls and runs them simultaneously. It monitors the state of the running processes and prints a progress percent on screen. If a particular Zonation process is already running, it hooks up to the process and starts monitoring it until it finishes. A Zonation process is identified by the absolute path of its output file.

The Zonation calls in the batch file must each be on their own lines and they must not span multiple lines. There can not be any other command on the same line as a call. Zig4run will not recognize any constructs that affect the control flow in the batch files. For example, if a Zonation instance is called in a loop, only one call will be recognized.

Example batch files

A legal batch file

```
zig4.exe -r set.dat splist.spp output.txt 0.0 0 1.0 0 --use-threads --image-output-form

zig4.exe -r set2.dat splist.spp output.txt 0.0 0 1.0 0 --grid-output-formats asc tif
```

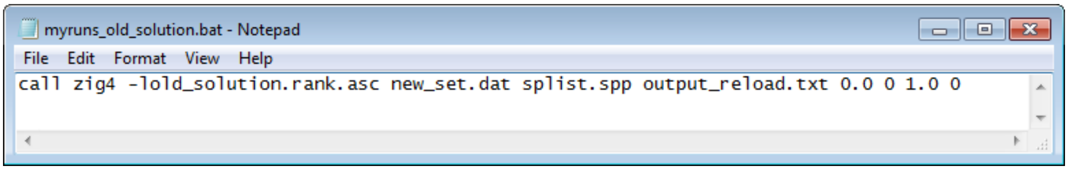
An illegal batch file

```
call zig4.exe -r set.dat splist.spp output.txt 0.0 0 1.0 0 \
        --use-threads

zig4.exe -r set.dat splist.spp output.txt 0.0 0 1.0 0 && zig4.exe -r set.dat splist.spp
output.txt 0.0 0 1.0 0

for %%f in (set.dat set2.dat) DO zig3.exe -r %%f splist.spp output.txt 0.0 0 1.0 0
```

3.2.2 Loading previously calculated Zonation solutions

It is also possible to load previously calculated Zonation solutions. This is a useful feature to conduct further analysis with an old solution or to test the performance of an old solution in different circumstances (see section 3.5.2 for solution cross-comparison using solution loading).

When operating the program from batch files, type "`-lfilename`" as the second parameter of your call. "`-l`" points out that an old solution is loaded. For "filename", enter the name of the priority rank file from the old solution (**.rank.asc file**, one of the output files produced during each run; see section 3.4.1). A typical call when loading an old solution looks like this:



```
call zig4 -lold_solution.rank.asc new_set.dat splist.spp output_reload.txt 0.0 0 1.0 0
```

Remember to give a new output name in the call to avoid overwriting the old solution.

Make sure that the extent of the solutions is the same when using different sets of biodiversity features in the original and the reload runs. Otherwise, it would not be obvious what to do with ranked cells that are not relevant anymore (no occurrences of any of the new features) or with cells that were not ranked but have occurrences in the new set of features. The Zonation core checks this before starting the removal process. If the check fails, the core reports the error and shows what cells are wrong (up to a maximum of typically 10 messages per type of error). The extent of the output rank maps will be the same only if the extent of the union of the non-missing cells of biodiversity feature layers are the same in the original run and in the new (reload) run. Note that some of the options normally given in the settings file do not make sense when Zonation is run in "-l" or re-load mode. For convenience these options (for example, edge removal, add edge points, etc.) are ignored. In principle it is possible to re-load solutions using different hierarchical removal masks.

3.3 Input files & settings

3.3.1 Introduction

A set of input files is needed to begin a Zonation analysis. Some of the input files are compulsory while others are optional. Some of the tutorial files can be used as templates when creating the input files for an analysis.

When entering file names, remember to **specify the relative file path** as well. If the files are in the same folder as the batch file, the file name alone will suffice:

```
my_file.asc
```

However, most often it is convenient to have subfolders for different analyses. If the subfolders partition from the folder where the batch file is saved, please specify the relative path from that folder:

```
subfolder/my_file.asc
```

If the Zonation project (batch) file and the analysis files are saved in different places (e.g., the batch file in C:/Program files and analysis files at D:/Data), then the entire (absolute) path to the analysis files should be specified like:

```
D:/Data/subfolder/my_file.asc
```

Said in another way, all file paths are expected to be relative to the folder where the Zonation project .bat file is. If input or output files should be elsewhere, the full file path starting from the drive letter (e.g., c:/) needs to be specified. This applies to all input files that are referred to at any stage of the analysis setup; these file names occur for example in the biodiversity feature list file and in the run setting file. The same rules also apply when specifying the location where the output files should be saved. For the sake of simplicity, the file paths have been left out in the examples.

Warning: It is recommended that special characters, including spaces, should be avoided in file names. Earlier, spaces in file names or very long paths / file names have also caused problems.

3.3.2 Compulsory files

There are three types of compulsory files needed for every Zonation analysis:

1. **Distribution grids of species or other biodiversity features**: Describes the distribution and local density of each feature across the landscape (section 3.3.2.1).

2. **Biodiversity feature list file** (section 3.3.2.2): Indicates which biodiversity feature grids are to be used in the analysis. Certain settings, such as weighting, are adjusted in the biodiversity feature list file.

3. **Run settings file** (section 3.3.2.3): Defines the settings and Zonation features used in the analysis.

3.3.2.1 Biodiversity feature map files

A biodiversity feature map file is a GIS raster file of feature (species, habitat type, etc.) distribution. There is one file for each biodiversity feature. Note that in Zonation v.1.0. and v.2.0. these grids needed to be ascii grids (described below). Zonation v.3.1 and higher uses the GDAL library for handling raster maps, so in theory these grids can also be in any standard GIS raster format, whether ascii or binary. See the GDAL library for supported file formats.

GIS raster files Zonation uses a variety of GIS raster files such as biodiversity feature (species) distribution layers, cost layers, mask files, or uncertainty map layers. All these files need to be exported from a GIS software (e.g., from ArcGIS) or produced with some other appropriate software. Zonation v.3.1 and higher utilize the GDAL library for the handling of GIS raster files. This adds flexibility compared to Zonation v.2.0, in which the feature raster files and other rasters could only be in the ascii raster format. Now, any standard GIS format can be used for raster map inputs. GeoTIFF (.tif), Erdas Imagine (.img), Band Interleaved by Line (.bil), Portable Networks Graphic (.png), JPEG-2000 (typically .jpg or .jpeg), and Arc/Info ASCII Grid (.asc) are well tested. A suffix is not necessary wherever a grid raster can be used as Zonation will recognize the format automatically. The benefit of the alternative formats is that some of them save 80-90% of disk space compared to ascii rasters. They also load in just 10-20% of the time needed before. However, only an ascii raster can be examined with a text editor such as notepad, but most of the other (binary coded) formats can only be examined using a GIS software.

Whatever the grid file formats, it is important that all feature distribution rasters have the same grid size! **This means that in all files the number of columns and rows as well as the size of cells should be equal**. It is equally important to have **at least one row of no data on each edge of the raster matrix.** This is due to computational reasons and the lack of these empty rows leads to a situation where the program automatically transforms the values on edge rows to missing data. This in turn may alter feature distribution information—a possibility one should be aware of. Missing data in the feature distribution rasters do not necessarily need to be congruent between all features; the program will run if cells marked as "no data" for one feature has values for other features. However, if missing data is not aligned, there are implications for the use of the BQP.

Description of ascii grid file format

Note

While ascii raster files are easy to manipulate and understand, as shown in the example below, Arc/Info ASCII Grid (.asc) is a very inefficient format. It is strongly discouraged to use ascii raster files, especially when working with a large analysis area. With large grids in ascii format you can easily fill up your hard disk and slow down Zonation runs painfully.This applies to both input and output rasters maps. Starting with Zonation v.4 the default format for output rasters is (compressed) GeoTIFF

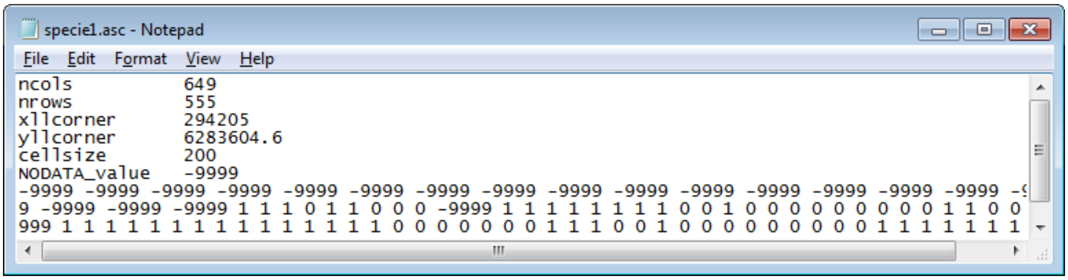
Ascii files need to include the standard GIS raster file header:

`ncols`	Number of columns
`nrows`	Number of rows
`xllcorner`	X-coordinates of the low-left corner
`yllcorner`	Y-coordinates of the low-left corner
`cellsize`	Cell size used in the raster file
`NODATA_value`	Definition of "no data" values. In example files, "no data"

has a value of either -1 or -9999.

The matrix showing the distribution of the biodiversity feature is entered after the header (see example below). Each value in the matrix describes the occurrence of the feature in a specific cell. Values can be of any form of data (e.g., probability of occurrence, presence/ absence data, number of population, etc.) as long as the data is in same format across all biodiversity feature map files. Note that value 0 in the matrix indicates that the feature does not occur in the cell with certainty, whereas lack of data must be marked as -1 or -9999, or with a similar value indicating "no data".

Remember also to use **decimal points, not commas**, in all the input files.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      200
NODATA_value  -9999
-9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -
9 -9999 -9999 -9999 1 1 1 0 1 1 0 0 0 -9999 1 1 1 1 1 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 1 0 0
999 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1
```

In Zonation v.3.1 and higher, biodiversity feature map files can contain several types of biodiversity features and considerations, including:

- The distributions of individual features (observed or predicted occurrences, abundances, or other measures of suitability)
- Distributions of community types (see section 5.3.1)
- Negative conservation value of areas that are desirable for competing land uses (see section 5.3.4)
- Distributions for the same sets of biodiversity features for multiple time steps (see sections 5.3.6 and 5.3.7)
- Any properties of the habitats that can be assigned a value that reflects the conservation value of the site

Different types of biodiversity features can be combined in a single analysis simply by listing them in the same biodiversity feature list file. In such an analysis, it is advisable to pay special attention to the values assigned to the features, as well as their weighting, to attain a desirable balance of prioritizing different kinds of features.

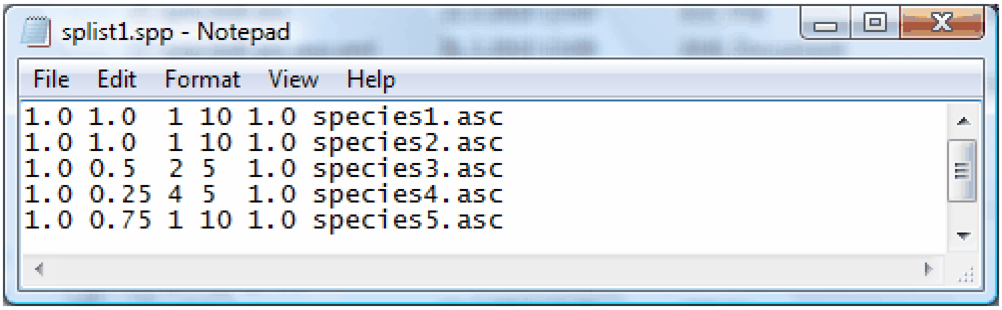
Please note also that Zonation can make use of point occurrence data as well as raster files. Use of point occurrences is most appropriate when a biodiversity feature has only a few occurrences that are not enough to fit reliable distribution models to produce a comprehensive distribution map. The input for a species of special interest (SSI) is a list of observation locations instead of a map. See section 3.3.3.1 for details about using point occurrence data.

3.3.2.2 Biodiversity feature list file

A biodiversity feature list file (.spp) contains a list of all the biodiversity feature map files that will be used in an analysis. [Note that "feature" or "species" is used here as a shorthand for any biodiversity feature, including species, habitat types, ecological communities, environmental conditions, ecosystem services, or any other information that acts as the basis for decision making.]

.spp files A suffix for feature list files such as a biodiversity feature list file or an uncertainty analysis weights file. In Zonation, these files would contain a list of species or other biodiversity feature raster files with set-up parameters. All .spp files must be created with Notepad. The tutorial files can be used as templates. When saving a .spp file, remember to add the suffix (.spp) after the file name. The files could also be named .spp.txt to emphasize that they are text files containing biodiversity feature lists.

In the biodiversity feature list file, each feature is represented by one row of information. This file always contains six columns, except when using the generalized benefit function as a cell removal rule, in which case there are nine columns (only in Zonation v.2 and higher). The six-column version used with the other cell removal rules has the following parameters:



```
1.0 1.0  1 10 1.0 species1.asc
1.0 1.0  1 10 1.0 species2.asc
1.0 0.5  2 5  1.0 species3.asc
1.0 0.25 4 5  1.0 species4.asc
1.0 0.75 1 10 1.0 species5.asc
```

Example of a biodiversity feature list file

Each number describes the corresponding column shown in the biodiversity feature list file above.

1. This column indicates the species weight (w_j) in the Zonation algorithm. If no weights are used, this should be set to 1.0. In the example above all features (species) have equal weights (1.0). Typically weights have positive values, but in Zonation v.3.1 and higher weights can be zero or negative as well. Using 0.0 as a weight means that the program calculates the performance of the features during cell removal, but that the feature does not in any way influence the priority ranking of the landscape. Thus, other features are acting as surrogates for zero-weight features.

   A feature with a negative weight is something undesirable that should be removed from the landscape as soon as possible. One way to think of negative-weight features is that they can represent (multiple) opportunity costs.

Note that feature-specific weights have no influence on the analysis if the target-based benefit function is used as the cell removal rule. This is because all features have targets when using this cell removal rule so weights have no influence on the outcome.

The weight parameter can include considerations such as degree of historical distribution loss, feature local or global scale priority, taxonomic uniqueness, etc. All such considerations should be aggregated to a single weight reflecting the importance of the biodiversity feature in the analysis. Two important notes about weighting are that (1) weighting is a political decision, and there is no general method for determining correct weights and (2) weights should still be used. If weights are implicitly taken as equal, then that is also a weighting system. In analysis, a rare feature or a species with unique taxonomic history should have a higher weight than a feature that has many rather similar features.

2. Column 2 contains the α -value of the biodiversity feature-specific scale of landscape use (parameter of negative exponential, α_j). This parameter is only used when distribution smoothing or matrix connectivity is enabled. However, **this parameter should always be specified**, and any numerical value (e.g., 0.0) will do fine if distribution smoothing is not used. The α value is the parameter of a dispersal kernel and it has an inverse relationship to the scale of connectivity of biodiversity features. This value can also be calculated based on, for example, the dispersal capability or the home range sizes of a species. The following three points are relevant to understand how to use this parameter:

   - *How to enable distribution smoothing*. Remember that the value of the sixth parameter in the batch file (see section 3.2.1) determines if distribution smoothing is used or not. It cannot be enabled or disabled in the settings file. Thus, in order to check whether distribution smoothing is being used you have to look in the batch files or command lines.

   - *How to define the degree of smoothing for individual features*. The α value indicates how species use the surrounding landscape and can be calculated based on, for example, the dispersal capability or the home range sizes of the species. It can be calculated as:

$$\alpha = \frac{2}{[\text{Use of landscape (in same units as cell size)}]}$$

     It is extremely important to make sure that the distance units used in this calculation are consistent with the cell size units used in the feature distribution files. The above equation assumes that 'use of landscape' is given in the same units that are used for the cell size value in the feature distribution files (if these are .asc files, the cell size value is given in the 'cellsize' field in the first rows). For example, if the cell size in the distribution files is 1 km and the known/ estimated mean dispersal capability of a species is 3 km, then the value of alpha for this feature is:

$$\alpha = \frac{2}{3} \approx 0.667$$

One of the most common errors in the use of connectivity in Zonation, is that dispersal α is calculated assuming everything is in kilometers, but the cell size in the GIS layer actually uses meters as the unit. In this case, the alpha-value calculated becomes 1000x too large and connectivity is effectively retained within the cell and use of connectivity makes no visible change to the output. Usually when you switch on connectivity, you should expect a visible change in the priority ranking - if nothing changed, please verify that connectivity is in fact used in the new run. Correct is to specify the 'use of landscape' parameters using the same units as the cell size in input layers. For example, if the use of landscape is initially given in kilometers (say 1.5 km), but the feature distribution files specify a cell size in different units (say 500 m), α should be calculated specifying the use of landscape in meters as well:

$$\alpha = \frac{2}{1500} \approx 0.00133$$

The alternative, although more prone to errors, is to use the value 'use of landscape' specified in different units. In this case, the value of α needs to be calculated using the following scale adjustment:

$$\alpha = \frac{2}{[\text{Use of landscape (in any units)} * \{\text{ratio between units}\}]}$$

where 'ratio between units' is the ratio between the units used for the 'use of landscape' parameter and the cell size. For example, if the use of landscape is given in kilometers (1.5 km) but the feature distribution files specify a cell size of 500 meters, the units ratio will be 1000. This is the ratio between the units used for the use of landscape (1 km = 1000 m) and the units of the cell size of input layers (1 m). In this example α would be calculated as follows:

$$\alpha = \frac{2}{[1.5 * \{\text{ratio between units}\}]} = \frac{2}{1.5 * 1000} = 0.00133$$

The 'ratio between units' is needed to keep the α –**value in same unit of length as the cell size** given in the feature distribution map file. Another common unit used in raster files is degrees. These also need to be converted to get the correct α -value. Let us assume that the cell size in our example was 0.0083 degrees, equaling approximately 0.860 kilometers. Thus the α -value in this case is:

$$\alpha = \frac{2}{1.5 * 0.0083/0.860} \approx 138.2$$

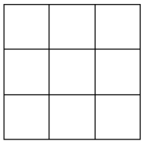
It is important to understand that this parameter is NOT the same α -value that is used in uncertainty analysis! The two parameters only happen to have been denoted with the same symbol in literature. Note also, that if distribution smoothing is not used, you should nevertheless enter a dummy value in this column. This can be for example zero (or any positive number). **Zonation will not run if this column absent from the biodiversity list file!**

- *How to disable smoothing for individual features*. An α value of 0 for a biodiversity feature means that distribution smoothing is not applied to that feature. This makes it possible to include in the same analysis both raw distributions and connectivity transformed distributions simultaneously. In order to do so, enable distribution smoothing in the project/command line, and use a non-zero α only for the features for which distribution smoothing should be applied.

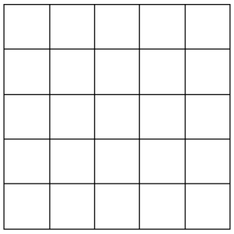
Columns 3 and 4 together define either Boundary Quality Penalty (BQP) or Neighborhood Quality Penalty (NQP, in directed connectivity) settings for the feature, depending on which one of the two options is used. The information of these two columns is only used if BQP or NQP is included to the analysis. Note that a value should be entered in these columns even if BQP or NQP is not used. They can be any positive number (e.g., a dummy values of 1). **Do not leave the columns empty!**

3. When using BQP, this parameter indicates the row number in the BQP definitions file linking the feature to a correct penalty curve. When using NQP, this parameter serves the same purpose, but only for **upstream connectivity**. Thus, in NQP, this number links the feature to a penalty curve that describes how the value of the focal planning unit changes when other planning units are lost upstream from the focal planning unit.

4. When using BQP, this parameter gives the feature-specific buffer size (number of cells). The buffer size indicates the area around the focal cell in which any fragmentation (removal of cells) influences the quality of the focal cell. For a species with a large home range, the buffer size should be larger; for a species with a small home range, a smaller buffer size is adequate.



buffer size = 3

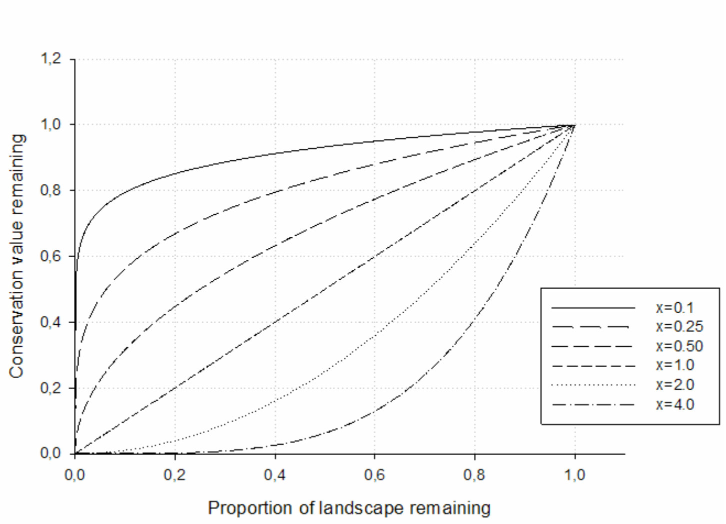


buffer size = 5

When using NQP, this parameter indicates the row number in the BQP definitions file linking the feature to a correct penalty curve for **downstream connectivity**. Thus, when running NQP, every feature has two penalty curves, one for upstream connectivity and another for downstream connectivity. As the NQP option uses planning units instead of singular cells, no buffers are needed. The connectivity of separate planning units (i.e., which ones are upstream or downstream), is defined in the directed connectivity layer.

Column 5 is used when using the additive benefit function, target-based planning, or the generalized benefit function as your cell removal rule. With core-area Zonation, a dummy number (e.g., 1.0) can be used.

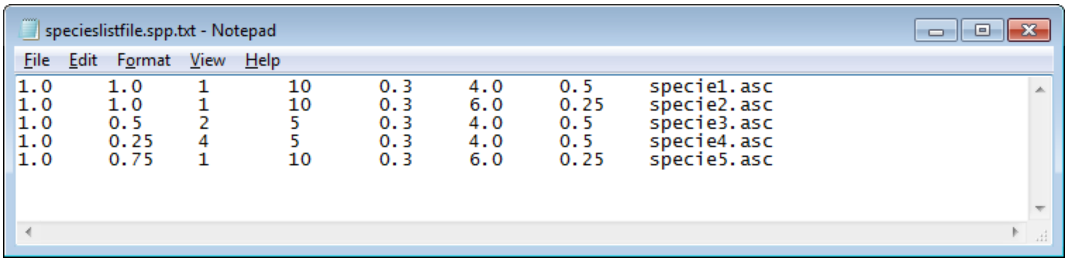
5. This column has three functions depending on which cell removal rule is used.
    - If the **additive benefit function** is used as the cell removal rule (see section 2.3.2 for methodological details), this parameter is the exponent (x) of the feature-specific power function (rjx) that translates representation to value. The power function determines the rate of loss of conservation value from the remaining landscape as cells are removed. The exponent can be any positive number, but zero is not a valid value.



Graph of power functions with differing x-values

- If **target-based planning** is used as the cell removal rule, this parameter determines the target proportion (from 0.0 to 1.0) of the feature distribution that is required in the final solution. *For a negatively weighted feature, this is still the proportion of the feature that should be included in the solution. However, Zonation will remove all but that proportion as soon as possible. For example, if you wish to remove 80% of a negatively weighted feature, the target should be set as 0.2.*

- If the **generalized benefit function** is used as the cell removal rule, four extra parameters are needed for each feature. Essentially, the final numerical column (the single ABF/TBF parameter) is split into four numerical columns, which in column order correspond to variables w_2, T_j, x, and y, in the two-piece power function. These parameters are used to determine the shape of the function, as explained in section 2.3.4. The fifth parameter, w_1, is the ordinary weight given for the feature as the first column in the biodiversity feature list file.



```
1.0   1.0   1   10   0.3   4.0   0.5    specie1.asc
1.0   1.0   1   10   0.3   6.0   0.25   specie2.asc
1.0   0.5   2   5    0.3   4.0   0.5    specie3.asc
1.0   0.25  4   5    0.3   4.0   0.5    specie4.asc
1.0   0.75  1   10   0.3   6.0   0.25   specie5.asc
```

Picture of the species list file when using the generalized benefit function as the cell removal rule.

Note that a value should be entered in the fifth column even if benefit functions or target-based planning are not used. This can be any positive number (e.g., a dummy value of 1). **Do not leave the column empty!**

6. This column indicates the name of the biodiversity feature map file (.asc raster file, .tif, etc.). Remember to type the correct path in front of the file name if the distribution maps are saved in a different directory than the biodiversity feature list file. Note that if the GBF is used, the contents of this column is shifted to column number nine.

There should not be any empty rows at the end of the biodiversity feature list file. If necessary, you can enter comments in your biodiversity feature list file on separate rows starting with the pound symbol *(#). Remember also to use **decimal points, not commas**, in all the input files.*

Supplementary material: Tips for using the command FOR to automatically produce biodiversity feature list files

Note that the `FOR` command can be used in creative ways to automatically create biodiversity feature list files. For example, the following single command row, typed and run from the command prompt (or run from a .bat batch file),

```
FOR /L %i IN (1,1,900) DO @ECHO 1.0   1   1   1   0.25   p%i.asc >> my_spp_list.spp
```

generates a file my_spp_list.spp, which has rows and relevant parameters for files p1.asc, p2.asc, ..., p900.asc. The ">>" at the end of the command indicates redirection of output into the following file. Without the ">> my_spp_list.spp", output is shown on the screen (command prompt).

Another variant of the `FOR` command allows one to loop through a set of files using the normal wildcard file name specification:

```
FOR %i IN (species_*.asc) DO @ECHO 1.0  1  1  1  0.5  %i
```

See the help menu for the `FOR` command for further information.

3.3.2.3 Run settings file

The run settings file (.dat) contains all of the basic Zonation settings. This input file is needed only when running the program from the command prompt or batch files.

.dat files	This is the suffix used in the tutorial and examples for the run settings file. This file should be created manually; the tutorial files can be used as templates. The run settings file is technically an ascii file, but it can be created using any text editor, including notepad.

In the run settings file, each parameter must be written on a separate row. The parameter names are case-sensitive, so **it is very important** that all the parameters in the run settings file are written **exactly** as presented here. If there are errors in the spelling of parameters, the program cannot find them and will use default settings instead. The order of parameters, however, is not obligatory. If a parameter is missing from the run settings file, it will be assigned a default value. If necessary, comments can be entered on separate rows starting with the pound symbol (#). Remember also to use **decimal points, not commas,** in all the input files.

The new settings allowed by Zonation v.4 are:

General settings
```
add to edge borders between removal mask levels = 1
output weighted range size corrected richness = 1
output proportional loss ranking = 1
save vmat = vmat_all_species_and_habitats
load vmat = vmat_all_species_and_habitats
load vmat directly and risky = vmat_all_species_and_habitats.bin.gz
```

Corridors
```
[Corridor loss penalty]
use corridors = 1
```

```
strength = 0.001
minimum width = 21
#use domain layers = 1
#domain layers list file = ../../domain_list.txt
#start at top percentage = 95
#penalty formula = 1
#variant = 3
#info period = 10000
#output corridor boundaries layer at percentage = 95 90 75 50

(optional settings indicated with a '#' or comment character)
```

Arbitrary Kernels

```
[Arbitrary kernels]
use arbitrary kernels = 1
default kernel = 100
arbitrary kernels constant = 1
arbitrary kernels files prefix = tutorial_input
```

The new settings allowed by Zonation v.3.1.x and higher (in addition to those in v.2.0) are:

Groups settings

```
use groups = 1
groups file = group_list.txt
```

Landscape condition and retention

```
use condition layer = 1
condition file = condition_layers.txt
use retention layer = 1
retention file = retention_file_name
retention layers relative weight = β
```

Community analysis settings and matrix connectivity

```
[Community analysis settings]
load similarity matrix = 1
community similarity matrix file = comm_sim_matrix.txt
apply to representation = 1 OR

[Community analysis settings]
load similarity matrix = 1
connectivity similarity matrix file = matrix_file_name.txt
apply to connectivity = 1
```

Edge correction in connectivity

```
connectivity edge effect fix file = fname.tif
```

Administrative units

```
[Administrative units]
use ADMUs = 1
ADMU descriptions file = ADMU_weights_all_local.txt
ADMU layer file = HV_ADMUs.tif
ADMU weight matrix = ADMU_weights_matrix.txt
calculate local weights from condition = 0
```

```
ADMU mode = 1
Mode 2 global weight = 1.0
row count for per ADMU output curves = 0
```

Analysis area mask

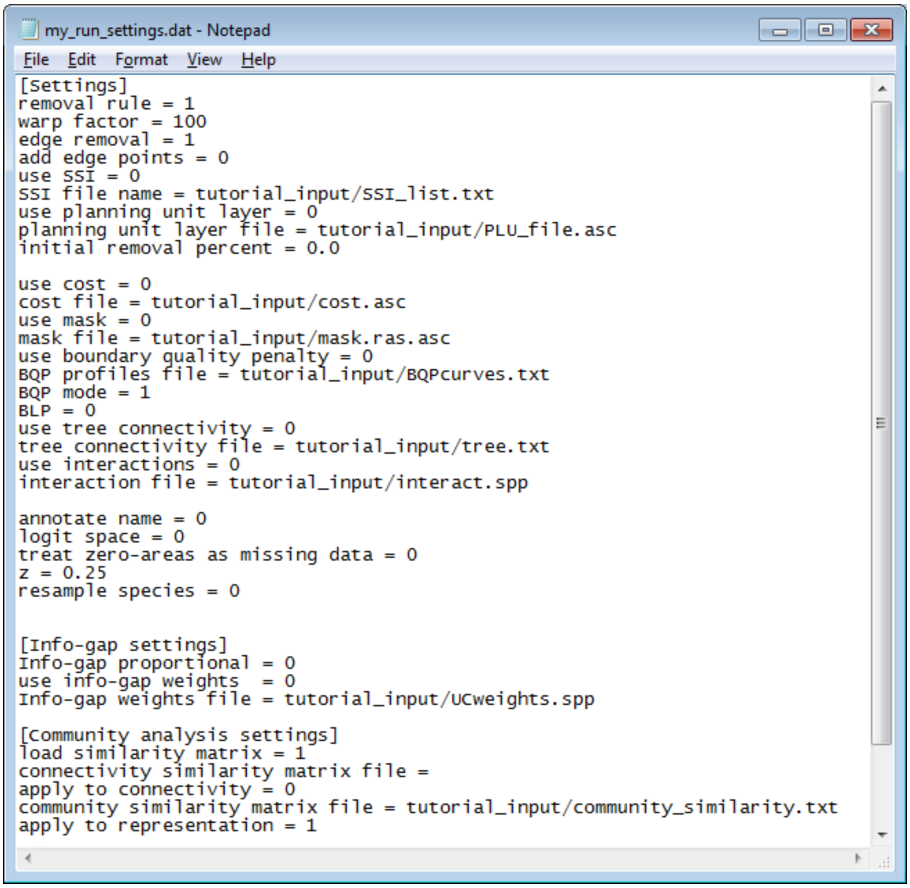
```
mask missing areas = 1
area mask file = file_name
```

Automated post-processing

```
post-processing list file = ppa_list_file_name.txt
```

Transformed biodiversity feature maps/layers

```
[Transformed layers]
output final transformed layers = 1
output info-gap transformed layers = 1
output distribution smoothing transformed layers = 1
output community similarity transformed layers = 1
output condition transformed layers = 1
output retention transformed layers = 1
output matrix connectivity transformed layers = 1
output interactions transformed layers = 1
```



```
[Settings]
removal rule = 1
warp factor = 100
edge removal = 1
add edge points = 0
use SSI = 0
SSI file name = tutorial_input/SSI_list.txt
use planning unit layer = 0
planning unit layer file = tutorial_input/PLU_file.asc
initial removal percent = 0.0

use cost = 0
cost file = tutorial_input/cost.asc
use mask = 0
mask file = tutorial_input/mask.ras.asc
use boundary quality penalty = 0
BQP profiles file = tutorial_input/BQPcurves.txt
BQP mode = 1
BLP = 0
use tree connectivity = 0
tree connectivity file = tutorial_input/tree.txt
use interactions = 0
interaction file = tutorial_input/interact.spp

annotate name = 0
logit space = 0
treat zero-areas as missing data = 0
z = 0.25
resample species = 0


[Info-gap settings]
Info-gap proportional = 0
use info-gap weights  = 0
Info-gap weights file = tutorial_input/UCweights.spp

[Community analysis settings]
load similarity matrix = 1
connectivity similarity matrix file =
apply to connectivity = 0
community similarity matrix file = tutorial_input/community_similarity.txt
apply to representation = 1
```

Example run settings file.

`[Settings]`

`removal rule`	Determines which cell removal rule (section 2.3) will be used. 1=Basic core-area Zonation, 2=Additive benefit function, 3=Target-based planning, 4=Generalized benefit function, 5=Random removal. Default=1.
`warp factor`	Defines how many cells are removed at a time. If warp factor is 100, it means that 100 cells are removed at each iteration. Thus, a lower warp factor leads to a finer solution, but a prolonged running time. A high warp factor keeps the running time short, but might result in a more coarse solution. If the warp factor is more than 1% of the remaining cells, then only 1% is removed. For example, if there are only 100 cells remaining in the landscape, then only one cell can be removed regardless of what the warp factor is. Note that if planning units (PLU; section 3.3.3.11) are used the warp factor is automatically set to 1. If planning units are not used, the warp factor can be defined freely. In our tests, having a warp factor of 100 has had little influence on the solution compared to lower warp factor values, but the run times have been considerably shorter. We recommend using a warp factor of 1 mainly for the final runs, if run times allow it. The effects of different warp factors can be compared with Landscape comparison. Default=1.
`edge removal`	Determines whether the program removes cells from the edges of remaining landscape (value=1) or anywhere from the landscape (value=0). Note that setting this parameter to 0 will increase the running times with large landscapes. Default=1. This option is ignored if running in re-loading (-l) mode.
`add edge points`	Randomly selects additional cells inside the landscape that will be initially classified as edge-cells from which removal can proceed. The value of this parameter determines the number of cells that are selected. This parameter allows a compromise between using and not using edge removal. Default=0. This option is ignored if running in re-loading (-l) mode.

When adjusting the settings, it is good to understand the function of the **edge removal** and a**dd edge points** options. The main benefits of using edge removal is that it keeps computation times short and, to some extent, increases the connectivity of high quality habitats in the landscape structure. Hypothetically however, this option can have downsides in cases where a large area of "poor" habitat is completely surrounded by "good" habitats, and the Zonation program should first remove all the good habitats from the edge to reach the poor area. Naturally, by not selecting the edge removal option the program would easily find all poor habitats deep inside the landscape, but with the cost of lost structural connectivity and increased computation times. **To prevent losing any valuable areas while keeping the computation times short, using the add edge points option together with edge removal is recommended.** By adding hypothetical edge cells into the landscape the program can spot any larger poor areas surrounded by good habitats without the risk of removing valuable cells. Cities and other poor quality areas can also be masked out to provide Zonation with sufficient low-quality edge from which ranking can naturally proceed.

`use SSI`	Determines whether Species of Special Interest (SSI) are included in the analysis (value=1) or not (value=0). The distribution data of these species is not given as maps but rather as a list of single point occurrences. Default=0.
`SSI file name`	Similar to the biodiversity feature list file, except this indicates a file that contains the list of SSI species used in the analysis (see section 3.3.3.1). Default is that SSI is not used.
`use planning unit layer`	Determines whether a planning unit layer (section 3.3.3.11) is used (value=1) or not (value=0). With this option cells are grouped into defined planning units which are then removed as a whole during the landscape ranking process.
`planning unit layer file`	Indicates the planning unit layer file (integer grid) to be used.
`use cost`	Determines whether land costs are included in the analysis (value=1). If no land costs are used, this parameter should be set to 0. Default=0.
`cost file`	Indicates the land cost file (section 3.3.3.6) to be used. Default is that a cost file is not used. Note that this cost layer is used in cost-efficiency analysis where the conservation value of a cell is divided by local cost. If multiple opportunity costs are to be considered, these should be handled as negatively weighted grid layers (see section 5.3.4).
`use mask`	Determines whether a (hierarchical) removal mask layer (section 3.3.3.9) is used (value=1) or not (value=0). Default=0.
`mask file`	Indicates the (hierarchical) mask layer file (integer grid) to be used. Default is that a mask file is not used.
`add to edge borders between removal mask levels`	Normally, when using "edge removal", the list of edge cells of the landscape is defined by borders between missing and non-missing cells. With this new option, the borders between regions with different levels in the removal mask are added to the list of edge cells as well. The default value is 1. When using "hierarchical removal mask" ('use mask', 'mask file'), borders between lower and higher removal levels are added to the edge list. Before, lower level cells could be removed too late (potentially after many higher level cells) if they were enclosed in

	higher level regions. Note: when using "edge removal" the run time of Zonation is directly related to the length of the edge list. This new option has been added in Zonation 4 and it may cause a significant run time increase for large analyses with intricate hierarchical removal masks, especially when the borders between removal levels are long. If you are completely sure that the hierarchical removal mask does not have any influence disabling this option may provide a significant speedup. This might be safe in cases where the hierarchical removal mask does not define any region of lower priority enclosed in regions of higher priority.
`use boundary quality penalty`	Determines whether BQP connectivity is used (value=1) or not (value=0). Use of BQP leads to solutions that include structural aggregation at scales relevant for individual features. Default=0.
`BQP profiles file`	Indicates the BQP profiles file (section 3.3.3.2) to be used. Default is that a BQP profiles file is not used.
`BQP mode`	Determines how the program will calculate the effects of fragmentation from feature distribution data. Essentially, this parameter tells the program what type of feature distribution layers are being using. The mode can be either 1 or 2 depending on your data (see section 3.3.3.2). Default=2.
`BLP`	Defines a penalty given for the boundary length of the reserve. The value of BLP should be a small decimal number. The value of this penalty cannot be decided in advance based on some numeric criterion; rather a suitable value needs to be found by experimentation. Use of BLP leads to a more aggregated solution, and it is desirable that only a small loss of representation follows from the use of BLP. Try first a small value (e.g., 0.01), to perceive the effect of BLP on the solution. **When including BLP in the analysis, it is preferable to use a warp factor of 1**. If BLP is not used, this parameter should be set to 0. Note potential difficulties in interpretation of results if multiple aggregation methods are used simultaneously. Default=0.
`use tree connectivity`	Determines whether directed connectivity (Neighborhood Quality Penalty; NQP) is used (value=1) or not (value=0). Unlike the other aggregation methods, NQP models a directed connectivity measurement up and down a linked system of planning units, which would typically model hydrologically linked water catchments. Note that planning units always need to be used together with NQP. Default=0.
`tree connectivity file`	Indicates which tree connectivity file (section 3.3.3.3) will be used when including NQP in the analysis. Default is that the file is not used.

`use interactions`	Determines whether [ecological interactions]() (section 2.6) are included (value=1) in the analysis or not (value=0).
`interaction file`	Indicates which [interactions definitions file]() (section 3.3.3.8) will be used.
`annotate name`	This option will mark the output file names to show which analyses have been used to produce them (value=1). The program will add letters and numbers in the middle of the output file name depending on the analyses used: • CAZ_/ABF_/TBF_/GBF: Shows whether basic core-area Zonation (CAZ), the additive benefit function (ABF), the target-based function (TBF), or the generalized benefit function (GBF) have been used as the cell removal rule • M: Mask used • C: Costs used • E: Edge removal used • A: Edge points added • Sxxx: Distribution smoothing used. The following numbers (indicated here with 'xxx') show the factor that has been used to multiply the feature-specific α –values. Note that for output, the factor has been multiplied by 100. Thus using factor=1.0 would result a suffix of S100, factor=0.1 results peciesa suffix of S10, etc. • IGxxx: Uncertainty analysis included. Again the following numbers show the info-gap α –value multiplied by 100. • BQP: BQP included • BLPxxxx: BLP included. The following numbers show the penalty given for the boundary length multiplied by 1 000 (i.e., using BLP=0.5 results a suffix of BLP500, BLP=0.05 results a suffix of BLP50, etc.). Default = 1.
`logit space`	A special option relevant for probability of occurrence models using logistic link functions. Determines whether the biological values of cells will be transformed from logit space (value=1) for processing. In this case the raster files (.asc files) should contain the values of the linear predictor part of a logistic link function. If data is not to be transformed from logit space, this parameter should be set as 0. Default=0.
`treat zero-areas as missing data`	This option changes all cells with no feature occurrences to missing data. This function might be useful in some cases, for example, if the missing data is in fact marked with the value 0 in your feature distribution files due to some technical reasons. Note however, that there is a fundamental biological difference between features not occurring somewhere (value 0) and not having any information from that same place (missing data). Thus, use this option with care. The use of this option does not change the input files in any way (thus cells with value 0 will remain as they are)—it will only change the way Zonation interprets the files. Default=0.

`z`	This value is used to calculate the extinction risks of features as their distribution sizes are decreasing. z is the exponent of the species-area curve ($S = cA^Z$), which has been widely used in ecological studies. In theory, z can have any positive value, but a commonly found empirical value is 0.25. If using an exponent equal to *z* in an ABF analysis, then Zonation is essentially minimizing the SA curve predicted extinction risk across features. Default=0.25.
`resample species`	This option tests analyses using only a subset of features. The program selects a random set of features from the biodiversity feature list file and uses them to run the analysis. Thus several analyses can be conducted to see how the selection of features influences the outcome. Note that the random sets do not include multiple selections of one features—all features in the set are different ones. To use this option enter the number of features you wish to include in one set. If this value is zero, any negative value, or equal to the total number of features, no sampling is done. Default=0.
`post-processing list file`	Indicates the list file of automatically executed post-processing analyses (sections 3.3.3.17 and 3.5.1) to be performed after landscape ranking. Default is that automated post-processing is not used.
`memory save mode`	This option uses Zonation in a memory save mode (value=1). This can be useful if the memory capacity of the computer is too small to complete the analysis. In other words, using this option allows slightly bigger analyses to be done. However, the best solution for memory problems is to use a computer with more memory. If this value is set to zero, Zonation will not operate in the memory saving mode. Default=0.
`use groups`	Determines whether a groups file (section 3.3.3.12) is used (value=1) or not (value=0). A groups file is needed for grouped output and also when condition or retention is used. Default=0.
`groups file`	Indicates the groups file to be used in the analysis. Default is that the file is not used.
`use condition layer`	Determines whether feature group-specific landscape condition (section 2.10) is used in the analysis (value=1) or not (value=0). Default=0.
`condition file`	Indicates the file describing linkage of features to landscape condition (section 3.3.3.14).
`use retention layer`	Determines whether feature group-specific retention layers (section 2.10) are used in the analysis (value=1) or not (value=0). Default=0.
`retention file`	Indicates the file describing linkage of features to landscape retention (section 3.3.3.15). Default is that a retention file is not used.

`retention layers relative weight`	Determines the relative weight of retention layers (as a group) compared to representation (normal grids not transformed for retention). Default=1.0.
`mask missing areas`	Determines whether some areas of the landscape are masked out (filled with missing data) for all input grid layers (value=1) or not (value=0). This option is useful if only a subregion of the landscape would be analyzed and one does not wish to redo all input grids. Default=0.
`area mask file`	Indicates the raster file to be used for masking the areas with missing information (section 3.3.3.10). Default is that an area mask file is not used.
`output weighted range size corrected richness`	Indicates whether to generate the weighted range size corrected richness output map (see section 3.4.1). Default: 1 (generate map). Disabling this option can be convenient when you run different variants of a setup with the same set of biodiversity features.
`output proportional loss ranking`	Indicates whether to generate the proportional loss ranking output map (see section 3.4.1). Since Zonation v.4 this is disabled by default. Default: 0 (do not output map).
`[Info-gap settings]`	This title in brackets is obligatory before the info-gap settings.
`Info-gap proportional`	Determines whether the errors in feature occurrences are uniform errors (value=0) or proportional errors (value=1). Uniform error is the default setting and works for most of the data sets, but in some cases it is more appropriate to use proportional errors (see e.g. Ben-Haim 2001).
`use info-gap weights`	Determines whether feature-specific distribution uncertainty map layers (section 3.3.3.7) are used in the info-gap analysis (value=1) or not (value=0). Default=0.
`Info-gap weights file`	Indicates the file that includes the list of uncertainty maps (that correspond to feature grids; section 3.3.3.7). Default is that the file is not used.
`[Community analysis settings]`	This title in brackets is obligatory before the community analysis settings. Note: the capability to use the following community-level analyses is available in Zonation v.3.1 and higher.
`load similarity matrix`	Determines whether a similarity matrix for connectivity or community composition (section 3.3.3.4) is used in the analysis (value=1) or not (value=0). Default=0. If the parameter is set to 1, either similarity expansion for representation, or similarity in connectivity, or both can be accounted for, as specified by the following parameters.

`connectivity similarity matrix file`	Indicates the connectivity similarity matrix file (section 3.3.3.4) to be used. Default is that a similarity matrix file is not used.
`apply to connectivity`	Determines whether the similarity matrix is applied to connectivity (section 2.4.5) in conservation area prioritization (value=1) or not (value=0). Default=0.
`connectivity edge effect fix file`	Indicates the file for edge effect fix for connectivity (section 3.3.3.5). Default is that an edge fix file is not used.
`community similarity matrix file`	Indicates the community similarity matrix file (section 3.3.3.4) to be used. Default is that a similarity matrix file is not used.
`apply to representation`	Determines whether community similarity expansion (section 2.8) is applied to the stack of biodiversity feature grids. (value=1) or not (value=0). Default=0.
`[Administrative units]`	This title in brackets is obligatory before the administrative units settings. Note: the capability to use the following administrative units settings is available in Zonation v.3.1 and higher.
`use ADMUs`	Determines whether administrative units (section 2.12) are accounted for in the analysis (value=1) or not (value=0). Default=0. If 1, all information below needs to be given for the analysis to succeed.
`ADMU mode`	Determines whether the administrative regions are weak (value=1) or strong (value=2) See section 2.12 for details. Default=1.
`ADMU layer file`	Indicates the administrative units layer (integer grid; section 3.3.3.16) to be used. Compulsory if ADMUs are used.
`ADMU descriptions file`	Indicates the file describing weights, etc. for the administrative units (section 3.3.3.16). Compulsory if ADMUs are used.
`ADMU weight matrix`	Indicates the file containing a matrix of administrative units x feature weight (section 3.3.3.16). Compulsory if ADMUs are used.
`Mode 2 global weight`	Indicates balance between global representation and local (ADMU-specific) considerations in landscape ranking. When ADMU mode 2 is used, this is the weight given to global weights of biodiversity features, from the perspective of the global decision maker. The value can vary between 0 and 1. Default=0.5.
`row count for per ADMU output curves`	Indicates the number of rows for each of the per ADMU curve files, that is, the number of removal levels reported. The files contain administrative unit specific results, as described in section 3.3.3.16.

	Default value is 0, meaning that ADMU specific files are not generated.
`[Transformed layers]`	This title in brackets is obligatory before the transformed output layer settings. This feature is available in Zonation version 3.1.2 and higher.
`output * transformed layers`	Indicates which layers will be transformed. * can be any of: final, info-gap, distribution smoothing, community similarity, condition, retention, matrix connectivity, and interactions. If enabled, different types of transformed layers will be generated in their respective output folders/directories. See section 3.4.4 for more details.
`[Corridor loss penalty]`	This title in brackets is obligatory before the corridor building settings. This feature is new since Zonation version 4.
`use corridors`	Enables the corridor building mechanism of Zonation. Default value: 0. Unless this option is set to 1, all other options related to corridor connectivity (see below) will be ignored.
`strength`	Strength or emphasis on corridor building. Real number bigger than 0 (0 would effectively disable corridor building). Appropriate values are application dependent, and finding them requires exploration of different alternatives in a way similar to the boundary length penalty (BLP).
`minimum width`	The minimum width of corridors expressed as a number of cells (integer number). Connections of width lower than this parameter will not be considered effective corridors. This is also an application dependent parameter. Default value: 0 (disabled). This parameter must be set as an integer value, but at present only odd values are supported, if you specify an even number it will be rounded down.
`use domain layers`	Enables the use of domain layers. Default: 0 (disabled). If enabled, the list of layers will be loaded from the file specified in the next option.
`domain layers list file`	Name of the file that gives the list of corridor domains. There is no limit to the number of different corridor domain layers (as long as enough RAM memory is available). The format of this file is similar to the condition and retention file: two columns, with the first column specifying a weight (real number) and the second file giving the name of a raster map (domain layer).
`start at top percentage`	This parameter defines the percentage of landscape remaining at which the corridor penalty mechanism starts working. This makes it possible to start the ranking process without corridor building until a certain fraction of the landscape has been prioritized (removed). This implies that corridors will not be built through the lowest quality areas (without considering the corridor loss penalty). Default: empty (equivalent to 100 which means that the corridor building process is in effect throughout the whole ranking process). Using this option is not recommended if you want to explore all the possibilities for

	corridor connectivity across the landscape, including low quality areas.
`output corridor boundaries layer at percentage`	Zonation can generate output raster maps with the boundaries of corridors. This requires setting one or more remaining percentages of the landscape. Multiple layers can be generated, as corridors are generated during the ranking process. Example value: `output corridor boundaries layer at percentage = 95 90 75 50` which will output four maps, with the boundaries of corridors identified at the 95%, 90%, 75%, and 50% top fractions of the landscape, respecitvely. Default: empty (do not generate corridor boundary maps).
`penalty formula`	This option chooses the exact penalty rule (formula) to use. We recommend using the default at first. Default: 1. Possible values for the fast/approximate variant: 1 (use richness), 2 (use size). Possible values for the exact/slow variant: 1-3 (based on richness), or 11-13 (based on size). 1 and 11: second maximum richness/size after split divided by initial richness/size. 2 and 12: initial richness/size. 3 and 13: minimum richness/size after split divided by initial richness/size.
`variant`	With this parameter you can choose between two variants in the way the corridor loss penalty is calculated. The first option corresponds to rules (formulas) that use the (square) inverse distance from the split point to the center of the original network. This is the default and fastest option but may not be suitable in all applications. More sophisticated formulas can be used in the second (or exact) variant. In this second case, the penalty formula can use the size or richness of the sub-networks created after splitting a network. Default: 1 (approximate/fast). Alternative value: 2 (exact/potentially slow).
`redundancy radius x` `redundancy radius y`	These options make it possible to specify a maximum redundancy radius beyond which redundant corridor connections between networks or patches will not be considered duplicated or redundant (the corridor loss penalty does not apply to redundant connections). Redundant connections that are farther away than this parameter (on the x or y axis) will be maintained (as far as the strength of the corridor loss penalty enforces). Both options are specified as a number of cells, and they do not need to be the same value. x refers to the x axis (longitude) and y to the y axis (latitude). Default: empty (no redundancy radius limit).
`info period`	When this option is enabled, an additional output file (with suffix) is generated with further information about the corridor building process. Default: 0 (disabled). The value given (bigger than 0) indicates the output period in number of cells ranked/removed. Note that if you use a small number here Zonation will potentially produce huge text files with tons of information. You probably do not need to use this feature, unless you are performing very advanced tasks.

`[Arbitrary kernels]`	This title in brackets is obligatory before the arbitrary kernel settings. This feature is new since Zonation version 4.
`use arbitrary kernels`	Determines if arbitrary kernels as specified in the groups file is applied. If this option is off, the default (negative exponential decrease) kernel will be applied regardless of the kernel numbers given in the fifth column of the groups file.
`default kernel`	Redefine the default kernel (to be applied unless otherwise specified). Normally, the default kernel is exponential decay. If this option is set to 4 for example, the triangular kernel (4) will be applied for every distribution smoothing, interaction, connectivity matrix, and community similarity matrix transformation, unless an exception is specified in the respective column of the groups file. If the fifth, sixth, and seventh columns of the groups file are all set to -1, this option defines the kernel that is used globally (any feature and any type of transformation that uses a kernel).
`arbitrary kernels constant`	Constant that multiplies the (arbitrary) kernel functions. For example the kernel number 3 (Uniform) is defined as 1 for any distance lower than the distance (dispersal) parameter given. By using this constant (which is applied as a multiplicative factor) it is possible to redefine the amplitude of the kernel (1) to any other (positive) real value. The default value of this constant is 1 (do not modify kernels).
`arbitrary kernels files prefix`	Path or prefix where to look for arbitrary kernel files. Both absolute and relative paths are supported. For example, if the prefix is "tutorial_input", the arbitrary kernel number 100 will be loaded from "tutorial_input/zonation_arbitrary_kernel_100.txt". As an example of absolute path, if the prefix given in this option is "C:\akernels", the arbitrary kernel number 100 will be loaded from "C:/akernels/zonation_arbitrary_kernel_100.txt".
`Options to save and (re)load feature lists`	
`save vmat`	If this option is enabled (giving the name of an output file), after loading all the features given in the features list file, Zonation will output a binary file that contains all the information loaded into memory. The name of the file is derived from the Zonation general output name by adding the value given here as a suffix. This file can be used later on in the next Zonation runs to accelerate the loading process, as you would not need to load again a potentially large set of raster maps. The content of this file is called 'vmat' in Zonation. Default: empty (disabled). This option is normally used the first time that a particular list of features is used in a Zonation setup or run. The next times that the same list is used you can accelerate the loading process by using the option `load vmat.`
`load vmat`	Re-load a binary file generated with the option `save vmat.` By default, the name given is used as a suffix to the Zonation output name (as with most of the output files). This requires some fiddling with files.

`load vmat directly and risky`	This option modifies the behavior of the `load vmat` option. Default: 0 (disabled). If you enable it (value 1), the file name given to the option `load vmat` will be used as the path to the vmat binary file to load. This can be a more comfortable and apparently easy mode, but it is error prone especially if you are not familiar with the vmat saving/loading process. Use this option at your own risk.

3.3.3 Optional files

These files are needed only when certain options are used.

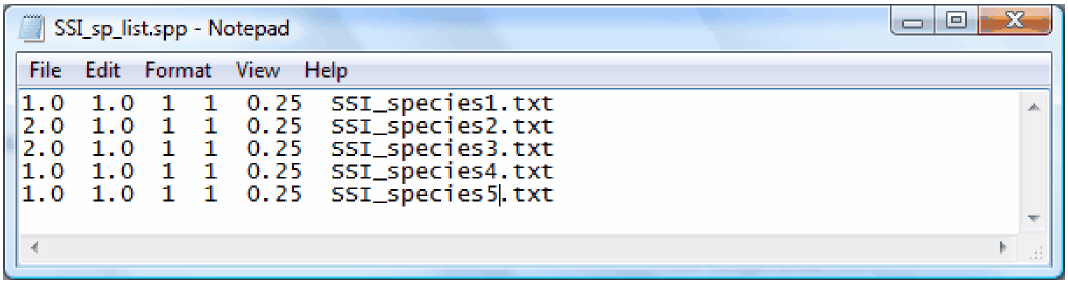
3.3.3.1 SSI list and coordinates

The Species of Special Interest (SSI species) are the second kind of feature occurrence information that can be entered into Zonation. The input for an SSI species or feature is a list of observation locations instead of a map. The SSI input can be used for a species or feature that either (1) has so few observations that the full distribution cannot be modeled or (2) has been completely surveyed and all occurrence locations are known. The idea of the SSI analysis is that if the species or feature occurs in just a few locations, it is wasteful to enter a full map for it—a one million element grid map takes ~3000 times as much memory to store as does a list of 100 observation locations + population sizes. Consequently, a very high number of SSI features can be analyzed in Zonation. Ordinary 'map' features and SSI features can be mixed in the same analysis; however, it is not currently possible to run Zonation only with SSI features. To conduct an observed-distribution-only analysis, one can just enter one zero-weighted map of the landscape and all of the rest of the features as point distributions.

SSI features or species are treated exactly as map features in the Zonation process: the marginal loss following the removal of a cell is based on the fraction of the distribution of the feature residing in the cell. However, there is the difference that distribution smoothing and boundary quality penalty connectivity methods do not operate on SSI features. Connectivity requirements for an SSI feature can be implemented indirectly by entering "buffering" locations for the SSI feature around the actual occurrence locations. Overall, it can be expected that full distributions of SSI features will be retained far into the cell removal process, especially if there are relatively few locations with observations of these features.

To include SSI features in an analysis, an SSI feature list file and a feature-specific coordinate file must be included as input files.

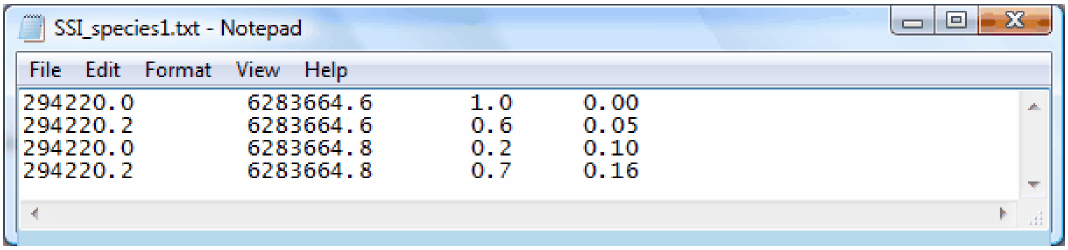
The SSI feature/species list file has an identical structure as the ordinary biodiversity feature list file. However, it should be understood that the columns for the dispersal alpha and BQP parameters contain dummy values—whatever is entered there will not influence computations. Thus, the relevant columns for an SSI feature are: the first column which defines the weight of the SSI feature, the last numeric column which gives the parameter either for the additive benefit function or for targeting analysis, and the last column indicating the coordinate files.



```
1.0  1.0  1  1  0.25  SSI_species1.txt
2.0  1.0  1  1  0.25  SSI_species2.txt
2.0  1.0  1  1  0.25  SSI_species3.txt
1.0  1.0  1  1  0.25  SSI_species4.txt
1.0  1.0  1  1  0.25  SSI_species5.txt
```

Example of an SSI feature list file.

The second set of files are the feature-specific coordinate files (one file for each feature) that give the exact locations for each observation. An SSI feature distribution file has four columns: (1) x coordinate of the observation point, (2) y coordinate of the observation point, (3) the biological value of the record (this can be population size or other measure of site suitability for the feature in question), and (4) the info-gap relative error measure, with exactly the same interpretation and operation as for the map species.The given coordinates must fall within the area of the maps loaded for map features (as defined in the biodiversity feature distribution map files for ordinary map features) or otherwise an error will be reported. If uncertainty analysis is used, distribution discounting will be applied to the population size (or to any other form of information) given for the location. Note that the fourth column can be omitted for SSI features. If this is the case, the uncertainty error measure will be taken as zero, and any uncertainty analysis will not influence population sizes given in column number three.



```
294220.0        6283664.6       1.0     0.00
294220.2        6283664.6       0.6     0.05
294220.0        6283664.8       0.2     0.10
294220.2        6283664.8       0.7     0.16
```

Picture of the feature-specific coordinate file. The first two columns give the x and y coordinates of the record. The third column shows the biological value of that record (any non-negative integer or decimal value) and the last column is the relative error measure.

Run settings for using SSI features

When including SSI features in your analysis, remember to type "`use SSI = 1`" (SSI option selected) and "`SSI file name = my_SSI_list.txt`" (name of the SSI feature list file) in the Run settings file.

Output with SSI features

Numbers for the mean and minimum representation curves are given in a special .SSI_curves.txt output file, which is produced together with the basic output files when SSI features are included in the analysis. Since Zonation v.4 this file also includes the individual proportions of remaining distributions for every SSI feature (as in the normal .curves.txt file).

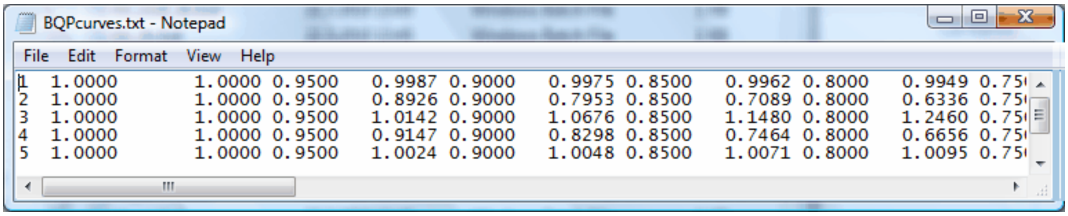
In this file there is also information about the level of landscape removal when the last occurrence of that SSI feature is removed. Output for SSI features is given in the "Plots" tab in the GUI Runtime viewer, where a graph displays the minimum and mean fraction retained across all SSI features. Locations with SSI observations are shown in red in the [Maps]() window. It is worth checking that the SSI locations display correctly, as errors in coordinates might otherwise easily go unnoticed.

3.3.3.2 Boundary quality penalty definitions file

To include boundary quality penalty ([BQP](); for a description of the method see section 2.4.3) in an analysis, the following considerations need to be made:

1. Create a BQP definition file which contains all penalty curves. This file determines different responses of biodiversity features to habitat fragmentation.
2. Link all features to the correct penalty curve by entering the correct row number of the respective curve in the third column of the biodiversity feature list file. Multiple features can, and commonly will, link to the same response curve.
3. Give a suitable buffer size (in cells!) for each feature in the biodiversity feature list file. The buffer size indicates the area in which any habitat loss and fragmentation will influence the biological value of the focal cell for that particular feature.
4. Decide how Zonation will treat missing data in a BQP analysis.

The BQP definition file is a text file where different feature responses to neighborhood habitat loss are displayed as points of penalty curves with each curve on its own row.



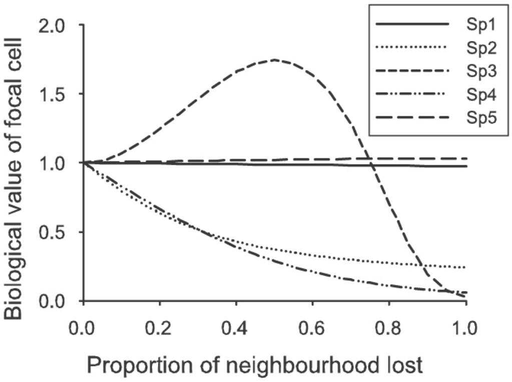
```
1  1.0000         1.0000 0.9500   0.9987 0.9000   0.9975 0.8500   0.9962 0.8000   0.9949 0.75
2  1.0000         1.0000 0.9500   0.8926 0.9000   0.7953 0.8500   0.7089 0.8000   0.6336 0.75
3  1.0000         1.0000 0.9500   1.0142 0.9000   1.0676 0.8500   1.1480 0.8000   1.2460 0.75
4  1.0000         1.0000 0.9500   0.9147 0.9000   0.8298 0.8500   0.7464 0.8000   0.6656 0.75
5  1.0000         1.0000 0.9500   1.0024 0.9000   1.0048 0.8500   1.0071 0.8000   1.0095 0.75
```

Here, the first column indicates the row number. The next two columns are the first column pair, in which the initial state (when no neighborhood habitat has been lost) is represented. The first number of the pair indicates the proportion of neighborhood habitat left and the second column indicates the fraction of the probability of occurrence retained in the focal cell. Hence, in the initial state, when no habitat has been lost yet and, therefore, the occurrence level of the focal cell has not yet changed, the two parameters always have a value of 1.000. The following column pairs describe the loss of neighborhood habitats and effects that this has on the occurrence level in the focal cell. Note that the change in probability of occurrence can be either negative (retained < 1) or positive (retained > 1) depending on biodiversity feature preference to fragmented habitats.

The column pairs should be considered as x,y -coordinates on a penalty curve. To draw a penalty curve, or any curve at all, it is necessary to have at least two points. Thus, in the BQP definition file at least two column pairs are needed. The two points could for example be the initial point, when no habitat has been lost and the final point, when all the habitat has been lost. For example, for feature A the two points could be (1.000 1.000) and (0.000

0.500), meaning that when all the neighborhood habitat has been lost, the biological value of the focal cell for that feature has decreased by half. Note that each of the penalty curve rows can contain a maximum of 20 points!

These BQP functions can be defined based on statistical analysis of habitat models or on expert knowledge. See Moilanen and Wintle (2007) for more information.

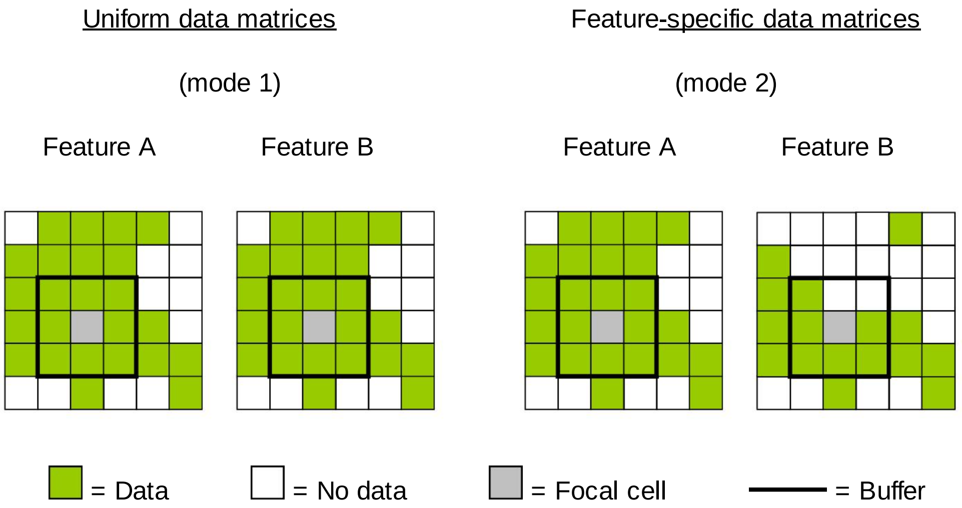


An example of different feature-specific penalty curves (redrawn from Moilanen & Wintle 2007). Note that the curve increasing over value 1 indicates a feature that prefers semi-fragmented habitats.

Boundary quality penalty can be included in two alternative modes. The choice will depend on the distribution grids of biodiversity features. BQP mode is defined in the run settings file.

Mode 1 indicates that the data/no data matrix in all biodiversity feature distribution map files should be uniform and aligned, and that there are no differences between features in terms of which cells are considered potential habitat, and which are then used in BQP buffer calculations. In other words, all features would be dependent on the same general habitat type, such as forest. With mode 1, Zonation automatically aligns missing data if different feature layers happen to have missing data at different locations. When aligning data, if feature A has missing data at location (x,y) where any feature B has a positive occurrence, then the missing value for feature A is replaced by a zero-level occurrence.

Mode 2 indicates that the data/no data matrixes are not uniform and aligned, and that the program needs to calculate feature-specific buffers for each feature/cell separately. Mode 2 is more realistic in the sense that fragmentation/loss in habitats that are not suitable for the feature will not influence the value of the focal cell. But, mode 2 also requires longer computation times due to more complicated feature-specific calculations. Also, use of mode 2 at least doubles the memory usage of Zonation, thus decreasing the number of features that can be run in one analysis. Mode 1 is a preferable when all features use approximately the same habitat type.



Run settings for using BQP in an analysis

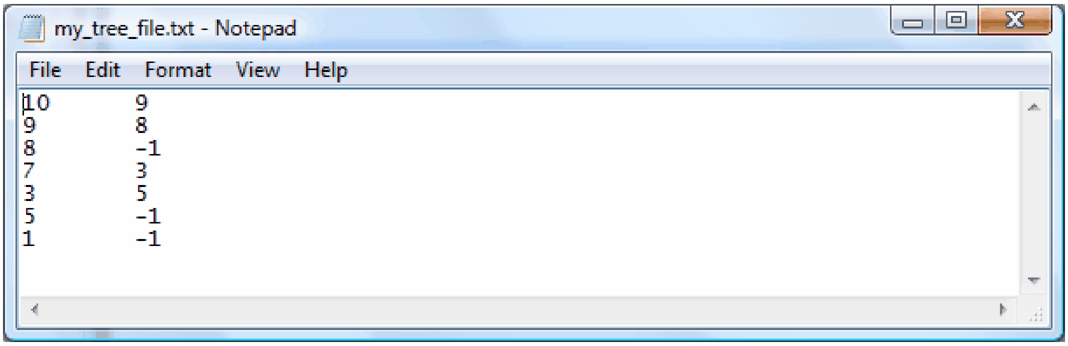
To run BQP from the command line, the following lines are needed in the run settings file:
`use boundary quality penalty = 1`
(option selected)
`BQP profiles file = myBQPdefinifionsfile.txt`
(the name of the BQP definitions file)
`BQP mode = 1 OR 2`
(depending on how Zonation will treat missing data about features.)

3.3.3.3 Directed connectivity layer

Directed connectivity is similar to the boundary quality penalty in the sense that the loss of habitat in surrounding cells influences occurrence levels in focal cells. However, rather than assuming a circular neighborhood, directed connectivity assumes that connectivity is strictly directed, such as in riverine systems. See section 2.4.4 for more information. To include directed connectivity (NQP) in an analysis, the following considerations need to be made:

1. Include a planning unit layer (section 3.3.3.11) to identify which cell belongs to which planning unit.
2. Create a directed connectivity definitions file describing the linkage between planning units.
3. Give both **upstream and downstream** connectivity responses for all biodiversity features. These responses are defined in a file that is identical to a BQP definition file (section 3.3.3.2) with the distinction that instead of one, every feature is linked to two penalty curves.
4. Link all features to the correct penalty curves by entering the correct row number of the respective curve in the **third (upstream) and fourth (downstream) column** of the biodiversity feature list file.

The directed connectivity definitions file is a text file that contains a description of the tree hierarchy (i.e., the linkages between planning units (PLU)). This information is entered simply as a file with two columns: a planning unit number (as given in the planning unit layer) and the number of the planning unit downstream, as in the following figure.



```
10	9
9	8
8	-1
7	3
3	5
5	-1
1	-1
```

This would be interpreted that PLU 10 flows into PLU 9, PLU 9 flows into PLU 8, and PLU 8 has no downstream component (e.g., if it flows into the sea). PLU 7 flows into PLU 3, 3 into 5, and 5 is, for example, at the root of a tree. PLU 1 is not linked to anything and it is taken as an independent entity. The ending of a linkage line is always marked as -1. Note that the planning unit numbers do not need to be consecutive. Warnings will be issued in the memo if linkage information is missing for a planning unit or if the linkage is confused (e.g., having multiple downstream connections for one planning unit). Remember that a PLU can have several upstream connections, but only one downstream connection.

In addition to the directed connectivity layer, loss functions (i.e., penalty curves) analogous to those used in the BQP technique need to be defined for each feature. As with BQP, the penalty curves represent the loss of biological value in the focal unit (here the planning unit) when neighboring units are removed. The difference from BQP is essentially that the neighborhood is not symmetric, but directional, and that for each biodiversity feature there are separate upstream and downstream response functions modeling effects of habitat loss upstream or downstream from the focal location (hence, instead of one, each feature may be linked to two penalty curves).

Note that the use of NQP also changes the interpretation of the biodiversity features list file. With NQP, there is no feature-specific radius like with BQP—the neighborhood is the set of linked planning units. Rather, instead of number of response and radius, the third and fourth columns in the biodiversity features list file are interpreted as the row number of the penalty curves for upstream and downstream losses, respectively. The curves are specified in the same input file as in BQP (see section 3.3.3.2).

Run settings for using directed connectivity in an analysis

To use directed connectivity in an analysis, adjust the run settings file to include:

1. Set `"use planning unit layer"` to 1
2. Give the name of your planning unit layer file `"planning unit layer file = my_PLUs.txt"`

3. Set "`use tree connectivity`" to 1 to indicate that NQP will be used.
4. Define the name of the directed connectivity definitions file in the run settings file with "`tree connectivity file = mytreeconnectivity.txt`".
5. Define the name of the BQP curves file by typing "`BQP profiles file = myBQPdefinifionsfile.txt`"

3.3.3.4 Connectivity and community similarity matrices

Connectivity matrix

A connectivity matrix is a file that is used in an analysis to apply matrix connectivity. The file is a numeric matrix with N columns and N rows and describes the extent to which the occurrence level of a biodiversity feature influences the connectivity of multiple other biodiversity features (see section 2.4.5). The values in the matrix describe pair-wise connectivity effects (i.e., how much feature *n* (column) influences the connectivity of feature *k* (row)). The matrix does not need to be symmetric; feature *n* may contribute more or less to the connectivity of feature *k* than the other way around. When Zonation reads in the connectivity matrix, it is applied to the first N features in the biodiversity feature list file.

Run settings for including connectivity matrix

To include the connectivity matrix in an analysis, adjust the run settings file to include the following lines:

```
[Community analysis settings]
load similarity matrix = 1
connectivity similarity matrix file = connectivity_matrix.txt
```
(the name of the connectivity matrix)
```
apply to connectivity = 1
```

Please refer to section 5.1.6 for more detailed analysis setups.

Community similarity matrix

The community similarity matrix is applied in the community level analyses. It describes the similarity between community types. Most commonly this would be the proportion of biodiversity features in common between two community types, but it can be other things as well. The values can vary between 0.0 and 1.0 and describe pair-wise similarities between community type A (in rows) and B (in columns). The similarity matrix values are used to expand the occurrence levels of the primary community types to effective occurrence levels; in other words, the extent to which features of community type A are present in type B and all the other types. The matrix does not need to be symmetric: community type A may contain more features of type B and vice versa. The dimension of the matrix does not need to be identical to the number of biodiversity features used in the analysis. When Zonation reads in a community similarity matrix with N rows and N columns, it applies it to the first N features in the biodiversity feature list file. This allows for combined community and feature level analysis (see section 5.3.2).

A community similarity matrix is applied in a different way than a connectivity matrix. A community similarity is used for transforming the actual occurrences of community types, whereas a connectivity matrix is used to transform connectivity calculations. Distribution smoothing following community similarity expansion would have a similar effect to matrix connectivity.

Run settings for applying a community similarity matrix

To include the community similarity matrix in an analysis, adjust the run settings file to include the following lines:

```
[Community analysis settings]
load similarity matrix = 1
community similarity matrix file =
  community_similarity_matrix.txt
apply to representation = 1
```

Please refer to sections 5.3.1 and 5.3.2 for more detailed analysis setups.

3.3.3.5 Connectivity edge effect fix file

A connectivity edge effect fix file can be applied in an analysis that utilizes matrix connectivity. It is a raster grid (for example, .ascii, .img, .tfi, etc.), in which cell values indicate the fraction of the cell that belongs to a habitat that does not harm connectivity. This fix can be useful, for example:

1. To account for national borders beyond which suitable habitats may continue but the connectivity appears to be lower on the edge as no data is available from the other side. In this case, cells on the other side of the national border could be marked as non-harmful base habitat. Then, it is effectively assumed that habitat outside the border will influence connectivity as habitat inside the border. If a cell has a connectivity of 2.0 that has been aggregated from a neighborhood that is only 1/3 within the country, and if cells outside the country have been marked as non-harmful base habitat, then the value of connectivity becomes 2.0/(1/3)=6.0.

2. In a situation where a mosaic of different habitat types is beneficial for biodiversity, but the distribution maps implicate that patches of different habitat types have discontinuous and patchy distributions. For example, consider the connectivity of a forest. The connectivity of a forest will necessarily be reduced at the edge of a large lake. But, there may be many cases where such an edge-effect is not desirable. If not, mark water as non-harmful base habitat in the connectivity edge effect fix file. This option could also be relevant at the border of a forest and a marshland as some species will not perceive the marshland as bad for connectivity.

The connectivity edge effect fix file corrects the connectivity values to $c'=c/(1-f)$, where f is the fraction of non-harmful habitat. The value for f is the value entered in the connectivity edge effect fix file. See section 2.4.6 for more information.

Run settings to include the connectivity edge effect fix

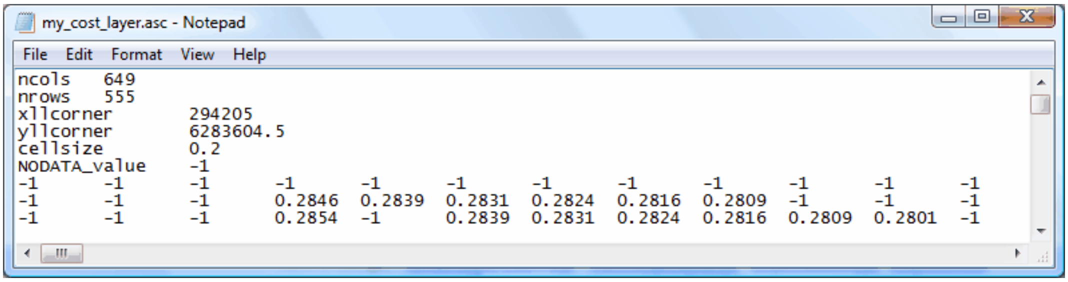
To include a connectivity edge effect fix file for matrix connectivity, first adjust the run settings to include matrix connectivity (see section 3.3.3.4). In addition, the following line is needed so that Zonation will also read in the edge effect fix file:

```
connectivity edge effect fix file = fixfile_name.asc
```

In addition, a cost layer (`use cost = 1` and `cost file = habitatproportion_name.asc`) may be useful to correct for the amount of habitat within each cell. In that case the cost layer should include the proportion of habitat in each cell.

3.3.3.6 Cost layer

The cost layer is a standard GIS raster file (for example, .ascii, .img, .tif, etc.) about land cost**.** This file includes all basic raster information as explained in biodiversity feature distribution map files and a map of land costs for each cell. The land cost value in the matrix can be any positive number larger than 0. Thus, zero or any negative value can not be used as land cost! If areas with no land costs need to be included in the analyses, the land cost value for these cells can be set very low (e.g., 0.00001). Note that here the costs do not need to be measured in terms of money; any other measure of economical loss can be used as well. For example, in economical fisheries the fishing intensity of a landscape can be used as a cost layer—the higher the fishing intensity, the higher is the cost of protecting the particular site. The cost layer is an optional file, and Zonation would most commonly be run without cost data. If land costs are not included in the analysis, all cells implicitly have an equal cost value of 1.



```
ncols     649
nrows     555
xllcorner           294205
yllcorner           6283604.5
cellsize            0.2
NODATA_value        -1
-1        -1        -1        -1        -1        -1        -1        -1        -1        -1        -1        -1
-1        -1        -1        0.2846    0.2839    0.2831    0.2824    0.2816    0.2809    -1        -1        -1
-1        -1        -1        0.2854    -1        0.2839    0.2831    0.2824    0.2816    0.2809    0.2801    -1
```

Example of a cost layer file.

It is important that the cost layer raster has the same grid size as the feature distribution map rasters! **This means that in all files the number of columns and rows as well as the size of cells should be equal**. It is equally important that all cells which contain biodiversity feature data also need to have a cost value. In other words, all cells that have data for any feature need a cost value > 0, otherwise undefined program behavior may occur. Remember also to use **decimal points, not commas**, in all input files.

Run settings for integrating land cost in the analysis

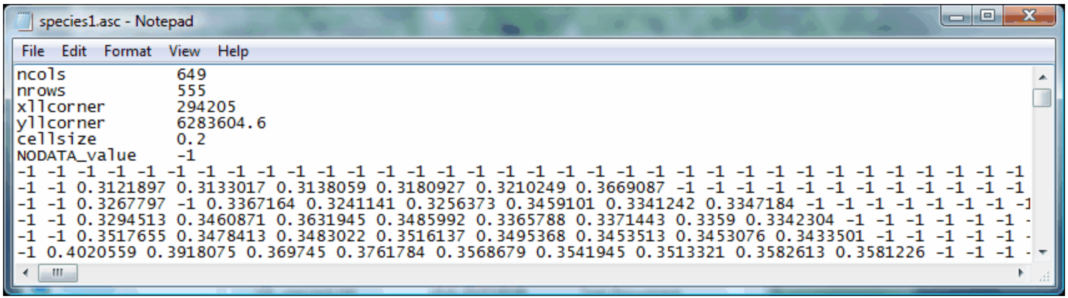
When using the cost layer, remember to type "`use cost = 1`" (cost option selected) and "`cost file = yourcostfile.asc`" (name of the cost layer file) in the run settings file.

3.3.3.7 Distributional uncertainty map layer

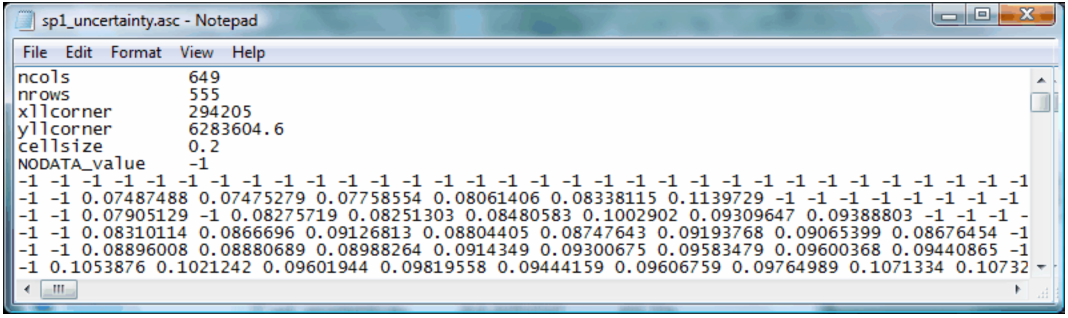
To account for uncertainty in biodiversity feature distribution data in conservation planning with Zonation, two types of files are needed: **a set of uncertainty map layers** (one for each feature) and **an uncertainty weights file**.

An **uncertainty map layer** is a standard GIS raster file (for example, .ascii, .img, .tif, etc.) of uncertainties in feature occurrence. These files are needed if the uncertainty in feature distributions (section 2.5.1) is included in an analysis. One uncertainty map layer is needed for each biodiversity feature used in the analysis. The file includes all basic raster information as explained in the biodiversity feature map files (section 3.3.2.1) and a matrix of feature occurrence uncertainties in each cell (parameter **wsc** in the info-gap uncertainty model). The feature- and cell-specific uncertainty value (wsc) can be any measure of error in prediction, any uncertainty about whether the feature will persist there, or a combination of the two, as long as the data for all features is in the same format. Zonation can treat the values as uniform errors or proportional errors. Uniform error is the default setting and works for most of the data sets, but in some cases it is more appropriate to use proportional errors (see e.g., Ben-Haim 2001).

For example, if confidence intervals are available for the probabilities of occurrence of feature A in a given cell, the uncertainty value for feature A in that same cell can be the size of the (lower half of the) respective confidence interval. The probabilities of anthropogenic threat (e.g., the uncertainty of occurrence of a feature in a given cell due to human activities in the near future) can also be used as uncertainty value. A combination of both types of uncertainty could also be utilized. The higher the uncertainty value, the greater the risk that the feature does not actually occur there (although the feature distribution data might suggest so). Thus, an uncertainty value of 0 indicates that the observed occurrence of a feature A in a given cell is trusted to be completely accurate.



```
ncols          649
nrows          555
xllcorner      294205
yllcorner      6283604.6
cellsize       0.2
NODATA_value   -1
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.3121897 0.3133017 0.3138059 0.3180927 0.3210249 0.3669087 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.3267797 -1 0.3367164 0.3241141 0.3256373 0.3459101 0.3341242 0.3347184 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.3294513 0.3460871 0.3631945 0.3485992 0.3365788 0.3371443 0.3359 0.3342304 -1 -1 -1 -1 -1 -1 -
-1 -1 0.3517655 0.3478413 0.3483022 0.3516137 0.3495368 0.3453513 0.3453076 0.3433501 -1 -1 -1 -1 -1 -
-1 0.4020559 0.3918075 0.369745 0.3761784 0.3568679 0.3541945 0.3513321 0.3582613 0.3581226 -1 -1 -1 -
```



```
ncols          649
nrows          555
xllcorner      294205
yllcorner      6283604.6
cellsize       0.2
NODATA_value   -1
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.07487488 0.07475279 0.07758554 0.08061406 0.08338115 0.1139729 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.07905129 -1 0.08275719 0.08251303 0.08480583 0.1002902 0.09309647 0.09388803 -1 -1 -1 -
-1 -1 0.08310114 0.0866696 0.09126813 0.08804405 0.08747643 0.09193768 0.09065399 0.08676454 -1
-1 -1 0.08896008 0.08880689 0.08988264 0.0914349 0.09300675 0.09583479 0.09600368 0.09440865 -1
-1 0.1053876 0.1021242 0.09601944 0.09819558 0.09444159 0.09606759 0.09764989 0.1071334 0.10732
```

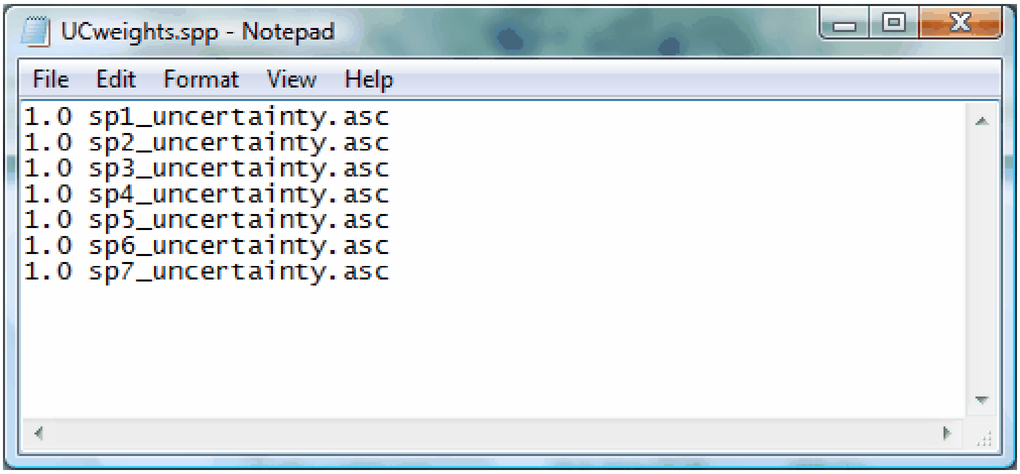
Example showing both the feature distribution layer (top, species1.asc) and the uncertainty layer (bottom, sp1_uncertainty.asc) for a biodiversity feature. During the distribution discounting process the value in each cell of the distribution layer will be discounted by a multiple of the corresponding value in the uncertainty layer.

It is compulsory that the uncertainty layer raster has the same grid size as the biodiversity feature distribution map rasters! **This means that in all files the number of columns and rows as well as the size of cells should be equal**. Also, for each feature, all cells with occurrence data should have a respective uncertainty value. Remember also to use **decimal points, not commas**, in all input files.

The uncertainty weights file contains a list of names for all distributional uncertainty map layers that correspond to feature layers. This file always contains two columns:

1. **Feature-specific weights in the uncertainty analysis**. These stress the accuracy of occurrence of a certain feature (e.g., very rare species). The higher the weight, the more strongly the program prefers cells with low uncertainty. Feature-specific weights can have any positive value larger than 0. If no feature-specific weighting of uncertainty is used (as is most commonly the case), this should be set as 1.0, indicating an equal weight for all features.

2. Name of the **distributional uncertainty map layer for the feature**. If the uncertainty maps are in a different directory than the weights file, remember also to type the correct path in front of the names. Note that the uncertainty layers for features have to be in the same order as the biodiversity feature maps in the biodiversity feature list file as these files are linked to each other solely via the order of listing in the two files.



```
1.0 sp1_uncertainty.asc
1.0 sp2_uncertainty.asc
1.0 sp3_uncertainty.asc
1.0 sp4_uncertainty.asc
1.0 sp5_uncertainty.asc
1.0 sp6_uncertainty.asc
1.0 sp7_uncertainty.asc
```

Example of the uncertainty analysis weights file.

Note that there shouldn't be any empty rows at the end of the uncertainty analysis weights file. This is because the program might interpret these as empty values or files that just don't have any names. Remember also to use **decimal points, not commas**, in all input files.

Run settings for including uncertainty of distributions in an analysis

To set up an analysis to account for uncertainty in biodiversity feature distributions, the IG-alpha parameter must be unequal to zero in the fourth column from the end of your Zonation call (see section 3.2.1). When **uncertainty** about distribution is considered **a negative thing** (see distribution discounting, section 2.5.1), **the value of** α **should be positive**. If uncertainty is considered **a positive thing** (see opportunity analysis, section 2.5.2), **the value of** α **should be negative**.

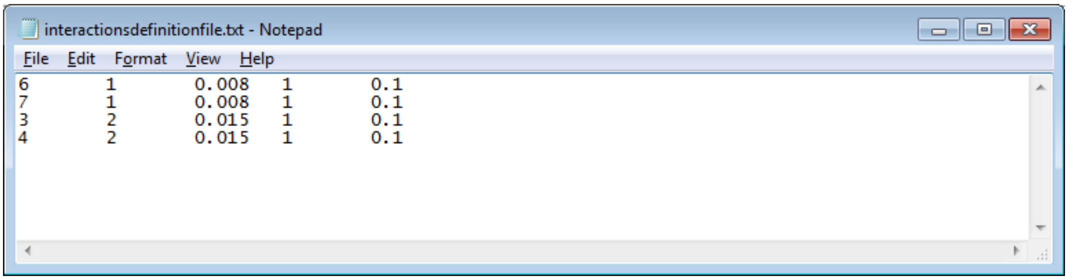
Additionally, the following rows must be included in the run settings input file:
`[Info-gap settings]`
`Info-gap proportional = 0 OR 1`
This value depends on whether uncertainty values are uniform errors (value = 0) or proportional errors (value = 1). If this row is not included, the value will be set to 0.
`use info-gap weights = 1` (uncertainty analysis selected)
`Info-gap weights file = yourweightsfile.spp` (the name of your uncertainty weights file)

3.3.3.8 Interactions definition file

A biodiversity feature interactions definition file is a text file which defines what interactions between features are included in the analysis. This option implements methods described by Rayfield *et al*. (2009). This file is only needed when ecological interactions are included in the analysis.

An interactions definition file is needed to include biodiversity feature interactions in an analysis. This file tells the program which features are interacting and to what extent. When preparing the biodiversity feature list file, please keep in mind the following points:

- **Every interaction changes the loaded, original distribution layer**. For example, assume an analysis which considers the original distributions of two features (A and B) as well as the connection of feature A's distribution to feature B's distribution. In this analysis, feature A must be listed twice in the biodiversity feature list file: the first layer will be used as it is (original distribution), and the second one will be transformed based on its connectivity to feature B. If feature A were listed only once, the landscape ranking would be done based on feature B distribution and feature A connectivity to feature B only. Thus the original distribution of feature A would not be included.
- **Make sure that an already transformed layer is not being used to transform other layers**. In this case, the interpretation of results is outside the scope of this user manual.
- **Note that every file listed in the biodiversity feature list file will be used for landscape ranking**. If a layer is used to transform another layer but should not be included in the analysis, set the weight of that layer to zero.



```
6        1        0.008    1        0.1
7        1        0.008    1        0.1
3        2        0.015    1        0.1
4        2        0.015    1        0.1
```

Example of an interactions definition file.

An example of the interaction definition file is shown above. The first column indicates the focal feature map (S1, the map that is transformed), and the second column is the feature map that the feature in column 1 is connected to (S2). Column 3 is the beta value. Beta is equivalent to the alpha value that is given in the biodiversity feature list file, and is calculated using the same formula:

$$\beta = \frac{2 * [\text{Cell size in km}]}{[\text{Distance dependence (km)}] * [\text{Input cell size}]},$$

where the feature-specific measure for landscape use refers to the distance that focal feature S1 in cell *i* can interact with feature S2 in the surrounding area of cell *i*. The fourth column indicates the type of interaction. The interaction can be positive (e.g., resource-consumer) or negative (e.g., competition). For positive interactions (type 1) use a 1 (as in the screenshot above); for negative interactions (type 2) use a 2. See section 2.6 for more information about interaction types. The fifth column gives the value of gamma, which in turn defines how the value of connectivity between S1 and S2 changes when moving away from the focal site. For a more detailed explanation about gamma, see section 2.6. By default, gamma should be 1.0. In the example file above, the first row indicates that feature map number 6 is transformed by connectivity to map number 1 using distance dependence (β_k) 0.008.

Run settings to include ecological interactions in your analysis

To include ecological interactions in an analysis, type "`use interactions = 1`" (option selected) and "`interaction file = myfile.txt`" (name of the interaction definition file) in the run settings file.

3.3.3.9 Removal mask layer

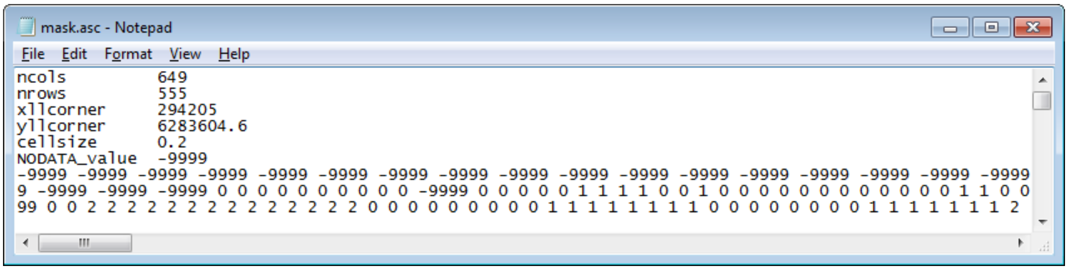
Zonation v.3.1 and higher includes an improved mask file function. **Please note that a removal mask layer composed in this new format is no longer compatible with Zonation v.2.0.** Also note that there are two different mask types for different purposes. Do not confuse them. The removal mask layer described here is sometimes called "hierarchical removal mask", in the sense that it defines a strict hierarchy or sequence of removal in the landscape.

The removal mask layer is a GIS raster file which determines the removal hierarchy of the edge cells. The typical uses of the mask layer include replacement cost analysis (section 2.7) and conservation prioritization when some predetermined information about land zoning exists. This feature is different and improved from its precedent in Zonation v.2.0.

The new mask file in Zonation v.3.1. works as follows:

- It is a raster file (.asc, .img, .tif, etc.) where integer values are assigned to cells
- The cell-specific integer is the "mask level" for the cell
- Cells with lower mask levels (small numbers) are removed before cells with higher mask levels (large numbers), thus making a forced multi-level hierarchy possible in the analysis (see Lehtomäki *et al*. 2009)
- There can be an arbitrary number of levels (e.g., 0, 1, 2, 5, 7, 100, 1500, 100000, etc.) and the levels need not be consecutive

Cells with low mask levels are removed first as they may be undesirable for conservation (e.g., built-up areas, private areas, areas ear-marked for residential building, commercial fishing, etc.) or they may have any other reason to be primarily excluded from the final solution. The cells with high mask levels that are removed last may have a special conservation value or they may already be ear-marked for conservation. These cells will be removed only after there are no more cells with lower mask level values left, and thus will be included in the top fraction of the solution.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value  -9999
-9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999
9 -9999 -9999 -9999 0 0 0 0 0 0 0 0 0 0 -9999 0 0 0 0 0 1 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0
99 0 0 2 2 2 2 2 2 2 2 2 2 2 2 2 2 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 2
```

Example removal mask layer file.

It is compulsory that the mask layer raster has the same grid size as the biodiversity feature distribution map rasters! This means that in all files the number of columns and rows as well as the size of cells should be equal. It is equally important that all those cells which have data of any of the feature used in the analysis (that is to say, the cells that are NOT marked as "no data" in all biodiversity feature distribution files), also have a mask value.

Run settings for using a removal mask layer

When using the removal mask layer, include the following lines under the `[Settings]` heading in the run settings file:
`use mask = 1` (mask option selected)
`mask file = yourmaskfile.asc` (name of the removal mask layer file).

Note that use of the removal mask layer is likely to result in a biologically non-optimal solution as the program is not allowed to remove cells only based on their conservation value. See replacement cost analysis (sections 2.7 and 5.2.4) for the analysis of the suboptimality of masked solutions with areas forcibly masked to low or high ranks of the solution.

3.3.3.10 Analysis area mask

This raster file (for example, .ascii, .img, .tif) indicates areas to be used in the analysis. If the analysis area mask is used, Zonation will only take into account those cells that are indicated by this file. Please note that this mask is not suitable for forcing areas into the top or bottom fraction of the solution; for that analysis, use the removal mask layer (section 3.3.3.9).

The analysis area mask file will be applied to all input raster files. It can be used for:
1. Forcing alignment of data
2. Cutting out areas that are not needed in analysis (e.g., areas outside the country), thereby saving memory and allowing more features to be analyzed
3. Targeting analysis to a subsection of the landscape without needing to develop a full new set of files—just the area mask file will suffice

This file includes all basic raster information as explained in the biodiversity feature distribution map files followed by a matrix where cells are categorized as follows: Cells with value >0 are included in the analysis, whereas cells with zeroes or "missing data" values (e.g., -1) are excluded. A simple operational area mask file will have "1"s for cells that should be analyzed and "0"s for cells that are outside the area of interest.

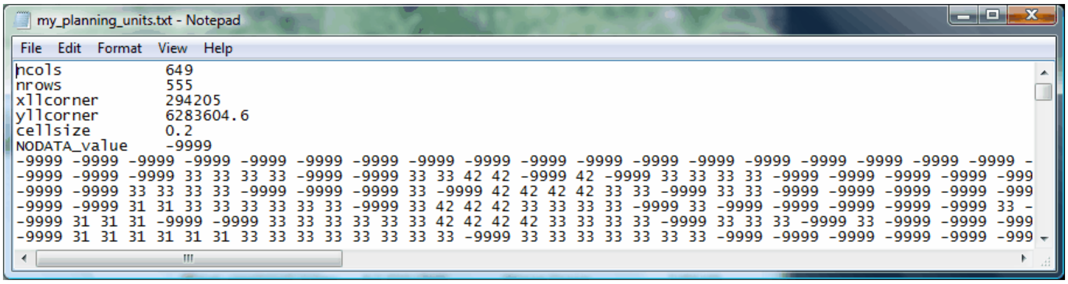
It is compulsory that the analysis area mask raster has the same grid size as the species distribution map rasters! **This means that in all files the number of columns and rows as well as the size of cells should be equal**.

Run settings for using analysis area mask

To use the analysis area mask, type "`mask missing areas = 1`" (analysis area mask is used) and "`area mask file = your_areamask.asc`" (name of the area mask raster) in the run settings file.

3.3.3.11 Planning unit layer

The planning unit layer file is a standard GIS raster file containing integer numbers. This file includes all the basic raster information as explained in the biodiversity feature map files (see section 3.3.2.1) and a matrix where the number given for a cell identifies the planning unit that the cell belongs to. Planning unit numbers must be positive integers, but they need not be sequential (e.g., they could be 1, 2, 5, 12, 101, etc.). Planning units may be used to model the situation when, for example, land ownership dictates that certain groups of grid cells should be treated as distinct units. Or, with directed connectivity, the planning units could correspond to hydrologically linked catchment areas. When the planning unit layer is in use, the entire planning unit is removed simultaneously. The cell removal rules operate as before, but they operate on value aggregated across the planning unit. Also, the cost of the planning unit is taken as either the summed cost of cells (if the cost layer is used) or as the area of (number of cells in) the planning unit if costs are not used. Note that each planning unit does not need to be spatially continuous—a planning unit may consist of a scattered collection of cells.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value  -9999
-9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -
-9999 -9999 -9999 33 33 33 33 -9999 -9999 33 33 42 42 -9999 42 -9999 33 33 33 33 -9999 -9999 -9999 -9999 -999
-9999 -9999 33 33 33 33 -9999 -9999 -9999 33 -9999 42 42 42 42 33 33 -9999 33 33 -9999 -9999 -9999 -9999 -999
-9999 -9999 31 31 33 33 33 33 33 33 -9999 33 42 42 42 33 33 33 33 -9999 33 -9999 -9999 -9999 -9999 -9999 33 -
-9999 31 31 31 -9999 -9999 33 33 33 33 33 33 42 42 42 42 33 33 33 33 -9999 33 33 33 -9999 33 -9999 -9999 -999
-9999 31 31 31 31 31 31 33 33 33 33 33 33 33 33 -9999 33 33 33 33 33 33 33 -9999 -9999 -9999 -9999 -9999 -999
```

Example of a planning unit layer file.

A planning unit number should be defined for every cell that has feature data. If not, some kind of error condition is likely to occur. It is not harmful to have planning units extending outside the area with feature data—the critical part is that all locations with feature data are covered with planning units.

Use of large planning units will automatically cause a decrease in the quality of results. The reason for this is that large planning units will probably contain areas that are both good and bad for conservation. Consequently, the performance curves will suggest lower protection levels than what can be obtained if selection is based on individual grid cells. With respect to the computation time needed by Zonation, the use of planning units is likely to cut computation times. The reduction will be larger the more cells are grouped into planning units.

Note

Make sure that every cell inside the analysis area has a planning unit number assigned. Otherwise the cells will not be ranked (removed) when using planning units analysis in Zonation. If some cells do not have any planning unit number, Zonation will still run and generate results, but those cells will not be included in the ranking. In the Zonation graphical user interface this is noted with an error message indicating how many cells have been removed (out of the total number of cells included in the analysis area).

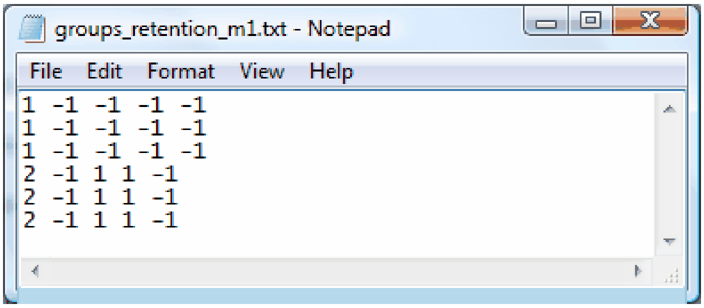
Run settings for planning units

When using planning units, remember to type `"use planning unit layer = 1"` (planning unit option selected) and `"planning unit layer file = my_plu_layer.asc"` (name of the planning unit layer file) in the run settings file. Note that Zonation automatically sets the warp factor to 1 when using planning units.

3.3.3.12 Groups file

The groups file allows linking input features to groups, upon which various operations are done. The groups file has as many rows as there are input features. Columns of the file define specific groupings (one per column). All numbers in this file are integers. A value of -1 for a feature in any group indicates that the feature is not grouped with respect to that particular criterion. The order of the features in the groups file must match that in the biodiversity feature list file. For each feature, there are five columns, one column for each group type:

column 1: output group
column 2: condition group
column 3: retention groups
column 4: retention mode
column 5: arbitrary kernel group (for distribution smoothing)
column 6: arbitrary kernel group (for interactions)
column 7: arbitrary kernel group (for matrix connectivity)



```
1 -1 -1 -1 -1
1 -1 -1 -1 -1
1 -1 -1 -1 -1
2 -1 1 1 -1
2 -1 1 1 -1
2 -1 1 1 -1
```

Example of a groups file.

(1) The **output group** is useful when mixed sets of biodiversity features are used and representation for each set should be assessed separately. Zonation will output mean, minimum, and maximum representation curves for each of the groups in a separate groups output file (section 3.4.3). Output groups can be used to separate higher taxa (birds, mammals, plants, etc.), community and species features, negatively and positively weighted features, habitat quality and connectivity layers, or creative combinations of any of the aforementioned. The information could, of course, be extracted manually from the standard representation curves file (section 3.4.1). Using output groups can save some manual work and make interpretation of the results more straightforward.

(2) The **condition group** defines linkage of features to condition groups.

The number is this column is linked to the condition layers list file (section 3.3.3.14), which specifies which condition number links to which condition layer file.

The condition column is different from the output group column in that only the one column of information is needed to define the output group. The condition feature additionally needs the condition layers list file and the condition layers themselves.

(3) The **retention group** defines linkage of features to retention groups.

The retention group is analogous to the condition group as the linkages to condition and retention features are structurally identical.

The number in this column is linked to the retention layers list file (section 3.3.3.15), which specifies which retention number links to which retention layer file.

(4) The **retention mode** can be 1 = stop loss or 2 = management intervention (see section 2.10 for descriptions).

(5-7) These columns define the arbitrary kernel group for different analysis features (distribution smoothing: column 5; interactions: column 6; matrix connectivity: column 7) . This feature is available since Zonation v.4. If not used, the default kernel (Gaussian) will be applied in all the transformations for all the features. To use a different kernel you need to specify a positive number here. The supported values are:

- 1: exponential kernel
- 2: gaussian
- 3: threshold (square block) or uniform distribution
- 4: triangular kernel (linear decay)
- 100 and higher: arbitrary kernels, loaded from files named 'zonation_arbitrary_kernel_#number.txt', for example 'zonation_arbitrary_kernel_101.txt'.

See section 3.3.3.18 for details on how to specify the shape of arbitrary kernels.

Example

If, for example, the group file contains the row:

```
3  1  -1  1  -1
```

This means that the feature belongs to output group number 3 and condition group 1. There is no retention group specified (-1 in column 3), meaning that the retention mode of column 4 is a dummy value. Columns 5 or higher are not used either.

Going back to the condition group, this means that in the condition file (section 3.3.3.14) there must be a row

`1  cond_group_1_grid_file_name.asc`.

meaning that condition group 1 is linked to the raster grid file cond_group_1_grid_file_name.asc.

Examples of using output groups are provided in Exercises 10 and 11.

Run settings for using a groups file

To include a groups file in an analysis, type "`use groups = 1`" (option selected) and "`groups file = mygroupsfile.txt`" (name of the groups file) in the run settings file.

3.3.3.13 Alternative land uses layer

In Zonation v.3.1 and higher, it is possible to consider multiple opportunity costs of conservation by including layers that describe landscape suitability for competing land uses. The alternative land use layers are compiled and treated precisely as the normal biodiversity feature layers (section 3.3.2.1) and listed in the biodiversity feature list file (section 3.3.2.2). The only difference is that the alternative land use layers are given negative weights instead of positive ones in the first column of the list file. If target-based planning is used as the cell removal rule, the value in the fifth column of the biodiversity feature list file denotes the fraction of the competing land use that should be excluded from the solution. Several alternative land uses can be entered in a single analysis.

It is compulsory that all feature grids, including negatively weighted ones, have the same number of columns and rows.

3.3.3.14 Condition layer

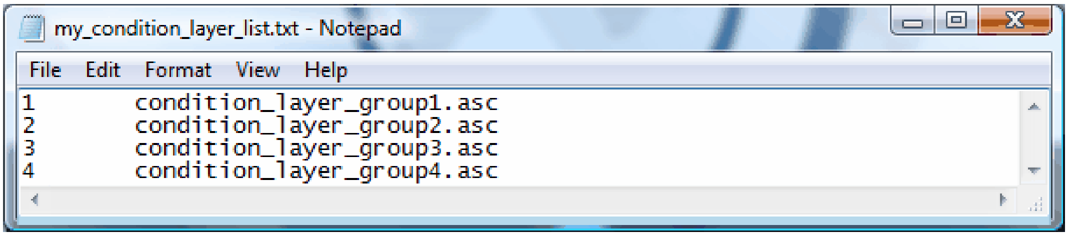
Condition layers are needed when landscape condition is included in an analysis to account for past habitat loss or degradation. They are raster grid layers (.ascii, .img, .tif, etc.) that describe the fraction of suitable habitat or occurrences that remain (relative to some historical baseline) for a group of biodiversity features in each grid cell. The condition values can vary between 0.0 and 1.0. A value of 1.0 indicates pristine condition where the habitat suitability or feature occurrence has not degraded. A value of 0.0 indicates a completely degraded condition. Any negative values (missing data) will be treated as zeroes.

It is compulsory that the condition layers have the same grid size as the biodiversity feature map rasters!

When a condition grid layer is included in a Zonation analysis, the occurrence level of a cell for a biodiversity feature in the distribution map layer will be multiplied by the condition value of the cell for the condition group in which the feature belongs to. The condition grid layers need to be accompanied by two files: a groups file and a condition layer list file.

A **groups file** (section 3.3.3.12) assigns the biodiversity features to condition groups based on their habitat preferences. Condition groups are assigned in column 2 of this file.

A **condition layer list file** links condition groups to the condition grid layers. This file has two columns:
Column 1 gives the number of the condition group. These numbers refer to those assigned to biodiversity features in the groups file.
Column 2 has the name of the condition grid layer for that group of features.



```
1       condition_layer_group1.asc
2       condition_layer_group2.asc
3       condition_layer_group3.asc
4       condition_layer_group4.asc
```

Example of a condition layer list file.

Run settings for landscape condition analysis

To include landscape condition in an analysis, include the following rows in the run settings file under the `[settings]` section:

`use groups = 1`
`groups file = mygroupsfile.txt` (name of the groups file)
`use condition layer = 1`
`condition file = condition_layers_list.txt`

3.3.3.15 Retention layer

A retention layer is a raster grid (.ascii, .img, .tif, etc.) that describes the fraction of local occurrences or habitat suitability that would be retained for a group of species (or other biodiversity features) in the absence of conservation action. Values in the cells can vary. A value of 0 indicates that in the absence of conservation, the cell will lose all of its biodiversity value. A value of 1 indicates that the cell will be retained as it is and no loss would occur. A value greater than 1.0 indicates that condition in the cell will improve with management. Values >1.0 are only appropriate for the 'management gain' mode of analysis, mode 2.

It is compulsory that retention layers have the same dimensions as the biodiversity feature map rasters!

When a retention layer is included in a Zonation analysis, it transforms the occurrence of the feature to expected loss by multiplying the value in the cell of the distribution grid by (1-retention level) of the same cell in the retention layer for that feature.

The retention analysis needs three files to be operational:

A **groups file** (section 3.3.3.12) assigns the biodiversity features to retention groups. The groups file also defines the mode of retention, which can be either that (1) reduced loss of habitat occurs or that (2) the quantity or quality of habitat is actually improved through management intervention. Retention groups are assigned in column 3 and retention modes in column 4 of this file.

A **retention layer list file** links retention groups to the retention grid layers. This file has two columns:
Column 1 gives the number of the retention group. These numbers refer to those assigned to biodiversity features in the groups file.
Column 2 has the name of the retention grid layer for that group of features.

The **retention layer** describes retention in terms of the fraction of cell condition retained even in the absence of conservation. A value of 1.0 indicates no change in condition even in the absence of conservation. A value of 0.0 (mode 1) would indicate total loss in the absence of conservation. A value of, say, 1.2 (mode 2), would imply a 20% management gain assuming the site is targeted for conservation.

Often, the most meaningful way to include landscape retention in an analysis is to use it in combination with species representation in the protected areas only. To do this, duplicate the biodiversity features in the biodiversity feature list file; see section 5.3.5 for the full setup. The first copy of the layers will be used to model representation and the second one will model loss in occurrence if the cell is not selected for the protected area network. The relative weights for representation and retention are defined by parameter β which is given in the run settings file (see below). The higher the value, the more emphasis is given to retention in the analysis.

When landscape condition and retention are both applied in the same analysis, the occurrence values of features in cells are first transformed by condition and then by retention levels.

Run settings for landscape retention analysis

To include landscape retention in an analysis, adjust the run settings file to include the following lines:

```
use groups = 1
groups file = mygroupsfile.txt (name of the groups file)
use retention layer = 1
```

`retention file = my_retention_layer_list.txt`

`retention layers relative weight =` value for β, a decimal multiplier for the retention layer weights for balancing between representation and retention

3.3.3.16 Administrative units analysis files

To run an analysis that considers conservation priorities over multiple administrative regions, three additional files are needed:

1. An administrative units description file that defines local and global factors for multiplying weights biodiversity features
2. An administrative units map raster that assigns cells of the whole planning region to administrative subregions
3. Administrative units weight matrix defining weights for each biodiversity feature in each administrative region

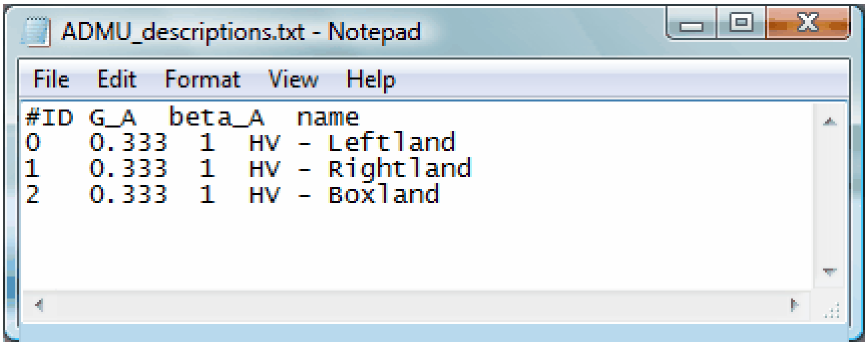
1\. An **administrative units description file** (shown below) is a list of administrative regions and their weights. The first row is a header row for column names. The file has four columns:

Column 1 has the identification number of the administrative region (these should be positive integers). The values are linked to the administrative units map raster and should match those.

Column 2 has the region-specific global weight, G_A. This number is the global priority given for a subregion. Effectively, conservation value aggregated from a region becomes multiplied by this number. A high number elevates the priority given to the region relative to other regions. Note that in the strong variant (ADMU mode 2) large global weights for the areas increase the relative importance of the local component because conservation values of the regions are summed (loss of conservation value is computed from 1+N components, where 1 is for global analysis and N is for N subregions). The weights in this column can be standardized to sum to one to keep the balance between global and local considerations directly dependent on parameter p.

Column 3 has the region-specific local weight balance between local and global feature weights. This number is bounded between 0 and 1. It gives the balance of local vs. global considerations (feature weights) in local decision making. A value of 1 here means that the region is concerned only with its own local priorities (w^L_{jA}). A zero here means that the region is willing to go with global priorities (w^G_j), thus foregoing local opinions of what is important.

Column 4 gives the name or ID of the region for output purposes.



```
#ID G_A   beta_A  name
0   0.333  1   HV - Leftland
1   0.333  1   HV - Rightland
2   0.333  1   HV - Boxland
```

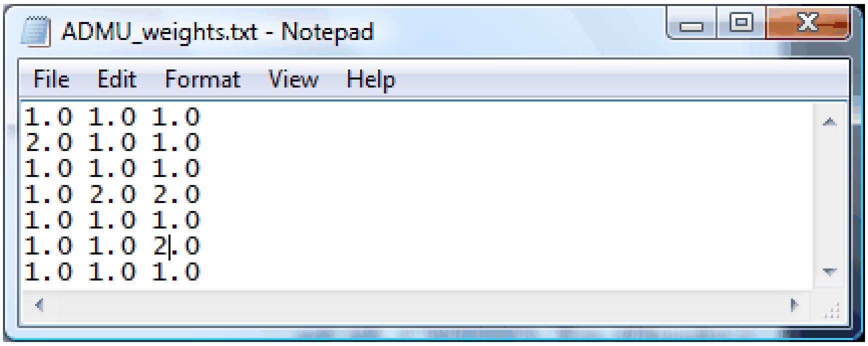
Example of an administrative units description file.

2. An **administrative units map raster** is a raster grid that contains integer numbers. It defines the division of grid cells to the administrative regions. Each administrative region has its own identification number, defined in the administrative units description file. Cells in this map raster should contain the identification numbers of the administrative regions they belong to.

It is possible to use the administrative units in combination with planning units. If you do this, please make sure that the planning units do not go across the borders of administrative regions.

It is compulsory that the ADMU raster has the same dimensions as the feature grids. The ADMU number should be defined for all locations for which feature data is entered. The ADMU number can be defined for locations without any feature data (i.e., one can use an ADMU layer based on simple polygons even if the feature data only covers parts of them).

3. An **administrative units weights matrix file** defines the a priori local weights of biodiversity features used in the analysis. Here, biodiversity features are in rows and administrative regions in columns. The local weights could reflect the regional conservation priorities or policies. The order of biodiversity features in the matrix file should match that of the biodiversity feature list file. This file has no header row.



```
1.0 1.0 1.0
2.0 1.0 1.0
1.0 1.0 1.0
1.0 2.0 2.0
1.0 1.0 1.0
1.0 1.0 2.0
1.0 1.0 1.0
```

Example of an administrative units weights matrix.

Run settings for conservation prioritization over multiple administrative regions

For an administrative units analysis, add the following lines to the run settings file:

```
[Administrative units]
use ADMUs = 1
ADMU descriptions file = my_ADMU_descriptions.txt
ADMU layer file = ADMUs_distribution_map.asc
ADMU weight matrix = ADMU_weights_matrix.txt
```

`calculate local weights from condition = 1` (Presently a non-operational dummy parameter.)

`ADMU mode = 1` OR `2`, depending on whether local representation for all features (mode 2) are enforced or not (mode 1; see section 2.12)

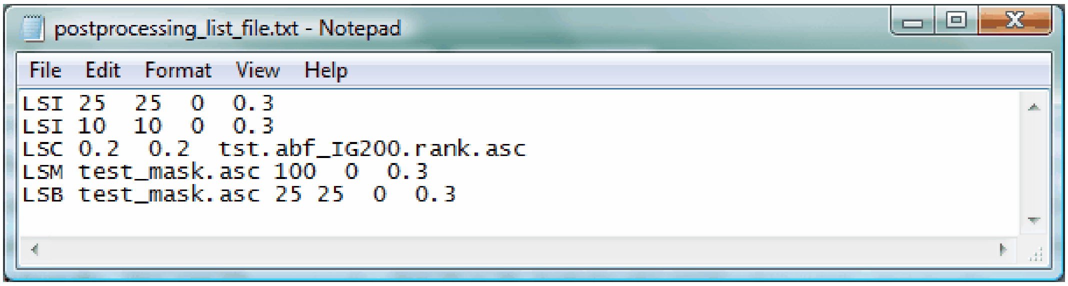
`Mode 2 global weight = 0.5` This parameter specifies the balance of global and local conservation value when using ADMU mode = 2. This parameter is bounded between zero (only local considerations) and one (only global considerations influence value).

`row count for per ADMU output curves = 0` Number of rows in the per ADMU .curves.txt files.

See section 5.3.8 for more detailed instructions for the analysis setup.

3.3.3.17 Automated post-processing file

The automated post-processing file describes post-processing analyses that are called automatically after the main computations have finished. The file contains a list of analyses to be done along with their parameter settings. Below is an example of what post-processing calls look like (there is one row per call in a text file; there can be multiple calls of each type and the calls need not be in any particular order):



```
LSI 25  25  0  0.3
LSI 10  10  0  0.3
LSC 0.2  0.2  tst.abf_IG200.rank.asc
LSM test_mask.asc 100  0  0.3
LSB test_mask.asc 25 25  0  0.3
```

Example of a post-processing file.

Run settings for automated post-processing

To enable automated post-processing analyses in a batch operation, include the following line under the `[Settings]` section in the run settings file:

"`post-processing list file = ppa_list_file_name.txt`" (name of the post-processing file).

Four different post-processing analyses can be computed:

1) **Landscape identification analysis, identified by analysis type LSI**
This option allows identification of separate management landscapes based on the distance and similarity in biodiversity feature composition between two sites. Spatially distinct areas (consisting of multiple grid cells) in a Zonation solution can be classified in management landscapes. An area is joined to a landscape if it is spatially close enough and similar enough with respect to biodiversity feature composition to *any* other distinct area in the same landscape. Landscape identification is done for a specific top fraction of a Zonation ranking.

To perform a landscape identification analysis, four parameters need to be defined:

1. First define the **percentage of landscape** that will be included in the classification, including areas from the top fraction of the landscape. Note: this number is given as percents (e.g., a value of 20 includes the best 20% area from the landscape to the solution).

2. Define **inclusion minimum**, which determines how highly ranked cells must be included in each management landscape (e.g., value of 10 means that each management landscape has to contain at least one cell which belongs to the top 10% fraction of the whole landscape). Note that if the inclusion minimum is equal to (or larger than) the **percentage of landscape**, all spatially separate areas will be joined into the management landscapes, whereas if the inclusion minimum is smaller, then management landscapes only with sufficiently high ranked core areas are included.

3. Give the nearest neighbor **maximum distance** (in cells!) that is allowed between spatially discrete patches that are included in the same management landscape (e.g., a maximum distance of 0 would mean that all separate groups of cells are identified as unique management landscapes).

4. Give the **maximum difference in feature composition.** This determines how much the feature compositions between two cells are allowed to differ (in terms of relative densities) for them to be joined to the same landscape. A value of 0 indicates that the feature composition in two patches is identical. A value of 1 indicates that the difference in relative density between two patches is on average log10 across features. For example, a maximum difference of 0.2 means that on average two features out of ten have a 1-log difference in their density or that one tenth of all the features have a 2-log difference. For more details, see Moilanen *et al*. (2005).

Thus, if you want to identify management landscapes within your solution, you need the following line in your automated post-processing file:

```
LSI  percent1  percent2  distance  similarity
```

in which "LSI" tells Zonation to perform the analysis, "percent1" is the fraction of landscape to be included, "percent2" is the inclusion minimum, "distance" is the maximum distance and "similarity" the maximum difference in biodiversity feature composition.

Each landscape identification analysis will produce two output files with extensions .nwout.#.ras.asc and nwout.#.spp_data.txt, described in section 3.5.1.1. The number (indicated here with #) in the file name is generated by the ordinal number of the call to an LSI analysis: the first call produces .nwout.1.ras.asc, the second .nwout.2.ras.asc, and so on. The beginning of the LSI output files will be the general output file name you have given in the Zonation call.

2) **Landscape comparison LSC**

The solution comparison calculates how much two solutions overlap with each other and what the average difference in the cell removal order is. The comparison is always made between the present solution and an older solution by using the rank.asc files of both solutions as input files. For landscape comparison, include the following line in the automated post-processing file:

```
LSC  fraction_of_present_solution
fraction_of_comparison_solution   comparison.rank.asc   output.asc
```

Here, "LSC" tells Zonation to perform the analysis. "fraction_of_present_solution" defines what (top) fraction of the present solution is accounted for in the comparison. The value can vary be between 0 an 1. "fraction_of_comparison_solution" defines what (top) fraction of a previous solution is accounted for in the comparison. Again, the value can vary be between 0 an 1. "comparison.rank.asc" refers to the rank output file of the previously computed solution that will be compared to the new solution. "output.asc" denotes the name of the output raster (section 3.5.1.3) that shows the overlapping areas.

3) **Landscape identification for masked subregion of landscape, LSM**

This analysis is, effectively, an LSI analysis done to an externally specified subregion of the landscape.

This analysis can identify management landscapes within subregions of the full landscape. This subregion could, for example, be areas owned by a particular land owner. A mask file identifies the areas that should be included. Do not confuse this mask file with the mask used in landscape priority ranking as these would typically be two different mask files.

To identify management landscapes for masked areas, add the following line to the post-processing file:

```
LSM   mask_file  percent  distance  similarity
```

in which "LSM" calls the analysis and "mask_file" is the name of the raster showing areas of interest (prepare this file so that areas of interest have number =1 and the rest of the grid cells are missing data, e.g., -1). "Percent" gives the inclusion minimum (see LSI), and "distance" and "similarity" are as in LSI.

The output of the LSM analysis is like that of the LSI analysis.

Since Zonation v.4, a special "-1" mode of the LSM analysis allows to get statistics of planning units. In this mode, no landscape identification is really performed, but the unit numbers provided in a layer of units (mask file) are directly used as network identifiers. The syntax of the "LSM -1" analysis line in the Zonation post-processing file is:

```
LSM units_layer.asc 0 -1 0
```

-1 distance indicates that rather than using a distance parameter to identify networks, the units/networks are defined as specified in the LSM mask layer (which is used as a layer of units rather than a mask). The other two parameters (percentage and similarity) are ignored, and can be set to 0.

The units layer must use only (consecutive) integer numbers starting from 1, each of them identifying a different unit for which statisitics are requested. These units will typically be planning units, and do not need to be contiguous.

The `nwout.#.spp_data.txt` output file (see section 3.5.1.1) is generated normally. However, as an exception to the normal functioning of LSM analyses, when the "-1" special mode is used, the `nwout#.ras.asc` output raster file is not generated (as it would be the exactly the same raster as the LSM mask).

4) **Landscape identification for top fraction inside masked areas LSB**

LSB can combine a top fraction analysis and a mask file to choose areas used in identification of management landscapes. It is different from LSM in that LSM uses all areas indicated by the mask while LSB only uses a given top fraction within masked areas. The line to type for LSB is:

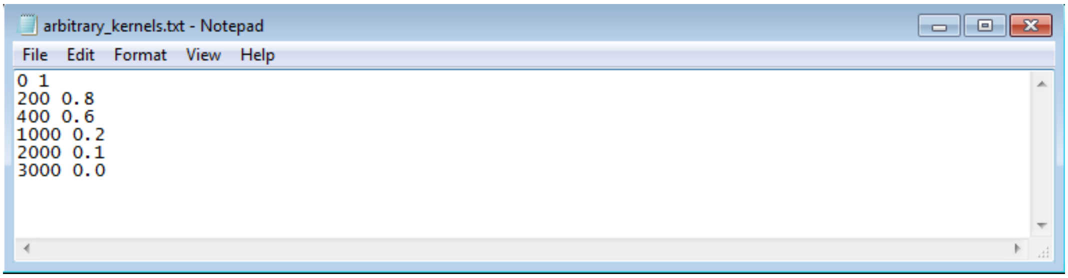
```
LSB   mask_file  fraction1 fraction2  distance  similarity
```

Here, "mask_file" is as for LSM analysis, and the rest of the parameters are the same as in LSI.

3.3.3.18 Arbitrary kernel defintion files

Arbitrary functions or shapes for connectivity kernels can be specified in this type of file. Each file of this type effectively defines a custom or user defined kernel function that can be used in different transformations, such as the distribution smoothing, interactions, or matrix connectivity methods. Arbitrary kernel files contain 2 columns of values that specify x-y coordinates (where x is the distance and y is the value of the arbitrary kernel function). The kernel functions are defined by linear interpolation between each consecutive pair of points. The sequence of x values must be strictly sorted from lower to higher. For distances higher than the last x value provided, the last segment is projected (constraining values to >=0).

These files must be plain text files with two columns of positive real values.



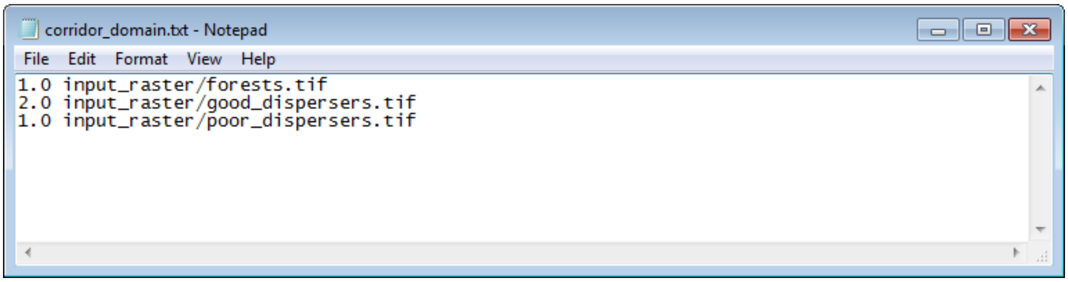
```
0 1
200 0.8
400 0.6
1000 0.2
2000 0.1
3000 0.0
```

Example of an arbitrary kernel definition file.

3.3.3.19 Corridor domain layers

When using the corridor building mechanism of Zonation, it may be desirable to exclude certain areas from the corridor building process (or to restrict corridors to certain areas). This can be done by specifying one or more so-called corridor domain layers. Two types of files are needed: the list of layers (as a plain text file), and the raster maps or layers (as GIS layers).

Corridor domain layers list file
This is a plain text file and its name is specified in the Zonation settings file using the option `domain layers list file`. Domain layers are listed in a file with two columns. The first column gives a layer-specific weight, and the second column the name of a domain layer (raster map).



```
1.0 input_raster/forests.tif
2.0 input_raster/good_dispersers.tif
1.0 input_raster/poor_dispersers.tif
```

Example of a corridor domain layer list file.

Corridor domain layers
The files given in the second column of the domains list need to be provided in the corresponding Zonation setup, as potentially multiple **corridor domain layers**, given as raster files (for example in GeoTIFF format) and with the path names specified in the corridor domain layers list file. In the example shown above, you would need the three GeoTIFF raster maps. For every layer, the values in the raster map cells are used as weights for calculating the corridor loss penalty. In their simple variant corridor domain layers can contain only two values: 1 or 0 (or no-data, missing). This defines a binary mask, whereby corridors will be built within some areas (those landscape elements or cells with value 1, whereas some other areas (those with 0s or missing/no-data) will be excluded from the corridor building mechanism. More generally, any real number (weight) can be used as a value in any of the cells of these raster maps.

3.4 Standard Zonation output

This section describes the basic output produced by the program. Zonation automatically produces two sets of outputs: visual and file. The visual output is shown in the graphical user interface, and the file output files are saved in the same directory as the program unless another file path has been specified for the output.

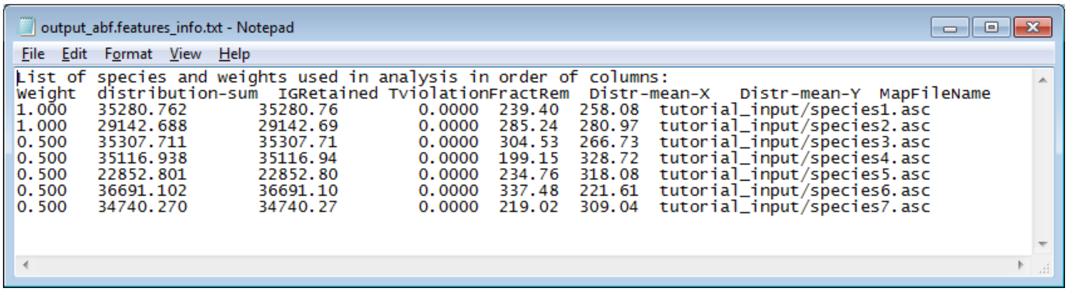
3.4.1 Automated file output

In addition to the visual output, Zonation automatically produces a number of different output files for each run. Here we describe the ones that are always produced, regardless of analysis variant. Analysis-specific files are described in the next section. The output filename (e.g., "output") specified in the command line will be used for each of the output files with a varying suffix: output.jpg, output.emf, output.curves.txt, output.prop.asc, output.rank.asc, output.wrscr.asc, and output.run_info.

.jpg file

This file is an image of the map of the area illustrating the Zonation results, ranked by using different colors to indicate the biological value of the site. Here the best areas are displayed in red, the worst areas in black, and the "no data" areas in white. See section 4.3.1 for a more detailed interpretation of the colors used in landscape ranking.

.features_info.txt file

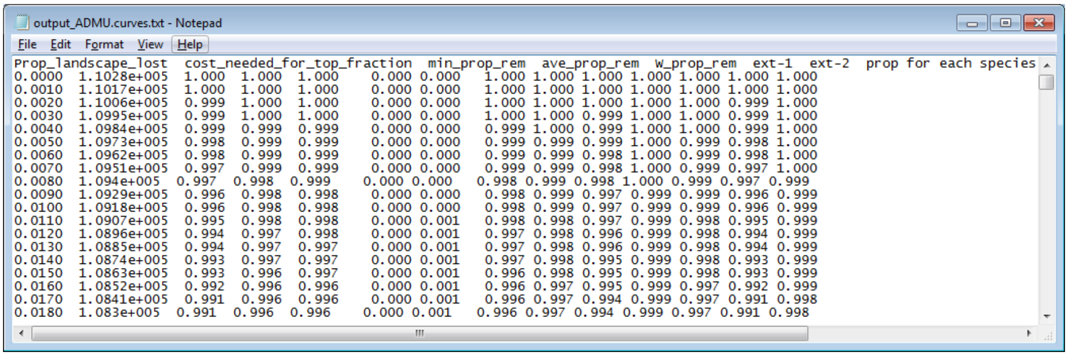


```
List of species and weights used in analysis in order of columns:
Weight   distribution-sum  IGRetained TviolationFractRem  Distr-mean-X   Distr-mean-Y  MapFileName
1.000    35280.762        35280.76        0.0000  239.40  258.08  tutorial_input/species1.asc
1.000    29142.688        29142.69        0.0000  285.24  280.97  tutorial_input/species2.asc
0.500    35307.711        35307.71        0.0000  304.53  266.73  tutorial_input/species3.asc
0.500    35116.938        35116.94        0.0000  199.15  328.72  tutorial_input/species4.asc
0.500    22852.801        22852.80        0.0000  234.76  318.08  tutorial_input/species5.asc
0.500    36691.102        36691.10        0.0000  337.48  221.61  tutorial_input/species6.asc
0.500    34740.270        34740.27        0.0000  219.02  309.04  tutorial_input/species7.asc
```

Example of a .features_info.txt file.

This is a text file containing a list of the biodiversity features and the **relative weights** `(Weight)` used in the analysis. This file also shows the initial **sum of feature distributions** `(distribution_sum)`and the level of cell removal at which point targets for particular feature have been violated. The initial sum of distribution is simply the sum of each feature's local occurrence levels throughout the landscape. For example, if the biodiversity feature data is in probabilities of occurrence, this is the sum of probabilities in all cells before any landscape has been removed. The third column shows the distribution sum after distribution discounting in an uncertainty analysis `(IGRetained)`. If target-based planning is used as the cell removal rule, each feature has a defined target (e.g., 25% of original distribution) which the program seeks to retain during the removal of landscape. However, these targets will inevitably be violated as all cells will eventually be removed. Thus, in the .features_info.txt file the program simply reports **when the targets of particular feature have been violated** (i.e., what fraction of landscape was still remaining when the proportion of each feature's original distribution dropped below the given target) (`TViolationFractRem`). **If target-based planning is not used as a cell removal rule, this column has only dummy values.** After the list you can find columns representing more detailed information about the proportion of each feature distribution remaining when landscape is iteratively removed. The last columns show **mean X and Y coordinates** of the original distribution of the feature (`Distr-mean-X and Distr-mean-Y`)and name of the **biodiversity feature map file** (`MapFileName`).

.curves.txt file

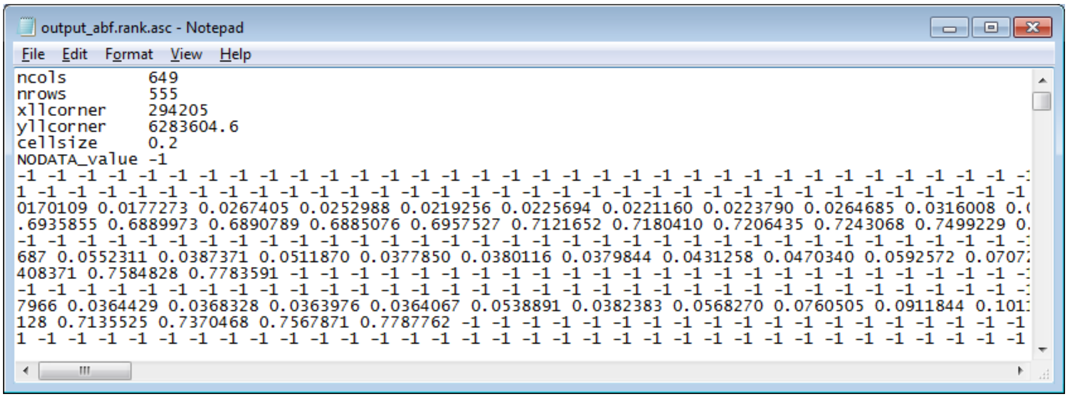


```
Prop_landscape_lost  cost_needed_for_top_fraction  min_prop_rem  ave_prop_rem  W_prop_rem  ext-1  ext-2  prop for each species
0.0000  1.1028e+005  1.000  1.000  1.000    0.000 0.000   1.000 1.000 1.000 1.000 1.000 1.000 1.000
0.0010  1.1017e+005  1.000  1.000  1.000    0.000 0.000   1.000 1.000 1.000 1.000 1.000 1.000 1.000
0.0020  1.1006e+005  0.999  1.000  1.000    0.000 0.000   1.000 1.000 1.000 1.000 1.000 0.999 1.000
0.0030  1.0995e+005  0.999  1.000  1.000    0.000 0.000   1.000 1.000 0.999 1.000 1.000 0.999 1.000
0.0040  1.0984e+005  0.999  0.999  0.999    0.000 0.000   0.999 1.000 0.999 1.000 1.000 0.999 1.000
0.0050  1.0973e+005  0.998  0.999  0.999    0.000 0.000   0.999 0.999 0.999 1.000 0.999 0.998 1.000
0.0060  1.0962e+005  0.998  0.999  0.999    0.000 0.000   0.999 0.999 0.998 1.000 0.999 0.998 1.000
0.0070  1.0951e+005  0.997  0.999  0.999    0.000 0.000   0.999 0.999 0.998 1.000 0.999 0.997 1.000
0.0080  1.094e+005  0.997  0.998  0.999    0.000 0.000   0.998 0.999 0.998 1.000 0.999 0.997 0.999
0.0090  1.0929e+005  0.996  0.998  0.998    0.000 0.000   0.998 0.999 0.997 0.999 0.999 0.996 0.999
0.0100  1.0918e+005  0.996  0.998  0.998    0.000 0.000   0.998 0.999 0.997 0.999 0.999 0.996 0.999
0.0110  1.0907e+005  0.995  0.998  0.998    0.000 0.001   0.998 0.998 0.997 0.999 0.998 0.995 0.999
0.0120  1.0896e+005  0.994  0.997  0.998    0.000 0.001   0.997 0.998 0.996 0.999 0.998 0.994 0.999
0.0130  1.0885e+005  0.994  0.997  0.997    0.000 0.001   0.997 0.998 0.996 0.999 0.998 0.994 0.999
0.0140  1.0874e+005  0.993  0.997  0.997    0.000 0.001   0.997 0.998 0.995 0.999 0.998 0.993 0.999
0.0150  1.0863e+005  0.993  0.996  0.997    0.000 0.001   0.996 0.998 0.995 0.999 0.998 0.993 0.999
0.0160  1.0852e+005  0.992  0.996  0.996    0.000 0.001   0.996 0.997 0.995 0.999 0.997 0.992 0.999
0.0170  1.0841e+005  0.991  0.996  0.996    0.000 0.001   0.996 0.997 0.994 0.999 0.997 0.991 0.998
0.0180  1.083e+005  0.991  0.996  0.996    0.000 0.001   0.996 0.997 0.994 0.999 0.997 0.991 0.998
```

Example of output.curves.txt file.

The first column (`Prop_landscape_lost`) in the .curves.txt file gives the proportion of the landscape removed. **The second column** (`cost_needed_for_top_fraction`) shows the cost of the remaining landscape. If land costs are not included in the analysis, this column represents the number of cells remaining in the landscape. **The third column** (`min_prop_rem`) shows the lowest biodiversity feature distribution proportion remaining in the landscape (i.e., the situation of the worst-off features). **The fourth column** (`ave_prop_rem`) represents the average proportion remaining over all features. **The fifth column** (`w_prop_rem`) gives the weighted average proportion of all features. **The sixth column** (`ext-1`) shows the average extinction risk of biodiversity features as landscape is iteratively removed (as calculated from the species-area relation using parameter *z*) and **the seventh column** (`ext-2`) is the weighted extinction risk where species-area extinction risk has been weighted by the feature weights. **The following columns** show the proportion of distribution remaining for each feature in the same order as the features are listed in the beginning of the file. Note, that for the output file to be readable, the program does not print every step of cell removal. This file includes a maximum of 10 000 rows.

.rank.asc file



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value -1
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -
1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
0170109 0.0177273 0.0267405 0.0252988 0.0219256 0.0225694 0.0221160 0.0223790 0.0264685 0.0316008 0.
.6935855 0.6889973 0.6890789 0.6885076 0.6957527 0.7121652 0.7180410 0.7206435 0.7243068 0.7499229 0.
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -
687 0.0552311 0.0387371 0.0511870 0.0377850 0.0380116 0.0379844 0.0431258 0.0470340 0.0592572 0.0707
408371 0.7584828 0.7783591 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -
7966 0.0364429 0.0368328 0.0363976 0.0364067 0.0538891 0.0382383 0.0568270 0.0760505 0.0911844 0.101
128 0.7135525 0.7370468 0.7567871 0.7787762 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
```

Example of a .rank.asc file.

This is a raster file representing the ranking of the landscape (i.e., the order of cell removal). The file includes all basic raster information as explained in the biodiversity feature distribution map files and a matrix of cell removal order. Here the cells receive a value between 0 and 1. Low values close to zero indicate that the cell has been removed in a early state of the process whereas cells with a high value are removed last.

.prop.asc file

This is a raster file similar to the .rank.asc file. Here, however, the matrix shows the proportional loss of distribution for the feature that has lost most of its distribution during the landscape removal process. For example, if a cell has a value of 0.7, it means that after removing that cell all features have at least 30% of their distribution left (the value 0.7 indicates that one of the features, which is doing worst after removing that particular cell, has lost exactly 70% of its distribution).

.wrscr.asc file

In addition to the .rank.asc and .prop.asc, a third map is output automatically. This is the .wrscr.asc file, where *wrscr* stands for weighted range size corrected richness. This map reports for each cell *i* the quantity

$$wrscr_i = \sum_j w_j q_{ij},$$

where w_j is the weight of feature *j* and q_{ij} is the fraction of the distribution of the features in the cell. The measure is simply a sum over features of the weighted fraction of features distributions occurring in the cell as measured from original input distributions. To illustrate, the cell could have many occurrences of widespread, low-weight features. In this case, despite high richness per se, the *wrscr* value would be low compared to another cell which does not contain many features but does have a significant fraction of the entire range of a few small range or endemic features with relatively high weights.

This map can be used as a scoring value for the cell, which can be useful when comparing two cells with a replacement cost value of zero: the cell with the higher *wrscr* value would be more important. *Wrscr* values could be used, for example, to inform agro-urban land use planning of the potential intrinsic conservation value of small land parcels.

It is emphasized that the *wrscr* measure does not take into account any complementarity or connectivity considerations and use of this measure does not replace a full Zonation analysis. Two areas could have equally high *wrscr* values but due to the occurrence of a completely different set of features, which is accounted for in a Zonation analysis but not by the *wrscr* measure. Note also that distribution smoothing and interactions influence *wrscr* values, as it is calculated from the data that is used in Zonation computations, and this data includes effects of all transforms done to input maps.

Useful information

.rank.asc, .prop.asc and, **.wrscr.asc** files are raster files that can be used to produce map images in GIS software. By default these output rasters are generated as ASCII files (.asc), but the default format can be changed (see Section 3.2.1 Command Line for more information). To import the output files into ArcGIS, choose "Import Data Source..." from "File" options in ArcView and select the correct raster as the import file type. Select the file to import, rename it, and **do NOT choose integer cell values**.
The rank-file can be loaded with the -l -command and can be used in landscape comparison.

.txt file

This file is written during the run. It keeps track of the files, settings, and analyses used during the runs and is useful for tracing errors and checking that everything happened as it was supposed to. Note that some error messages or warnings may appear here. The contents of this file should be checked after a serious analysis run to verify that the correct options have been used and that there are no worrisome error messages or warnings.

.run_info.txt file

This file will be created only after you have closed the program. You can use it to go back to see what happened in the analyses. The content of the .run_info.txt file is identical to that of the .txt file.

3.4.2 Optional output file formats

In addition to the output files Zonation produces automatically, it is possible to produce the raster file output in several different raster and picture formats. This is useful when formatting map outputs for publication purposes or further GIS analyses.

Image output

Zonation creates the color map with cell ranking in .jpg and .emf formats by default, but it is also possible to output the map in .png and .bmp formats. To output the additional image formats, type an additional parameter in the command line calling Zonation (see section 3.2.1 for details):

```
--image-output-formats [png] [bmp] [jpg] [emf]
```

Grid output

In addition to the ASCII rasters Zonation produces by default, the raster files can also be output in GeoTIFF (.tif) and Erdas Image file (.img) formats as well as compressed versions of both. All output formats except ASCII have 32 bit float element type. 'tif' produces a GeoTIFF file and 'img' an Erdas Imagine file. Compressed GeoTIFF uses DEFLATE compression. Specify the output raster formats by typing an additional parameter in the Zonation call (see section 3.2.1 for details):

```
--grid-output-formats [asc] [tif] [img] [compressed-tif] [compressed-img]
```

3.4.3 Output files from optional analyses

.grp_curves.txt

If the biodiversity features have been assigned to output groups, Zonation will create an additional output file: a .grp_curves.txt file. This file contains representation curves for minimum, mean, weighted mean, and maximum representation as well as weighted extinction risk during the course of cell removal for each group. The second column of this file specifies the solution cost (duplicated from the global .curves.txt file). In the per group weighted extinction risk columns, the species-area extinction risks are weighted using the weights of the features belonging to each particular group, similar to the ext2 column of the global .curves.txt file.

A groups file must be included in the analysis (see section 3.3.3.12 for details and settings) to get this output. Output groups are specified in the first column of the groups file. Each biodiversity feature can be linked to an output group.

Groups could be assigned, for example, to:

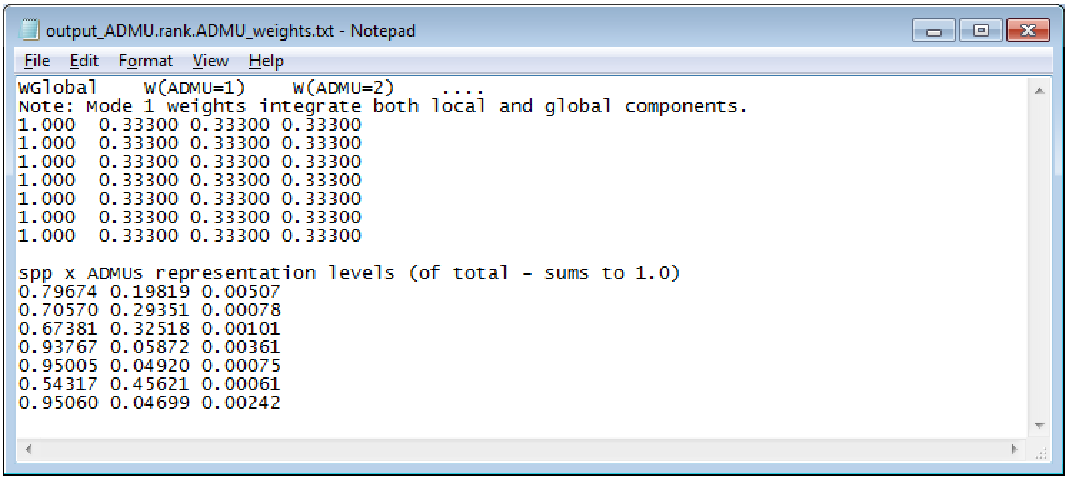
- Different higher taxa, birds, mammals, etc.
- Community and species features
- Negatively and positively weighted features
- Habitat quality and connectivity layers
- Any combinations of the above

All this information can also be computed from the .curves.txt file, but that requires manual operations and can be tedious. Thus, the grouped output is meant for saving manual work.

3.4.3.1 Output files specific to administrative units analyses

ADMU_weights.txt

This output file is produced by the administrative units analysis. The file includes a matrix describing the joint effective weights for each administrative unit and feature, as combined from global and local priorities and the local weights assigned a priori.



```
WGlobal     W(ADMU=1)    W(ADMU=2)    ....
Note: Mode 1 weights integrate both local and global components.
1.000  0.33300 0.33300 0.33300
1.000  0.33300 0.33300 0.33300
1.000  0.33300 0.33300 0.33300
1.000  0.33300 0.33300 0.33300
1.000  0.33300 0.33300 0.33300
1.000  0.33300 0.33300 0.33300
1.000  0.33300 0.33300 0.33300

spp x ADMUs representation levels (of total - sums to 1.0)
0.79674 0.19819 0.00507
0.70570 0.29351 0.00078
0.67381 0.32518 0.00101
0.93767 0.05872 0.00361
0.95005 0.04920 0.00075
0.54317 0.45621 0.00061
0.95060 0.04699 0.00242
```

Example of an ADMU_weights.txt file.

ADMU.redistributed.rank.asc file

As opposed to the normal rank file, the rank values found within each administrative unit in this map are redistributed in the [0,1] range. This way it is easy to identify a certain top fraction for every administrative unit. Similar to the .rank.asc., .prop.asc, and .wrscr.asc files, this is a raster file and can be used to produce map images in GIS software. This redistributed rank could be generated by means of GIS software, but the process can be tedious and/or complicated. It is produced automatically by Zonation for convenience.

It should be noted that the rank given in this file does not correspond to the values that are generated in the .curves.txt file and other curves files. Those results are generated for the normal .rank.asc output. To look at the curves for the redistributed solution, generate them by using Zonation in loading mode (-l option, as described in section 3.2.2, "Loading previously calculated Zonation solutions").

ADMU.rank.per_ADMU_outputs folder

This folder is generated by the administrative units analysis if either of the two following conditions hold:

- Output groups have been defined
- The option "row count for per ADMU output curves" is specified in the settings file

The following files are generated for the two cases above:

ADMU.XX.grp_curves.txt files

If output groups have been defined, a series of ADMU.XX.grp_curves.txt files will be generated, where XX is the administrative unit number. These files are generated inside the

.rank.per_ADMU_outputs folder and contain minimum, mean, weighted mean, and maximum representation values for groups within administrative units. This file is given in the same format as in the general .grp_curves.txt files (but without the extinction risk column). The .grp_curves.txt and ADMU.XX.grp_curves.txt files also include a column (the second one) with the costs at each level of cell removal.

ADMU.XX.curves.txt files

These files are generated in the ADMU.rank.per_ADMU_outputs folder if the option "row count for per ADMU output curves" is specified in the settings file. The file names are ADMU.1.curves.txt, ADMU.2.curves.txt, etc. Each of these files describes biodiversity feature representation levels in each administrative unit as cell removal proceeds. The format of these files is the same as that of the global .curves.txt file but without the costs column.

Warning

Please note that in planning cases where thousands of biodiversity features are defined, the .curves.txt file may require available disk space in the order of hundreds of megabytes. Since the format of the ADMU.XX.curves.txt file is the same, when tens of administrative units are defined, ADMU.XX.curves.txt files may require gigabytes of available disk space. This is the reason why ADMU.XX.curves.txt files are generated only if the option "row count for per ADMU output curves" is set.

3.4.3.2 Corridor_boundaries_dir

When using the corridor building function in Zonation, enabling the option `output corridor boundaries layer at percentage` in the run settings file will generate a directory or folder that contains GIS layers specific to corridor analysis.

.corridor_boundaries folder (directory)

This folder contains a list of output raster maps that represent the boundaries of corridors for different ranking levels (the list of ranking levels is specified in the option `output corridor boundaries layer at percentage`, see section 3.3.2.3 Run settings file for details). The format of the files corresponds to the general output GIS format, which is GeoTIFF by default. Each of these files is named as `corridor_boundaries_top_25.compressed.tif`, where 25 denotes the top 25% of the ranking, and the extension may vary depending on the output GIS format selected.

3.4.4 Transformed layers output

Some methods implemented in Zonation operate by transforming the input biodiversity feature maps. The transformed maps (or layers) are then used in the ranking (iterative removal) process. These transformed layers can be generated as optional output files that can be used for inspection, post-processing, or to better understand how Zonation works.

Up to 8 different types of transformed layers can be generated. For each type of transformation, a new folder/directory is created where the transformed layers are saved. These transformed layers can be visualized in the GUI from the project view (project -->

instance --> Transformed layers).

Below is a list of transformations, their mnemonics, and the corresponding settings and output folders/directories. The output folders are named by appending a suffix to the project output name.

Info-gap uncertainty analysis (IG) (or distribution discounting); see section 2.5	Settings: "output final transformed layers". Layers are written into the directory "OUTPUT_transf_info-gap_layers" (where the initial part, "OUTPUT" is variable and denotes the project output name)
Distribution smoothing (DS); see section 2.4.2	Settings: "output distribution smoothing transformed layers". Directory: "OUTPUT_transf_distrib_smooth_layers".
Community similarity (CST); see section 2.8	Settings: "output community similarity transformed layers". Directory: "OUTPUT_transf_comm_simil_layers".
Condition (CT); see section 2.10	Settings: "output condition transformed layers". Directory: "OUTPUT_transf_condition_layers".
Retention (RT): see section 2.10	Settings: "output retention transformed layers". Directory: "OUTPUT_transf_retention_layers".
Matrix connectivity (MCT); see section 2.4.5	Settings: "output matrix connectivity transformed layers". Directory: "OUTPUT_transf_matrix_conn_layers".
Interactions (IA); see section 2.6	Settings: "output interactions transformed layers". Directory: "OUTPUT_transf_interactions_layers".
Final: result of all transformations	Settings: "output final transformed layers". Directory: "OUTPUT_transf_final_layers".

Transformations are performed in the same order as in the list above, so the transformed layers corresponding to any transformation other than info-gap can potentially go through several transformations. The transformations performed are contained in the name of the output file. For example, assume the directory "OUTPUT_transf_distrib_smooth_layers" contains the file "feat_5_species5_IG_DS.asc". This means that the 5th biodiversity feature (the 5th one listed in the biodiversity features file) read from the "species5" map went through two transformations: first info-gap and then distribution smoothing. For a more complicated example, assume the file "feat_1_species1_IG_DS_CST_CT_RT.asc" is found in the directory "OUTPUT_transf_final_layers". This means that feature number 1 read from the "species1" map went through the following transformations: info-gap (distribution discounting), distribution smoothing, community similarity, condition, and retention.

Note that not all transformations are applied on every input feature. For example, the interaction transformation is performed only on those features specified in the interactions file (section 3.3.3.8). Distribution smoothing is applied only to features for which the α parameter is non-zero. Similarly, different methods are applied to a varying number or set of features depending on the settings. For example, the same analysis could output the files "feat_3_species3_IG_CT_RT.asc" and "feat_4_species4_IG_DS_CT_RT.asc". This implies that distribution smoothing has been applied on the layer of feature 4 but not on feature 3.

3.5 Post-processing analyses & options

This section describes three types of post-processing analyses that can be conducted for solutions produced by the main Zonation analysis. This discussion concentrates on automated post-processing, which is defined in the Zonation analysis setup prior to doing the analysis itself. There are also some useful post-processing tasks that can be done using the Zonation GUI, including identification of the best or worst locations in the landscape and identification of the overlapping areas of two solutions using the merged map.

Note that some of the post-processing analyses can be done using GIS software, but with significantly more effort.

3.5.1 Automated post-processing

After completing several Zonation priority ranking analyses, it may be desirable to examine the properties of a solution. With Zonation v.2.0, it was possible to do such post-processing manually. In Zonation v.3.1 and higher, two post-processing analyses have been automated: identifying management landscapes and comparing the spatial overlap between two solutions. These options are described in more detail below.

To use automated post-processing, adjust the run settings file (section 3.3.2.3) to indicate that automated post-processing will be used. To do this, include the name of the post-processing file in the run settings file:

```
post-processing list file = ppa_list_file_name.txt
```

Here, the file ppa_list_file_name.txt will include one or several post-processing calls, which will be done immediately after the Zonation priority ranking has finished. See section 3.3.3.17 for information about how to structure the calls for a post-processing analysis.

3.5.1.1 Landscape identification

Landscape identification options allow identification of separate management landscapes based on the distance and similarity in biodiversity feature composition between two sites. Spatially distinct areas (consisting of multiple grid cells) in a Zonation solution can be classified into management landscapes. An area is joined to a landscape if it is close enough and similar enough in the feature composition to *any* other distinct area in the same landscape. Landscape identification is done for a given fraction of the landscape.

Management landscapes can be identified among:
- All cells in a certain top fraction of the solution (the original LSI analysis)
- A set of areas specified via an external mask file (LSM analysis)
- Areas that are both in the top fraction of a solution and within the area specified by

the external mask file

Running landscape identification

To identify management landscapes from a Zonation solution, rows describing the correct LSI, LSM, or LSB calls must be added to the automated post-processing file. Information in these rows includes the fraction of landscape to be included, maximum distance, and minimum similarity of cells to be included in the same management landscape, as well as the possible mask file to be used. Please see section 3.3.3.17 for details about the input file contents.

The landscape identification analysis produces two output files:

nwout.#.ras.asc file	Here the matrix indicates which cells belong to which management landscapes. Each landscape has an integer value starting from 1. If a cell has a value of -2, it means that cell has not been included in the given top fraction (see "Percentage of landscape" above). Remember that this file, as any of the ASCII files produced with Zonation, can be imported to GIS programs. However, when importing this file, integer format should be selected for the cell values!
nwout.#.spp_data.txt	A text file containing statistics for the biodiversity features in management landscapes. See section 3.5.1.2 for contents of this file.

The pound symbol (#) in the file name is generated by the ordinal number of the call to an LSI analysis: the first call produces .nwout.1.ras.asc, the second .nwout.2.ras.asc, etc. The beginning of the LSI output files will be the general output file name given in the Zonation call.

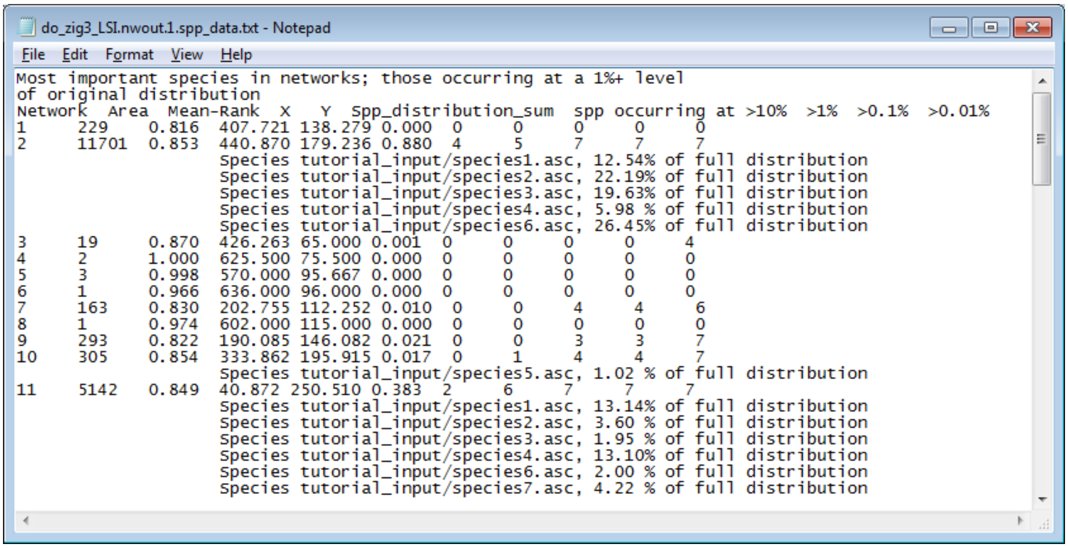
Literature

The landscape identification analysis is described by:

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and C.D. Thomas. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

3.5.1.2 Statistics for management landscapes

The landscape identification procedure also produces a text file (network feature data) containing statistical information about the management landscapes and occurrences of features within them.



```
Most important species in networks; those occurring at a 1%+ level
of original distribution
Network  Area  Mean-Rank  X   Y  Spp_distribution_sum  spp occurring at >10%  >1%  >0.1%  >0.01%
1     229    0.816  407.721 138.279 0.000  0     0     0     0     0
2     11701  0.853  440.870 179.236 0.880  4     5     7     7     7
                    Species tutorial_input/species1.asc, 12.54% of full distribution
                    Species tutorial_input/species2.asc, 22.19% of full distribution
                    Species tutorial_input/species3.asc, 19.63% of full distribution
                    Species tutorial_input/species4.asc, 5.98 % of full distribution
                    Species tutorial_input/species6.asc, 26.45% of full distribution
3     19     0.870  426.263 65.000 0.001  0     0     0     0     4
4     2      1.000  625.500 75.500 0.000  0     0     0     0     0
5     3      0.998  570.000 95.667 0.000  0     0     0     0     0
6     1      0.966  636.000 96.000 0.000  0     0     0     0     0
7     163    0.830  202.755 112.252 0.010  0     0     4     4     6
8     1      0.974  602.000 115.000 0.000  0     0     0     0     0
9     293    0.822  190.085 146.082 0.021  0     0     3     3     7
10    305    0.854  333.862 195.915 0.017  0     1     4     4     7
                    Species tutorial_input/species5.asc, 1.02 % of full distribution
11    5142   0.849  40.872 250.510 0.383  2     6     7     7     7
                    Species tutorial_input/species1.asc, 13.14% of full distribution
                    Species tutorial_input/species2.asc, 3.60 % of full distribution
                    Species tutorial_input/species3.asc, 1.95 % of full distribution
                    Species tutorial_input/species4.asc, 13.10% of full distribution
                    Species tutorial_input/species6.asc, 2.00 % of full distribution
                    Species tutorial_input/species7.asc, 4.22 % of full distribution
```

Example of the first part of the network feature data text file.

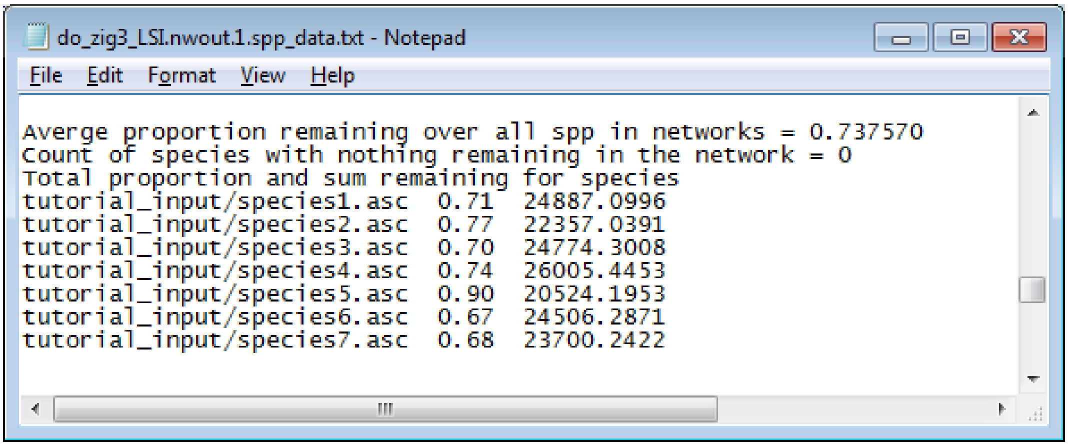
The first part of the file contains statistics about biodiversity feature occurrences in each landscape. This part is divided into eleven columns:

1. Identifier of the management landscape
2. Area of the management landscape in number of cells
3. Mean of the priority rank
4. Mean X-Coordinate
5. Mean Y-Coordinate
6. Sum of feature distribution proportions: this value shows what proportion of the **original** feature distributions the respective management landscape covers
7. Number of features which have more than 10% of their original distribution located in the management landscape
8. Number of features which have more than 1% of their original distribution located in the management landscape
9. Number of features which have more than 0.1% of their original distribution located in the management landscape
10. Number of features which have more than 0.01% of their original distribution located in the management landscape
11. Number of features which have more than 0.001% of their original distribution located in the management landscape

If the five last columns are marked as zero, it means that in that particular management landscape the feature occurrence is less than 0.01% of the full distribution. Note that if any of the features have a proportion larger than 1% of their distribution located in the landscape, the program automatically prints a list of those features and the precise proportions of their distributions on the next rows.

The second part shows the proportion of feature distributions remaining in the entire landscape (all management landscapes together) that was initially included in the analysis (the percentage of landscape defined in column 1 in the LSI command, section 3.3.3.17).

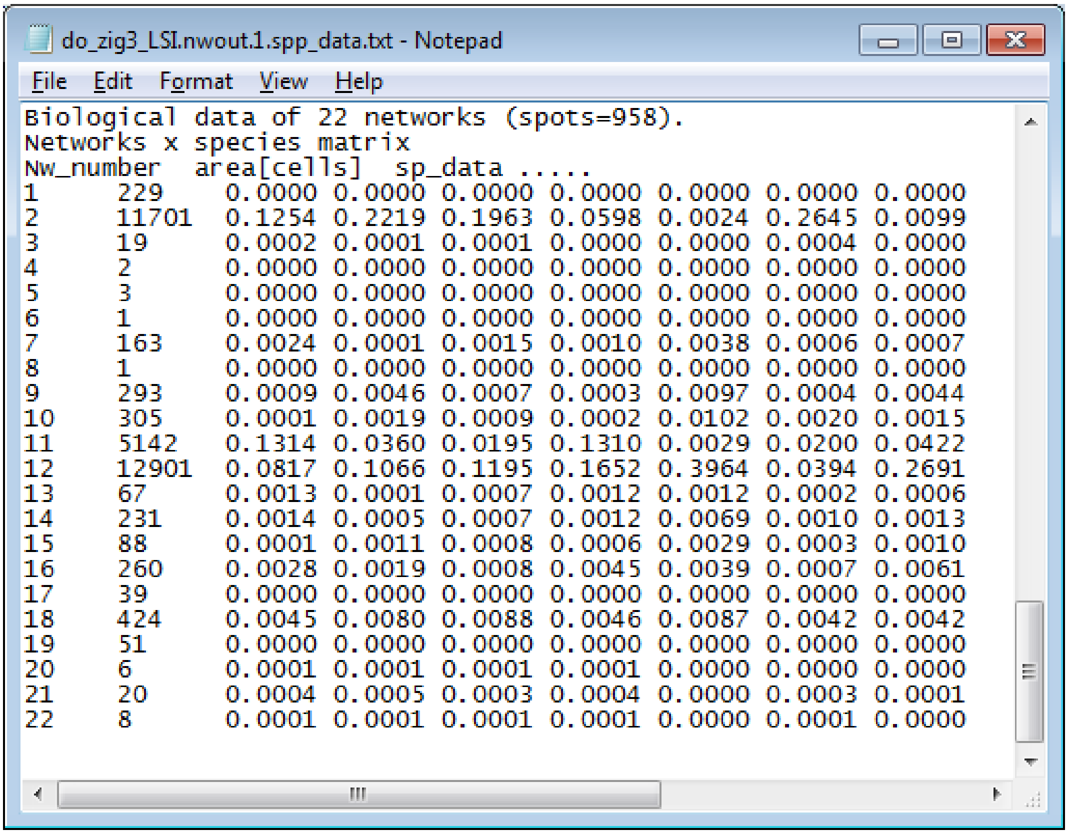
The program also automatically calculates an average of these proportions.



```
Averge proportion remaining over all spp in networks = 0.737570
Count of species with nothing remaining in the network = 0
Total proportion and sum remaining for species
tutorial_input/species1.asc  0.71  24887.0996
tutorial_input/species2.asc  0.77  22357.0391
tutorial_input/species3.asc  0.70  24774.3008
tutorial_input/species4.asc  0.74  26005.4453
tutorial_input/species5.asc  0.90  20524.1953
tutorial_input/species6.asc  0.67  24506.2871
tutorial_input/species7.asc  0.68  23700.2422
```

Example of the second part of the network feature data text file.

The third part contains a list of all management landscapes, area in number of cells, and the proportions of distributions for each feature in the respective management landscape. The features are listed here in the same order as they are in the biodiversity feature list file.



```
Biological data of 22 networks (spots=958).
Networks x species matrix
Nw_number   area[cells]  sp_data .....
1       229     0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
2       11701   0.1254 0.2219 0.1963 0.0598 0.0024 0.2645 0.0099
3       19      0.0002 0.0001 0.0001 0.0000 0.0000 0.0004 0.0000
4       2       0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
5       3       0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
6       1       0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
7       163     0.0024 0.0001 0.0015 0.0010 0.0038 0.0006 0.0007
8       1       0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
9       293     0.0009 0.0046 0.0007 0.0003 0.0097 0.0004 0.0044
10      305     0.0001 0.0019 0.0009 0.0002 0.0102 0.0020 0.0015
11      5142    0.1314 0.0360 0.0195 0.1310 0.0029 0.0200 0.0422
12      12901   0.0817 0.1066 0.1195 0.1652 0.3964 0.0394 0.2691
13      67      0.0013 0.0001 0.0007 0.0012 0.0012 0.0002 0.0006
14      231     0.0014 0.0005 0.0007 0.0012 0.0069 0.0010 0.0013
15      88      0.0001 0.0011 0.0008 0.0006 0.0029 0.0003 0.0010
16      260     0.0028 0.0019 0.0008 0.0045 0.0039 0.0007 0.0061
17      39      0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
18      424     0.0045 0.0080 0.0088 0.0046 0.0087 0.0042 0.0042
19      51      0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
20      6       0.0001 0.0001 0.0001 0.0001 0.0000 0.0000 0.0000
21      20      0.0004 0.0005 0.0003 0.0004 0.0000 0.0003 0.0001
22      8       0.0001 0.0001 0.0001 0.0001 0.0000 0.0001 0.0000
```

Example of the third part of the network species data text file.

Note that setting the maximum distance between cells to zero makes it possible to view the statistics of every single spatially distinct patch in the landscape. This, however, increases the running time. A larger maximum distance leads to fewer management landscapes. Note that the program presently only allows identification of up to 30 000 landscapes.

3.5.1.3 Solution comparison

Solution comparison calculates how much two solutions overlap with each other and the average difference in the cell removal order. The comparison is always made between the present and a previously computed result by using the rank.asc files of both solutions as input files. Please note that a function similar to automated solution comparison can be implemented using the map merge function of the new GUI.

Running automated solution comparison

To conduct a solution comparison, add the correct automated post-processing call into the post-processing file (section 3.3.3.17). This is a line that has the following information:

Column 1: Type "`LSC`" to indicate that two Zonation solutions will be compared
Column 2: The top fraction of the **new solution** that will be included in the comparison (value between 0 and 1)
Column 3: The top fraction of the **previously calculated solution** that will be included in the comparison (value between 0 and 1)
Column 4: The file name of the previously calculated solution ranking map to be loaded (e.g., `my_previous_solution.rank.asc`)
Column 5: The file name for the output map showing the top fractions included for both solutions and the overlap between them (e.g., `overlap_between_1and2.asc`)

Output

.ras.asc file	The values in the matrix are as follows: 0. The cell is not included in the top fraction 1. The cell is included in both solutions (overlapping areas marked as yellow) 2. The cell is included only in the present solution (light green areas) 3 The cell is included only in the older solution (dark green areas)

Remember that this file, as any of the ASCII files produced with Zonation, can be imported in GIS programs. However, when importing this file, select integers as the format of the cell values!

3.5.2 Solution cross-comparison using solution loading

Solution cross-comparison using solution loading is used in Figure 4 of Moilanen & Wintle (2007) *Conservation Biology*, 21:355-364, where effects of BQP on solutions originally calculated with the BLP are evaluated.

This is a major analysis, which can produce very important information including:

- The effectiveness of a solution produced without connectivity criteria if connectivity is actually needed
- How much (apparent) conservation value is lost if a solution is developed with connectivity when connectivity is not actually needed
- Analysis as in previous two points for the inclusion/exclusion of interactions between biodiversity feature distributions
- Surrogacy analysis: a solution developed for one set of features can be evaluated for performance across a completely different set of features
- Representation curves for the original, unexpanded community types in a community level analysis (section 5.3.1)

Overall, the idea is that a solution can be developed using one set of criteria but evaluated post hoc using another set of criteria.

When loading an old solution, the program does not just display the solution, but removes cells from the landscape based on the ranking order of the old solution. Thus it is possible to use new settings to evaluate the performance of a network with different settings. For example, a basic Zonation analysis can be conducted and then the performance of the resulting network if uncertainty or boundary quality penalty would be included in the analysis can be tested. Full output will be produced from the loaded analysis, but cell removal order will be enforced according to the .rank.asc file that is loaded.

To do this, load the old solution with new settings (see section 3.2.2 for instructions). Type "`-lfilename`" as the second parameter of your call. Remember also to enter the correct name of the adjusted settings file as the third parameter in the call.

The curves.txt files of the two solutions reveal the proportion of each feature's distribution remaining when landscape is iteratively removed. Tutorial examples of solution loading are provided in exercises 6 and 10.

Note that some options of the settings file are ignored when Zonation runs in load old solution mode (-l) because they would not make sense. This includes the options edge removal and add edge points. This implies that the same settings files that were used to generate the original solution can be safely reused to load that solution.

3.6 What Zonation does NOT do directly

Multi-action planning

Zonation does not directly select from multiple alternative conservation actions. Different scenarios where different actions could be applied can be generated, and then Zonation can be used to prioritize among those spatially fixed actions. (See Thomson *et al*. 2009; Moilanen, Leathwick & Quinn 2011).

Dynamics

Zonation does not model stochastic landscape or species dynamics. Dynamics can, however, be faked by including predictions for several time steps in the analysis (see e.g., Carroll *et al*. 2010; Thomson *et al*. 2009). Zonation is not a system for spatial population viability analysis (PVA).

Habitat modeling

Zonation is not a statistical habitat modeling tool, and it cannot produce species distribution models, niche models, or similar. Rather, Zonation uses habitat models as input. Zonation processes the pattern of biodiversity occurrence after distribution modeling has first been done.

Fixing inadequate data

Zonation does not fix data. Low quality data will produce low quality results. Zonation will produce a priority ranking no matter what data is used, but the value of the ranking depends on the quality of inputs.

Polygon data

Zonation does not operate on GIS data in polygon vector format. Such data needs to be converted into raster format before it can be used as Zonation input. Flexibility in the raster type and increased memory capacity alleviate the data intensity of raster files.

3.7 Data limitations & system requirements

Zonation v.3.1 and v.4 have been developed for both 32- and 64-bit operating systems. The 32-bit software can use approximately 2GB of memory. The 64-bit version can use 4GB*4GB, which is sufficient for most large analyses (see memory requirements below). Presently a desktop PC can easily have 32GB of memory or more. High-end workstations can have hundreds of GB or TB of memory. Increased memory capability allows the use of much more data than with v.2.0. More complicated setups and more feature layers are possible with increased memory.

As the memory capacity of Zonation has increased massively in version 3.1 and higher, the software now becomes limited by speed for large data sets. To alleviate this, the Zonation v.4 core has been optimized for large datasets. Also, the core has been made partially multi-threading, which multiplies computation speed by a small integer factor. Suitable organization of data and analysis settings allow for initially unlikely analyses to be possible. For more involved examples see (full references in section 2.1) :

- Thomson *et al*. 2009 (habitat restoration)
- Carroll *et al*. 2010 (climate change)
- Gordon *et al*. 2009 (urban planning)
- Lehtomäki *et al*. 2009 (extensions of forest conservation areas with relatively complicated connectivity arrangements)
- Rayfield *et al*. 2009 (utilizes several different connectivity components per species)

An additional advantage of increased memory is that it enables faking of landscape dynamics. Faking dynamic landscapes requires solutions that are balanced at all time steps and can by accomplished by entering layer sets for the present and for several time steps in the future.

Hard-coded limitations, which may be alleviated in later Zonation versions

All analysis	
Maximum number of biodiversity features:	~65 000
Maximum SSI species (point occurrence lists):	~25 000 (thus total maximum is 65 000 + 25 000)
Boundary Quality Penalty (BQP):	
Maximum number of penalty curve profiles:	50
Maximum number of points on a penalty curve:	20
Maximum number of different species-specific buffer sizes:	100
Landscape identification	
Maximum number of management landscapes:	Since Zonation v.4 there is in principle no hard-coded limit.

Memory requirements

Memory requirements depend on the data being used. Starting with Zonation v.4 it is preferable to mark areas (cells) where features do not occur and cannot occur as missing or no-data cells, as this can save significant amounts of memory. The amount of memory needed will depend on the extent of biodiversity features. The smaller the range of features, the more features that can be processed with approximately the same RAM memory. The Zonation core produces a report on the amount of memory required to load all the biodiversity features included in a setup. This can be checked in the output text files or in the graphical user interface (text output). Naturally, larger data sets (many features and/or high resolution and/or large landscape) will need more memory to run the computations. One raster file of one million effective elements (cells that have real data, not missing values) for one species requires 6 MB of memory. Thus the maximum number of species that can be used can be estimated with the help of this formula:

> Max. features or species = 0.7 * memory in MB / (6 * effective grid size [in millions of elements])

The 0.7 in the formula accounts for the memory needs of the operating system and the memory needs of Zonation in addition to the feature/species data matrices, and it gives a pessimistic or conservative estimate of memory requirements. Note that his is just an approximate rule of thumb. With 6 GB (= 6 000 MB) of memory you can have approximately (0.7 * 6 000)/(6 * 5) = 140 species with 5 million effective grid elements each. But the total amount of memory required will depend on the effective range of the features (as long as you use no-data for the cells outside of the features ranges). For example, if the size of the analysis area or landscape is 5 million cells but the effective range of the features is on average 1 million cells, then you would be able to process 700 features or species with the same amount of memory.

Note that some analysis features may significantly increase memory needs. For example, using BQP mode 2 (feature-specific missing data areas) approximately triples the memory consumption compared to an analysis where BQP is not used, leading to a respective loss in the number of features that can be used in the analysis. Interaction layers count as independent features for the purpose of memory computations.

The estimate of memory consumption described here is only indicative, but sufficient for getting an idea if an analysis definitely should or should not run. Check the amount of available RAM in Windows task manager to verify that Zonation has not run out of memory.

3.8 Troubleshooting

The following is a short list of things to check when encountering problems.

- **Directory paths**
  Check that the paths names are correct so that the program can find them. If the program is in the same directory as the input files (recommended), the paths are not needed—just the file names. However, if some or all of the input files are located in a different directory, a directory path is needed. If files are in a subdirectory from the exe-directory, then filenames can be entered as sub_dir_name\filename.

- **File names**
  One reason for problems can be long directory and file names, especially when running the program from the command prompt. Thus try to keep the directory names short (e.g., max. 8 characters). Also, do not use any spacings in the directory or file names. In Zv3 and higher this is not a problem in principle, but avoiding special characters and spaces in file names is nevertheless safest.

- **Computer memory capacity**
  If the program is running very slow during computations, check Task Manager/ Performance. If the RAM (Physical Memory) is close to zero, the memory is full. This does not mean that the program has jammed, but it will take a very long time for it to finish the calculations. Stop the computation and try again after closing all other programs to save memory. The analysis can also be run with a smaller data set or with a computer that has more memory. It is not ok to have Zonation running using virtual memory (the hard drive) because that will simply run too slow. The Zonation core reports the amount of memory needed to load biodiversity features in the output text files (or text memo in the graphical interface).

- **Operating system**
  Zonation 4 work on Windows and GNU/Linux operating systems. Zonation is provided for Windows as 32 bit or 64 bit software, which should be operational on Windows XP, 2000, 7 and newer versions. The 32 bit software runs on both 32 and 64 bit Windows versions, and the 64 bit Zonation runs only on 64 bit Windows. Zonation 4 also works on GNU/Linux systems.

- **Check the memo**
  Some warnings or error messages appear in the memo. Read through the text to check for any information that might give a clue to solving the problem.

- **α values**
  Check the α values for any errors. Remember that these values have to be in the **same unit of length** as the cell size given in the biodiversity feature distribution map file. It is very easy to get these values wrong in the first calculation, so be sure to verify computations. A common error is that alpha is given per kilometer (alpha on the order of 1.0) when raster file units are in meters. In this case the α values are 1000x too large, leading to a solution which is for practical purposes identical to the one when connectivity is not used at all.

- **Decimal points and commas**
  Always remember to use only decimal points in the settings files—this is required by the core! NO commas! Zonation assumes decimal dots, and commas will result in undefined errors. The default decimal setting can be changed (from decimal commas to dots) from the Windows control panel.

- **Empty rows at the end of your input files**
  Check that there are no empty rows at the end of the input files. With bad luck this might cause some unexpected software behavior.

- **Differences in grid sizes/cell sizes**
  All raster files should have the same grid size. This means that in all files the number of columns and rows as well as the size of cells should be equal. This information can be verified from the beginning of each raster file.

- **No data rows at the edges of species distribution matrices**
  Computational efficiency requires the input data to have at least one row of no data

on each edge of the biodiversity feature distribution grids. Otherwise the program will automatically transform all values on the edge rows to missing data. Usually this is not an issue, but it can create problems if the data goes right to the edge of the area.

- **Differences in grid matrices between species distribution and other data files**
  Check that all biodiversity feature distribution rasters are congruent with any other raster files used in analysis (e.g., cost layer, uncertainty layers, etc.) This means that all cells in a grid which have data for any of the biodiversity features used in the analysis (that is to say, the cells that are NOT marked as "no data" in all biodiversity feature distribution files), also have to have a value in the optional raster grids. Equally, all cells marked as "no data" in all biodiversity feature distribution rasters should have the same definition in any optional rasters. Note that the analysis area can be masked using the analysis area mask.

Part

IV

4 Zonation Graphical User Interface

The Zonation graphical user interface (GUI) (executable zig4gui) is a front end to the Zonation core (or command line application, executable zig4) that also provides several useful visualization features. It is possible to create new Zonation projects or setup, manage analysis settings and input files, and monitor the runs while they happen in the GUI. With the GUI, outputs and results can be visualized and exported in various formats. This is a quick guide to the Zonation GUI. The GUI has significantly improved in versions 3.1 and 4, but you can always check the Zonation website (http://cbig.it.helsinki.fi/software/zonation/) for further updates.

GUI versions

Please note that the Zonation GUI has been under development, and there may be small differences in appearance between this manual and past, present or future GUI versions. The basic layout of the GUI has stayed relatively stable though, and it is expected that small additions or changes in appearance should not significantly alter the way Zonation is used.

GUI Features

- Zonation project management
- Create new projects with the "project maker", without having to edit text files
- Monitor parallel Zonation calculations at runtime
- Visualize input and output maps in a very simple GIS graphical interface
- Multiple predefined color schemes for the ranking and other maps, and flexible editing of custom schemes
- Plot results/performance curves for different features, groups, administrative units, etc. interactively
- Loading existing configuration and batch files
- View and edit input files
- Edit and save output maps with custom color schemes
- Compare maps between different analyses
- Raster calculator to pre- and post-process input and output maps.

Technical details

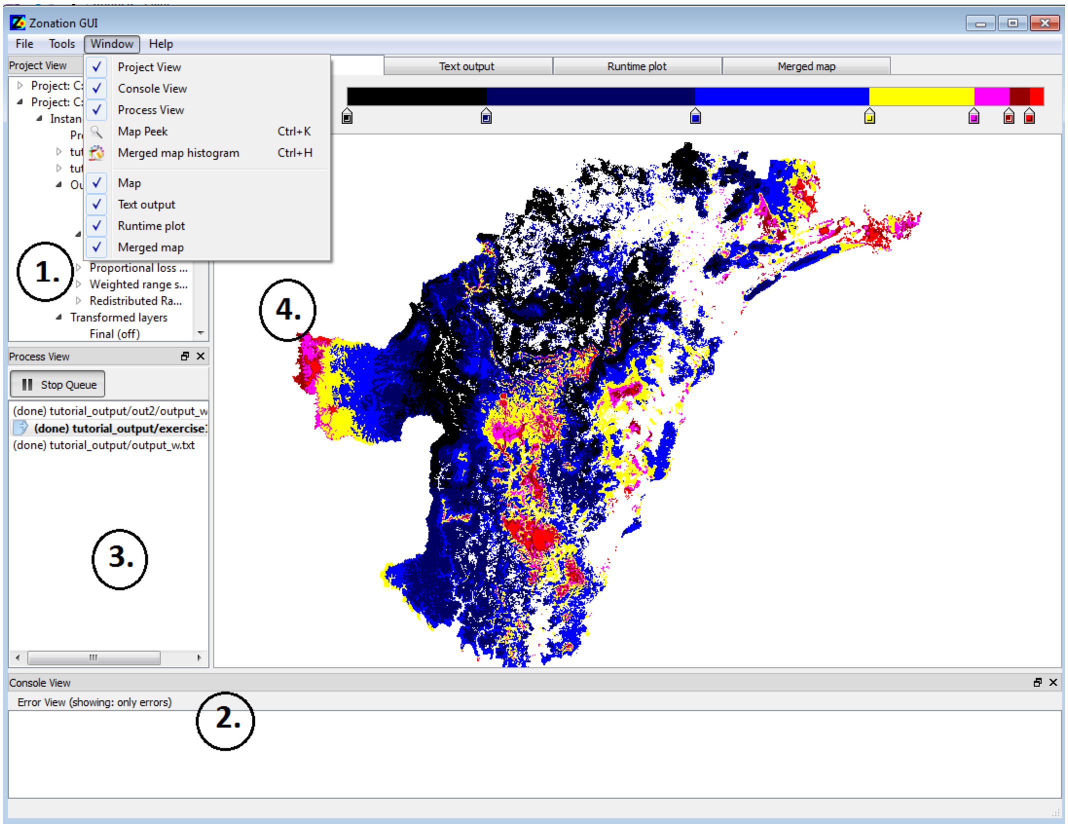
The Zonation GUI is written in C++ using the Qt toolkit. It utilizes GDAL for GIS raster support. The GUI uses the following libraries: Qt (http://qt-project.org/), Qwt (http://qwt.sourceforge.net), boost (http://www.boost.org) and GDAL (http://www.gdal.org). We also used various versions of GCC, the GNU compiler collection (http://gcc.gnu.org). These free open source software projects are greatly acknowledged.
The inter-process communication between the GUI and the command line process is handled through a shared memory mechanism.

4.1 Main menu

Four drop-down menus are available in the main menu of the Zonation GUI: File, Tools, Help, and Window. The File menu opens new or recent projects. The Tools menu is used for defining preferences. The Zonation User Manual can be accessed in html format from the Help menu, and the Window menu allows selection of visible windows.

The GUI has four different windows (shown and numbered in the image below):

1. **Project View:** Call, edit, and visualize input files (see section 4.2)
2. **Console View:** Track error messages and warnings during the runs
3. **Process View:** Monitor the progress of analyses (see section 4.2)
4. **Visual output:** Monitor the progress of analyses (see section 4.2) with the following tabs:
    4.1. **Map:** Visualize input and output maps (see section 4.3.1)
    4.2. **Text output:** Keep track of all settings, input files, and analyses included in one run (see section 4.3.2)
    4.3. **Runtime plot:** Visualize the performance of biodiversity features during cell removal (see section 4.3.3)
    4.4. **Merged map:** Visualize a color-wise combination of maps (see section 4.3.4)

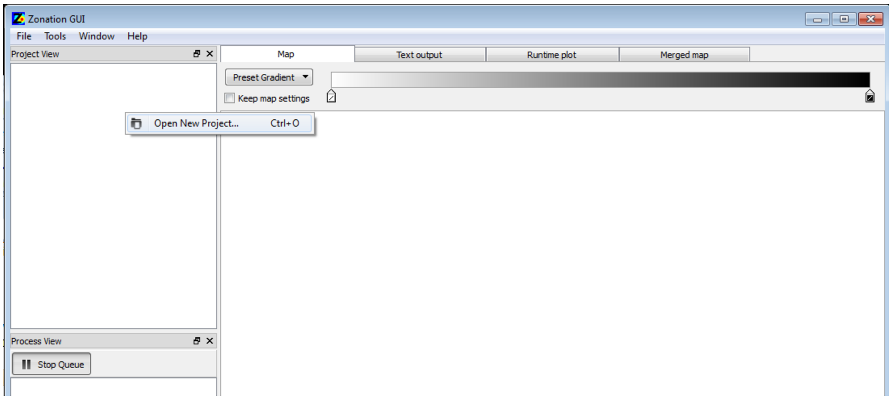


In addition to the four main windows described above, there are two auxiliary windows for inspecting map details (see section 4.3.1) and visualizing the list of colors in the merged map (see section 4.3.4).

4.2 Project management

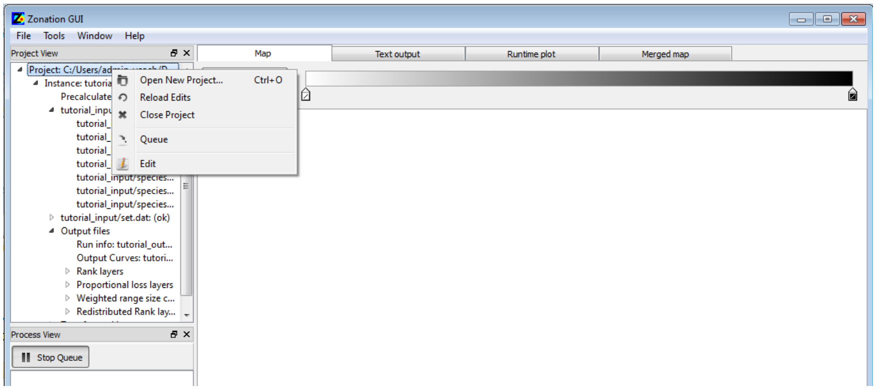
Open the Zonation GUI (zig4gui.exe) from the start menu or by double-clicking on the desktop icon.

The Project View window is located in the upper left section of the GUI. To load a project, right-click on the white area in the Project View window and select "**Open Project**" to browse and select the appropriate batch file for the analysis. After selecting the desired project, the batch file as well as all the input files referred to in the batch file will appear in the Project View window.



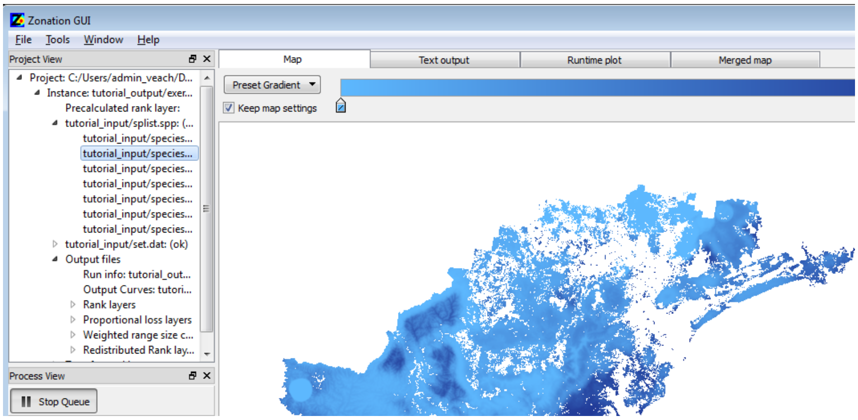
Example showing how to load a project from the Zonation GUI.

Clicking on the small icon to the left of the batch file expands the project file to show the hierarchy containing all input files that are referred to in the batch file as well as the linkages between the input files. The input files can be opened and edited by right-clicking on the files and selecting "**Edit**". Any changes will be saved in the files, not the GUI. The project must be refreshed for the changes to appear in the GUI. This can be done by right-clicking on a project and selcting "**Reload Edits**". This can also be done automatically by selecting "**Reload from disk before enqueuing projects**" in the global preferences dialog box. Editing project files may result in structural changes in the project. This happens, for example, if new instances are added to the project file or if new features are added to the biodiversity feature list files.

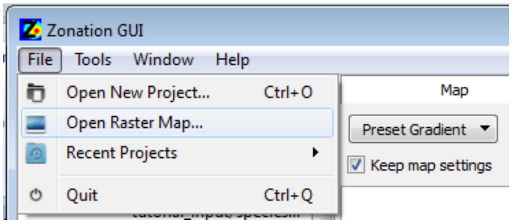


Example showing the 'Edit' and 'Reload Edits' options.

Input raster files can be viewed in the Map tab of the visual output window (section 4.3.1) by double-clicking on the file name. Raster maps can also be opened from the File menu by selecting **File -> Open Raster Map**. This option opens a new window to browse for the desired map file.

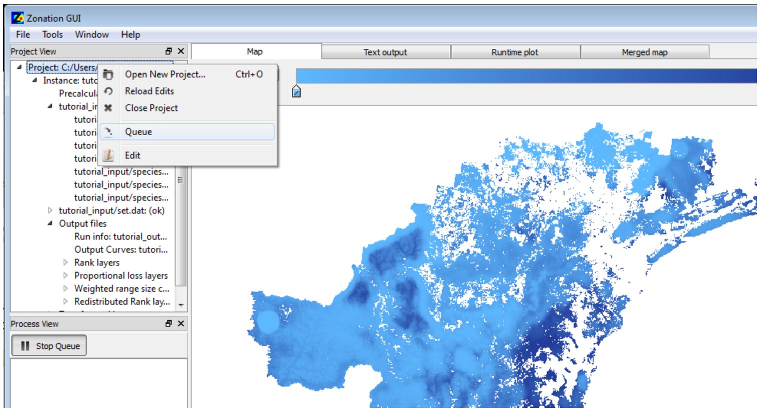


Example of input raster file show in the Map tab of the Visual Output window.



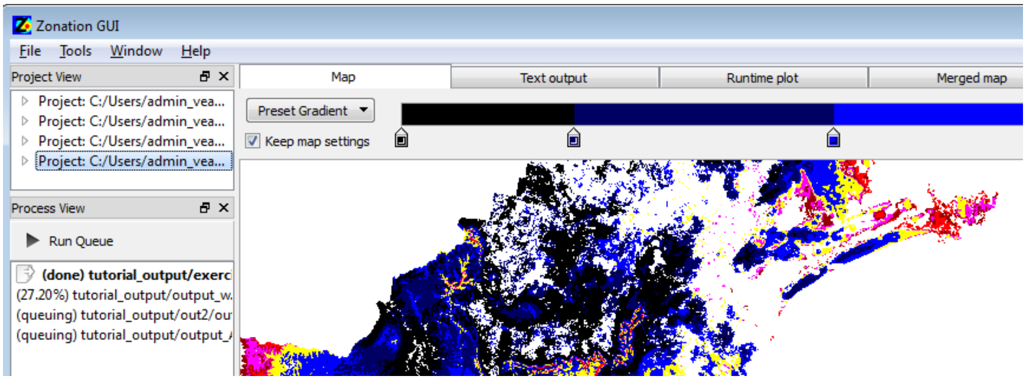
Example showing how to open a raster file from the File menu. A new window will open from which you can browse and select the desired file.

Each command line in a batch file corresponds to one Zonation run. In the GUI, each command line is listed as an 'Instance' and referred to by the output file name specified in the command line. It is possible to execute all command lines at the same time or one-by-one. Right-clicking either 'Project' (batch file) or 'Instance' (individual command line) in the Project View and selecting "**Queue**" adds the selected item to the work queue shown in the Process View window.



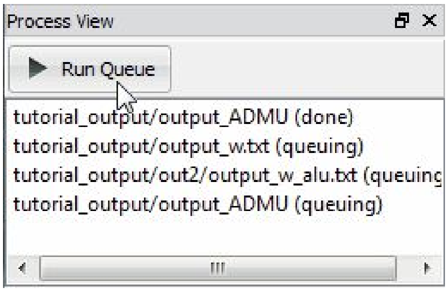
Example showing how to add a project file or an instance to the work queue.

The run will now appear in the Process View window located below the Project View. Double-clicking a line while Zonation is running will open the output map in the Map tab of the visual output window (see section 4.3).



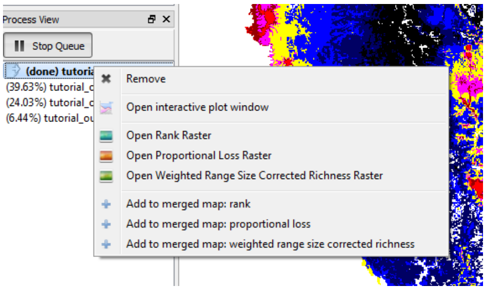
```
(done) tutorial_output/exerc
(27.20%) tutorial_output/output_w.
(queuing) tutorial_output/out2/ou
(queuing) tutorial_output/output_
```

The work queue can be started and stopped by clicking the Run Queue button located at the top of the Process View. Pausing the queue will prevent further processes from opening but will not stop the processes already running. Items can be removed from the queue by right-clicking on them and selecting Remove. Removing a running process will terminate it forcefully. In addition to the default rank raster, right-clicking the run provides options to display the proportion remaining (.prop.asc) and weighted range size normalized richness (.wrscr.asc) layers. Note that when displayed, the .wrscr.asc raster has been rescaled to the interval [0,1]. Thus, if looking at pixel-level information, the number shown is "weighted range size corrected richness relative to the maximum that occurs anywhere in the landscape".



tutorial_output/output_ADMU (done)
tutorial_output/output_w.txt (queuing)
tutorial_output/out2/output_w_alu.txt (queuinc
tutorial_output/output_ADMU (queuing)

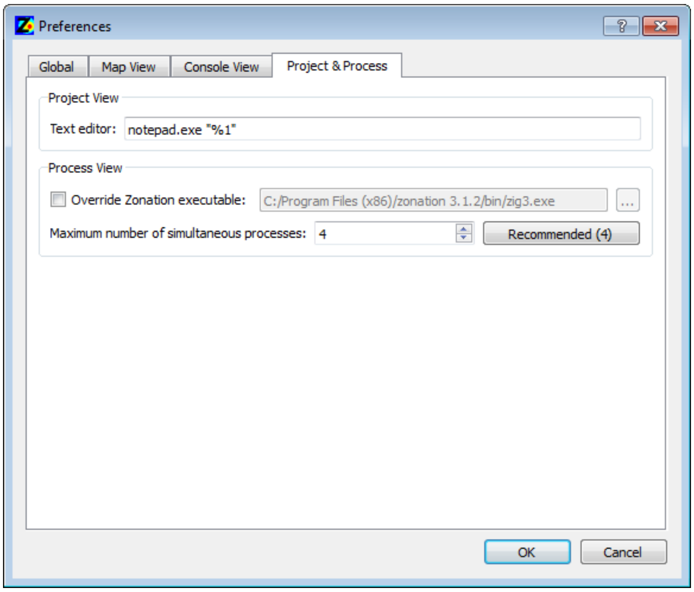
Process View window with work queue.



Example showing the options available to analyses in the Process View window.

By default, Zonation will call the core executable (zig4.exe) bundled with the GUI instead of the executable called in your batch file. This can be changed in the **Preferences** menu in **Tools-> Preferences**. The batch file can be any regular command prompt script. Please note, however, that the Zonation core executable called in the batch file must be reachable and must be of version 3.0.3 or higher. Otherwise the Zonation GUI will not be able to parse the batch file and your project will appear empty. Earlier versions (< 3.0.3) of Zonation have a different parser for batch files and are not inter-operable with later versions.

The Zonation GUI can run multiple analyses at once. By default, it won't try to run more processes at once than there are physical processors in the computer. The maximum amount of simultaneous analyses can be controlled from the Preferences menu in Tools-> Preferences. In general it is most efficient to have several runs progressing simultaneously, but the ability to do so may be limited by the RAM memory needs of the analyses and the memory available on the computer. Please keep in mind that all simultaneous Zonation analyses must fit into the physical RAM memory of the computer. You can check memory usage from the Windows task manager. Often in large Zonation analyses, the memory consumption rather than the number of runs will slow the computation. If this is the case, lower the maximum number of simultaneous processes allowed so that all of the analyses can fit in the memory. If the analyses need to use the virtual memory (the hard drive), the computations will become so slow that they will never finish.



Text editor: notepad.exe "%1"

☐ Override Zonation executable: C:/Program Files (x86)/zonation 3.1.2/bin/zig3.exe

Maximum number of simultaneous processes: 4

Example of Preferences dialogue box where the maximum number of simultaneous process can be set.

4.3 Visual output

The Zonation GUI produces three types of visual output accessible from the visual output window: the map from the cell ranking in the **Map** tab (section 4.3.1), a memo describing the analysis in the **Text output** tab (section 4.3.2), and performance curves describing the representativeness of the solution in the **Runtime plot** tab (section 4.3.3).

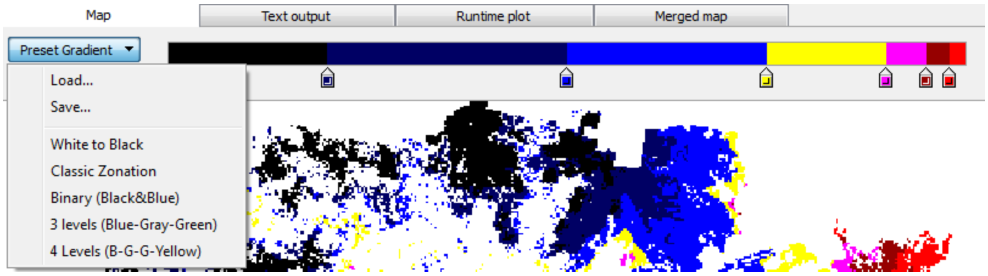
4.3.1 Map

The Zonation GUI shows the proceeding cell ranking and the resulting map in the **Map** tab of the visual output window. Double-clicking each run in the **Process view** (see section 4.2) will open the respective output map in the **Map** window.

Zooming in and out is possible with the scroll wheel on the mouse. The map can be panned by dragging it with the mouse.

Color schemes

The new Zonation GUI provides a tool for editing the output map to best suit individual analysis and visual preferences. The default map has a white to black gradient with black representing the highest priorities. Four additional color schemes can be selected from the **Preset Gradient** drop-down menu, including the following: **Classic Zonation, Binary (Black&Blue), 3 levels (Blue-Gray-Green),** or **4 levels (B-G-G-Yellow)**. All color schemes can be manually adjusted to suit your preferences.



Example showing the preset color schemes available in the Zonation v.3.1 GUI.

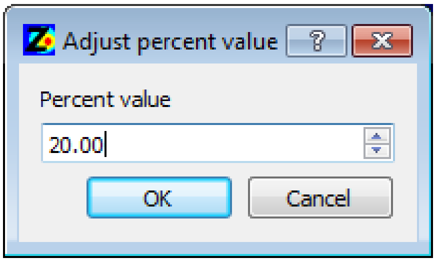
The Classic Zonation color scheme shows the nested ranking on a map. The ranking of sites is visualized by using different colors to indicate the biological value of the site:

Red	The best 2% of the landscape
Dark Red	The best 2-5%
Magenta	The best 5-10%
Yellow	The best 10-25%
Light Blue	The best 25-50%
Dark Blue	The best 50-80%
Black	The best 80-100% (or the least valuable 20%)

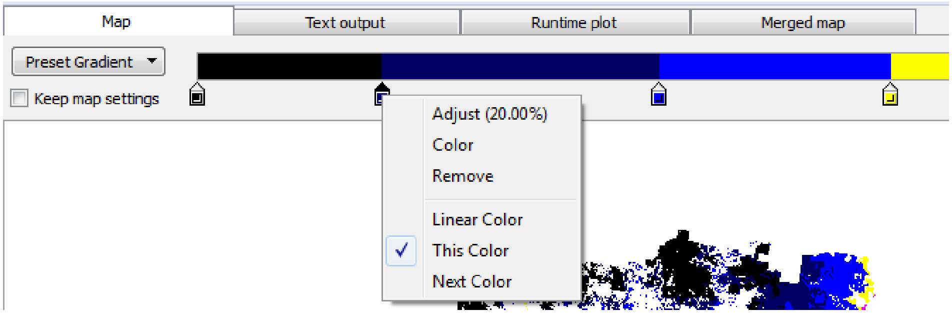
The information of this map is equal to the .rank.asc file that the program produces as part of file output. This map will also automatically be saved as a picture (output.jpg and output.emf files), but it can be saved again (e.g., with a different name or to a different directory) by right-clicking on the picture. The background (i.e., the cells for which no data exists) are shown in white. At the beginning of an analysis, before being overwritten by the ranking, locations with SSI species are shown as red dots.

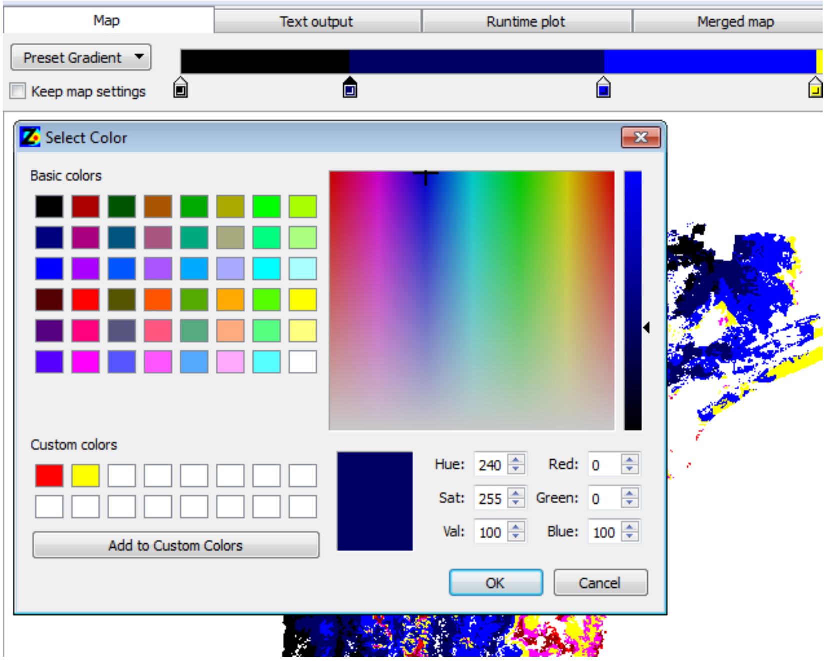
A **custom color gradient** can be designed and saved with the new GUI. The custom gradient can be used for many purposes including identifying any given top or bottom fraction of the landscape, identifying reserve network expansions in a hierarchic analysis, or identifying areas that satisfy a given minimum level for all species (features). The custom color gradient tool corresponds to the rank and remaining buttons of the Zv2 GUI.

Right-clicking on a color tab under the color scheme opens a menu from which the settings of that tab can be adjusted. By selecting "**Adjust**" the proportion of the priority ranking applied to that portion of the gradient can be defined. This is expressed in percentages 0-100% with 99 indicating the top 1%, 85 indicating the top 15%, 20 indicating the top 80%, etc. Selecting "**color**" opens a new window through which the color assigned to each tab can be adjusted. Right-clicking on the space beneath the color gradient adds a new color tab and opens a menu where the tab settings can be adjusted. Removing a tab is possible by selecting "**Remove**". The options "**This color**", "**Next color**", or "**Linear color**" adjust the appearance of the color scale. The options and windows described here can be seen in the screen captures below, but understanding how they work is perhaps easiest learned by experimentation.



Percent value: 20.00





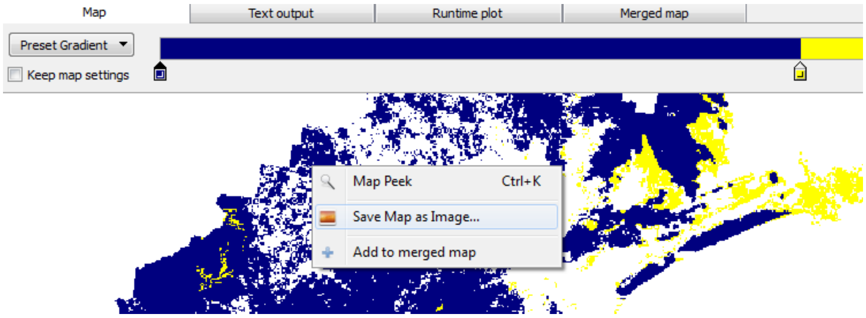
Hue: 240
Sat: 255
Val: 100
Red: 0
Green: 0
Blue: 100

A color scheme can be saved as a file (**.zgradient**) to use later with other analyses by selecting "**Save**" in the **Preset Gradient** menu. Previously saved color schemes can be loaded by selecting "**Load**" and browsing for the file through the same menu.

A custom color scheme can also be used to identify a reserve network expansion. Assume that a present reserve network covers 7% of the landscape and that a 1.5% expansion is planned. First, do a Zonation run where a hierarchy is forced in the analysis and the existing reserves are retained longest. Then adjust the output of this analysis so that one color tab is placed at 93 (corresponding to the top 7%) and another at 91.5 (the 1.5% expansion). Give these tabs different colors to show the expansion clearly.

Saving output map with a custom color scheme

During each run, Zonation automatically produces the ranking map in various picture formats (see section 3.4.1 for details). These maps follow the Classic Zonation color scheme. The output map with the custom color scheme can be saved in a **.png** format by right-clicking on the map and selecting "**Save Map as Image**" (see screen capture below).

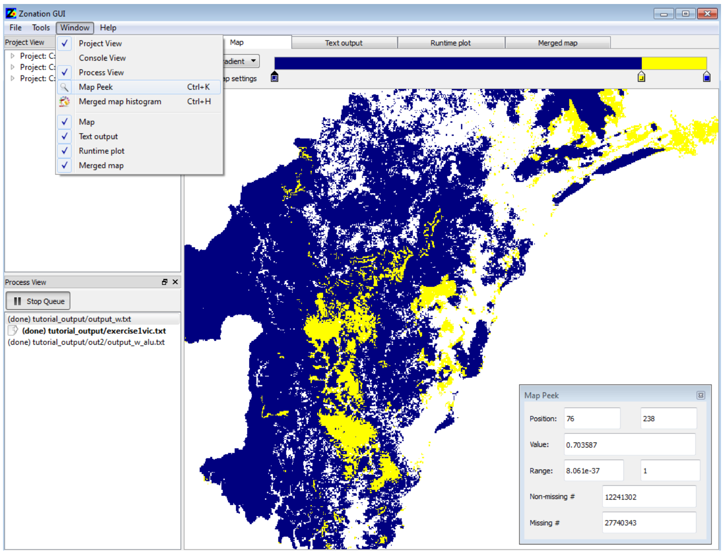


Visual comparison of outputs between runs

After a number of analyses have been completed, the outputs can be visually compared directly in the GUI. Double-clicking on an output file in the **Process View** will open the map of the ranking for that analysis in the **Visual Output** window. When switching between maps, it is often convenient to keep the zoom and pan settings as well as the color scheme the same. This is done by selecting "**Keep map settings"** in the Map tab of the **Visual Output** window. The merged map view (described in section 4.3.4) can be used for a visual analysis of overlaps between solutions.

Map Inspection

The Zonation GUI has a useful pixel inspection window which can be enabled from **Window ->** "**Map Peek**" or by selecting "Map Peek" after right-clicking on the map. With this option, the value and position of the pixel under the cursor as well as the range of values of the map and the number of missing and non-missing cells will be printed in the **Map Peek** window.

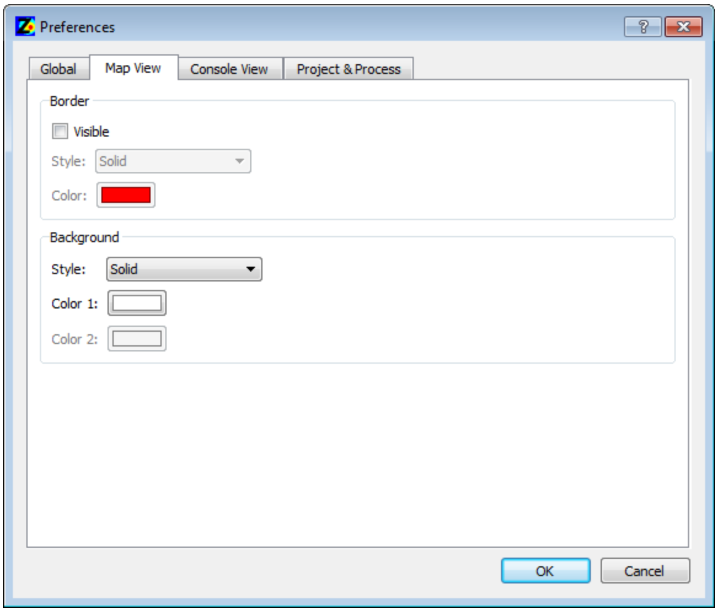


(done) tutorial_output/output_w.txt
(done) tutorial_output/exercise1vic.txt
(done) tutorial_output/out2/output_w_alu.txt

Position: 76 238
Value: 0.703587
Range: 8.061e-37 1
Non-missing # 12241302
Missing # 27740343

Map Preferences

The "**Map View**" tab in **Tools -> Preferences** contains settings for drawing the map border and background.



Border

☐ Visible

Style: Solid

Color: 🟥

Background

Style: Solid

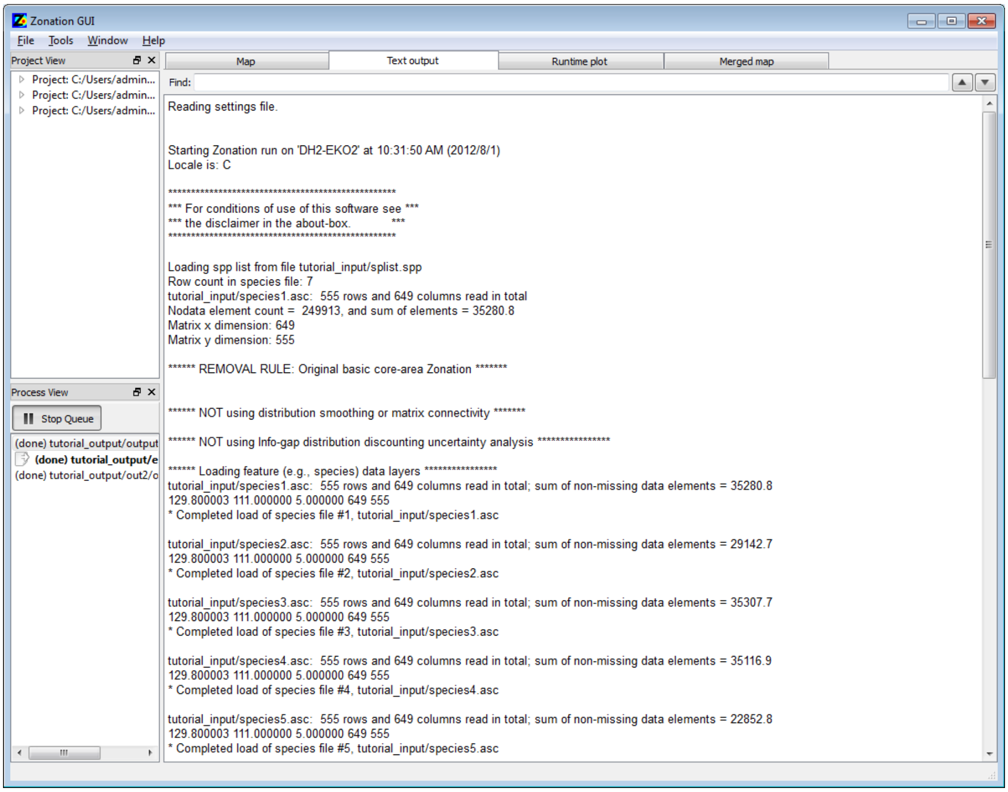
Color 1: ☐

Color 2: ☐

4.3.2 Text output

All calculations, settings, and input files used during a Zonation analysis are stored in a memo file (section 2.4.2). They are also automatically printed in the **Text** tab of the **Visual Output** window (see image below).

Error messages and warnings are also printed in this window. It is advisable to go through the memo to check that the correct input files and settings have been used and that no errors occurred during the analysis. The **Find** bar located at the top of the **Text** window can be used to search for certain key words. We recommend searching for 'error', 'warning', and 'unable'. Zonation v.4 is more verbose and produces extensive information on how features and other layers are loaded, how the ranking process works, how and what outputs are generated, etc. We recommend to read the text output to understand well what outputs Zonation produces, and whether there are any issues during the ranking process.



```
Reading settings file.


Starting Zonation run on 'DH2-EKO2' at 10:31:50 AM (2012/8/1)
Locale is: C

**************************************************
*** For conditions of use of this software see ***
*** the disclaimer in the about-box.           ***
**************************************************

Loading spp list from file tutorial_input/splist.spp
Row count in species file: 7
tutorial_input/species1.asc:  555 rows and 649 columns read in total
Nodata element count =  249913, and sum of elements = 35280.8
Matrix x dimension: 649
Matrix y dimension: 555

****** REMOVAL RULE: Original basic core-area Zonation *******


****** NOT using distribution smoothing or matrix connectivity *******

****** NOT using Info-gap distribution discounting uncertainty analysis ****************

****** Loading feature (e.g., species) data layers ****************
tutorial_input/species1.asc:  555 rows and 649 columns read in total; sum of non-missing data elements = 35280.8
129.800003 111.000000 5.000000 649 555
* Completed load of species file #1, tutorial_input/species1.asc

tutorial_input/species2.asc:  555 rows and 649 columns read in total; sum of non-missing data elements = 29142.7
129.800003 111.000000 5.000000 649 555
* Completed load of species file #2, tutorial_input/species2.asc

tutorial_input/species3.asc:  555 rows and 649 columns read in total; sum of non-missing data elements = 35307.7
129.800003 111.000000 5.000000 649 555
* Completed load of species file #3, tutorial_input/species3.asc

tutorial_input/species4.asc:  555 rows and 649 columns read in total; sum of non-missing data elements = 35116.9
129.800003 111.000000 5.000000 649 555
* Completed load of species file #4, tutorial_input/species4.asc

tutorial_input/species5.asc:  555 rows and 649 columns read in total; sum of non-missing data elements = 22852.8
129.800003 111.000000 5.000000 649 555
* Completed load of species file #5, tutorial_input/species5.asc
```

Example of the Text output file.

4.3.3 Runtime plot

Curves visualizing the proportion of distributions remaining, average extinction risk, cost needed to achieve a given conservation value, and proportions of distributions remaining for SSI features are plotted in the **Runtime plot** tab of the **Visual Output** window for each run. The lowest fraction across biodiversity features is plotted with a red line, whereas a blue line represents the average across all features. The curves from different analyses active in the

GUI can be viewed by double-clicking on a completed analysis in the **Process View.**

Four types of curves are plotted during each run:

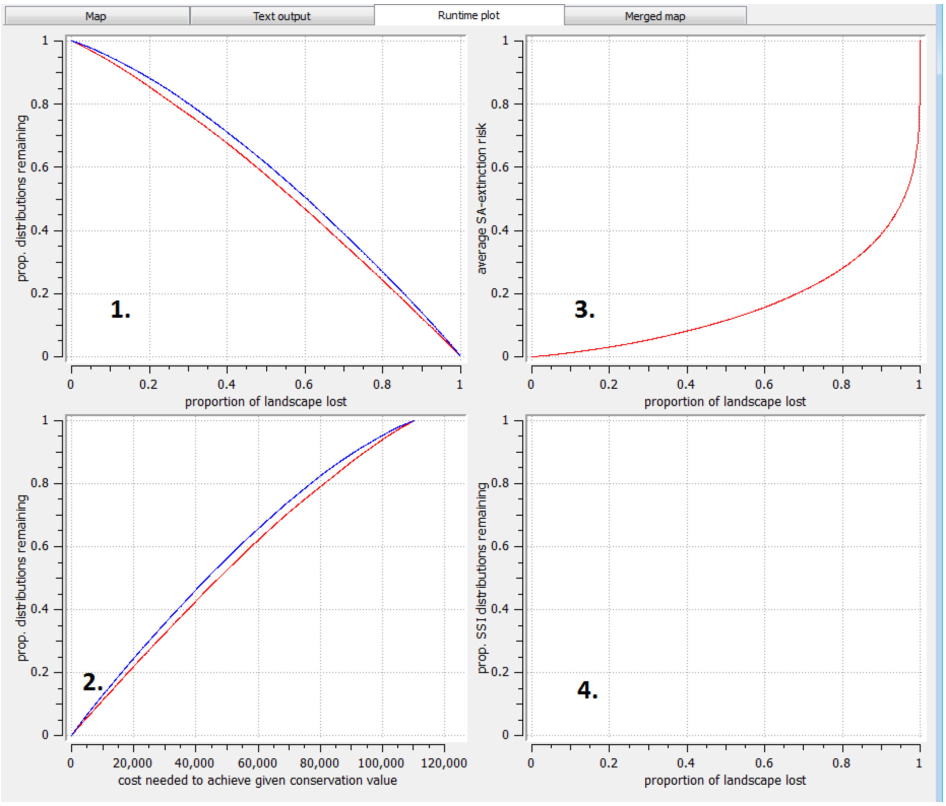
1. The first of the four plots shows summary information about the **proportion of distribution remaining** across features as landscape is removed. The red line represents the biodiversity feature with the lowest distribution remaining while the blue line represents the average over all features. The black line represents the weighted average over all the biodiversity features (using the feature-specific weights). This line may not be visible in some cases. For example, if the weights are equal for all features, the blue line will always appear on top of the black line. All this information can be found in numerical form in the **.curves.txt** file (section 3.4.1), which also contains the respective curves for each feature included in the analysis.

2. Below the distribution curve is the **cost curve** showing what costs are needed for buying the respective top fraction. If no cost layer is used, all cells receive an equal value of 1.0 and the curves show the number of cells needed for respective top fractions. The curves begin from the proportion of landscape that is left after the initial removal of the worst proportion. If nothing is removed, the curve starts from zero, as in the picture below.

3. The third curve shows how the **extinction risk** of biodiversity features increases as landscape is removed. This curve is based on the species-area ratio and shows the average extinction risk over all features, assuming the exponent *z* given in the settings.

4. The fourth and last curve displays the **proportion of distribution remaining for SSI features** (Species of Special Interest) when landscape is removed. As in the proportion of distribution remaining curve, the red line represents the features with the lowest distribution remaining and the blue line represents the average over all features. Note that, unlike the other panels, this panel will only be displayed when SSI features are included in the analysis.

The information of these four graphs is equal to the .curves.txt file that the program automatically produces as part of file output.



Example of graphs shown in the Runtime plot.

4.3.4 Merged map

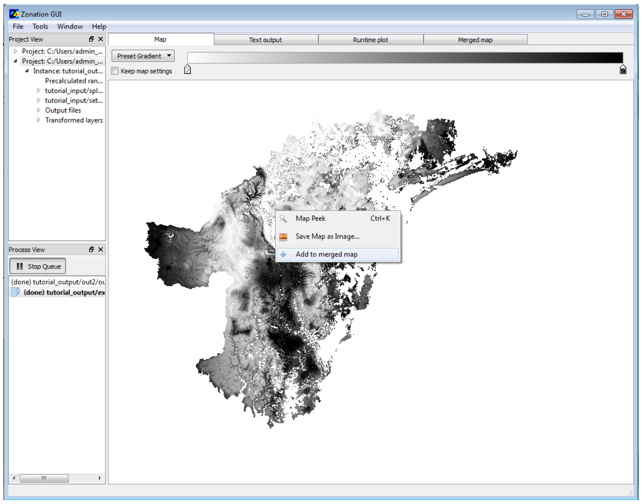
The **Merged map** tab in the **Visual output** window of the graphical interface can be used to visualize the overlap between different maps. The first map added to the merged map is an exact copy of the original and the starting point. From that point on, the merged map is updated every time a new map is added. When a map is added to the merged map, the color gradient defined in the map output (see section 4.3.1) is used. The merging is done by calculating the per-pixel average color for all the maps added.

This way, by selecting an adequate set of colors, it is possible to visualize the overlap between top and/or bottom fractions of different rank maps. Or, more generally, it is possible to visualize the overlaps between different segments in the range of values of the maps aggregated. In practice a reduced number of colors should be applied to gradients of the maps, otherwise the number of averaged (or combined) colors in the merged map can grow too large for a visual analysis. Please note that this merging process is not restricted to output rank maps, and in principle any input or output map can be added, including biodiversity feature distributions, administrative units layers, etc.

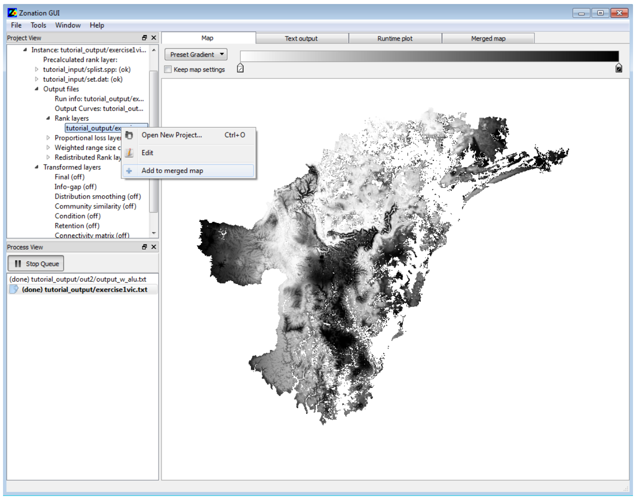
Note also the following difference between the display in the map and merged maps tabs. In the map tab the displayed object is a raster map, with values transformed into colors by applying the color gradient. In contrast, the merged map tab displays an image where each pixel's color is an average of the corresponding pixel colors in the original maps. In other words, the merging or averaging of maps is done in the color space. Merged maps are not meant for quantitative analysis as all numerical data associated with cells is lost when a map is added to the merged map tab. The output is meant only for visual use. If you need a numerical comparison of solutions, you should use the solution loading feature (section 3.5.2).

Maps can be added from the map tab and from the project and process views.

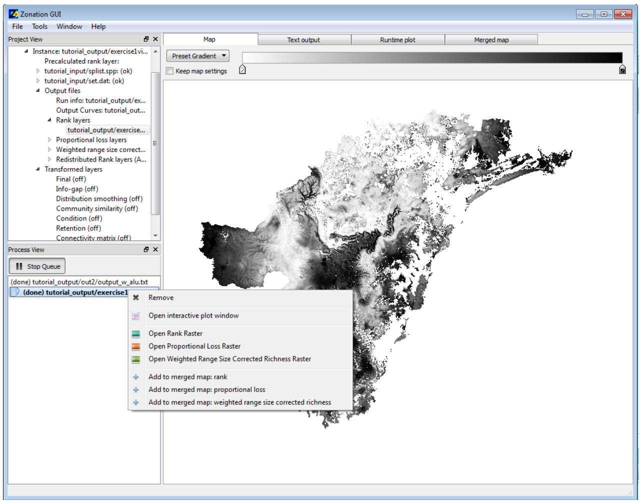
- **Map visual output**: Switch to the "Map" tab. If a map is being displayed, it can be added to the merged map by right-clicking and selecting "**Add to merged map**".



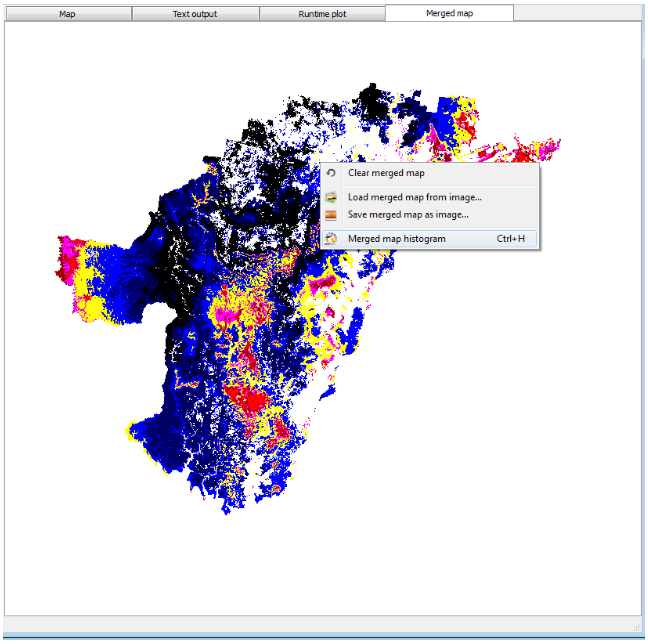
- **Project View**: In the project view, open the tree of a project and right-click on a raster file. Then, select "**Add to merged map**".



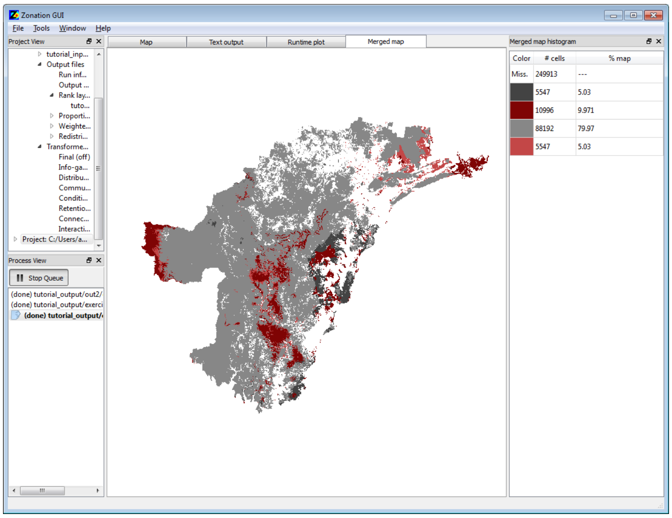
- **Process View**: Select one of the three "**Add to merged map**" options provided in the pop-up menu (only for processes already finished). The three options correspond to the following maps: normal rank, proportional loss, and weighted range size corrected richness.



Right-clicking in the Merged map window will bring up another menu with options to clear the merged map (start from scratch), load an image (which will be used as the starting point for merging maps), and save the current merged map as an image. An additional option opens a window containing the merged map histogram.



Merged map histogram: This window shows a table of colors present in the merged map. The table shows the number of cells with each color and the corresponding percentages of the map (with respect to the total number of non-missing cells). The maximum number of colors allowed is 21.



```
(done) tutorial_output/out2/
(done) tutorial_output/exerci
(done) tutorial_output/
```

Color	# cells	% map
Miss.	249913	---
■	5547	5.03
■	10996	9.971
■	88192	79.97
■	5547	5.03

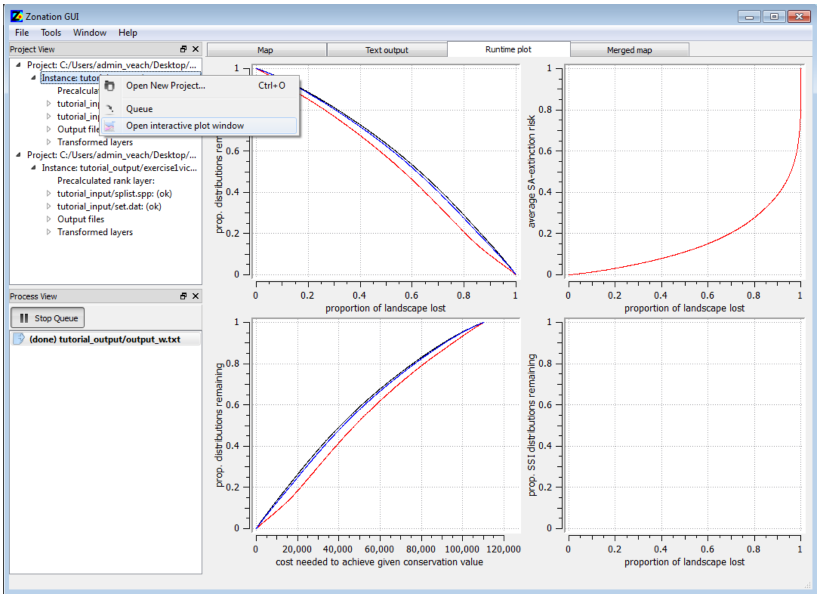
Please note the following clarifications:

- The merging process operates on maps with the same grid size. If, after merging a sequence of maps of a certain size, a new map of a different size is added to the merged map, it will be used as a new starting point.
- Every time a map is added to the merged map, the areas missing data in that map also become areas missing data in the updated merged map. Thus, if a map without any overlap of non-missing areas with the current merged map is added, the result will be an empty map.

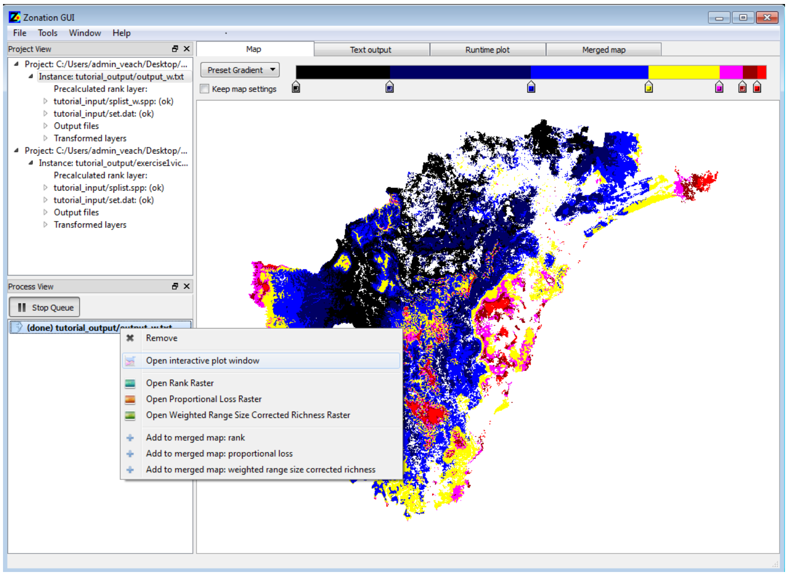
4.4 Interactive plots

Zonation v. 4 offers six interactive plot windows. With the interactive plot window it is possible to quickly generate plots and graphs about global results, results specific to biodiversity features or groups of features, and other plots that can be useful in understanding Zonation outputs. The plots can be copied and/or saved into files. The interactive plots window can be opened from the project and process views.

- **Project View**: In the project view, open the tree of a project and right-click on an instance. If there are results available for the selected instance, the interactive plot window can be opened by selecting "**Open interactive plot window**."



- **Process View**: To open the interactive plot window from the process view, right-click on a finished process and select "**Open interactive plot window**."



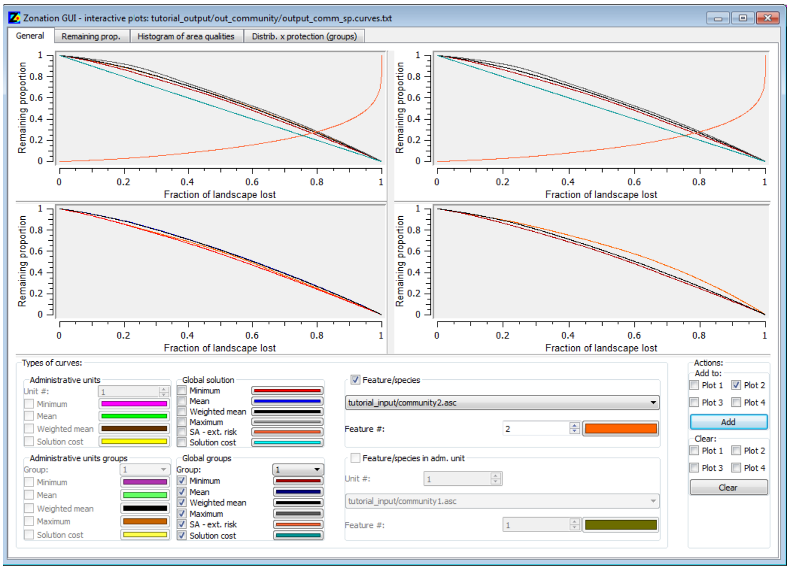
The interactive plots window contains six tabs (see images below): General plots, Remaining proportions, Interval plots, Box plots, Histogram of area qualities, and Distribution x protection (groups).

Note

In the general plots, negatively weighted features (if you are using negative weights) are excluded from the average, minimum, maximum and weighted average curves (this is also true in the Zonation output text files)

General Plots

This tab allows for plotting results relative to individual features, groups of features, administrative units, and groups of features within administrative units.


Zonation GUI - interactive plots: tutorial_output/out_community/output_comm_sp.curves.txt

Types of curves:

Administrative units
Unit #: 1
☐ Minimum
☐ Mean
☐ Weighted mean
☐ Solution cost

Global solution
☐ Minimum
☐ Mean
☐ Weighted mean
☐ Maximum
☐ SA - ext. risk
☐ Solution cost

☑ Feature/species
tutorial_input/community2.asc
Feature #: 2

Administrative units groups
Group: 1
☐ Minimum
☐ Mean
☐ Weighted mean
☐ Maximum
☐ Solution cost

Global groups
Group: 1
☑ Minimum
☑ Mean
☑ Weighted mean
☑ Maximum
☑ SA - ext. risk
☑ Solution cost

☐ Feature/species in adm. unit
Unit #: 1
tutorial_input/community1.asc
Feature #: 1

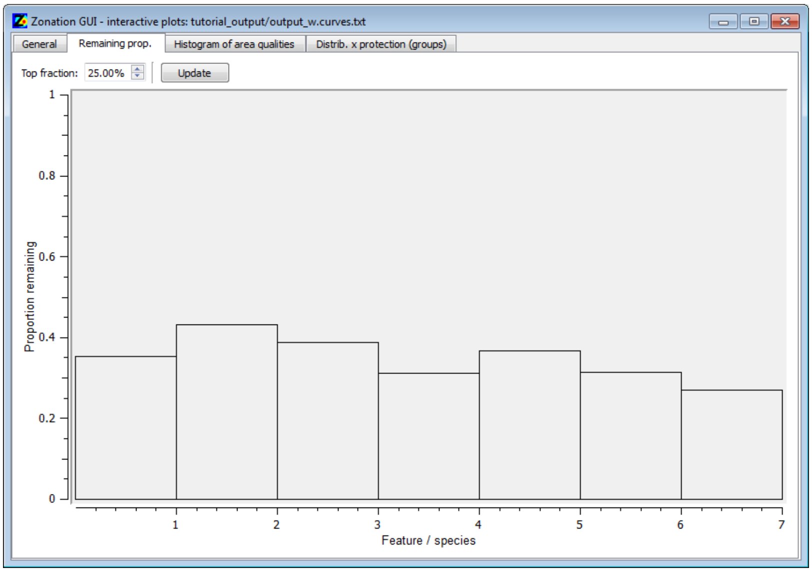
Actions:
Add to:
☐ Plot 1 ☑ Plot 2
☐ Plot 3 ☐ Plot 4

Clear:
☐ Plot 1 ☐ Plot 2
☐ Plot 3 ☐ Plot 4

Example of the interactive plots window 'General plots' tab.

Remaining Proportions

This tab shows the total proportion remaining for every feature/species for a given top fraction of the landscape. This information can be found in the rows of the .curves.txt file.

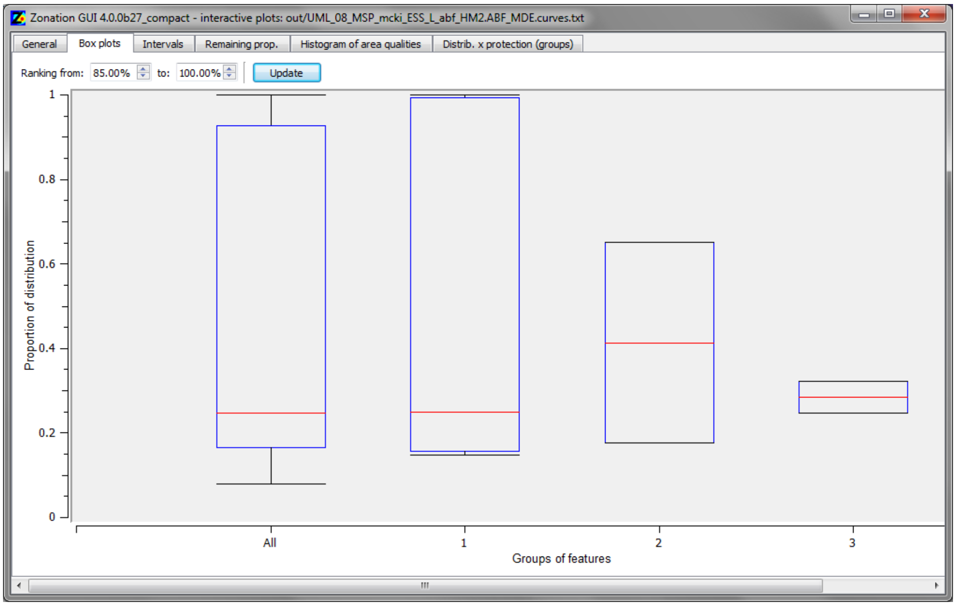


Top fraction: 25.00%

Example of the interactive plots window 'Remaining proportion' tab.

Box Plots

The box plots show the variability of results within groups of biodiversity features. For each group, the bottom of the box corresponds to the first quartile (25th percentile), the line inside the box (in red) shows the median, and the top of the box corresponds to the third quartile (75th percentile). The ends of the whiskers correspond to the maximum and minimum of the groups. In this plot, the first group corresponds to all the biodiversity features. The other groups and the list of features that are inside each group are the ones defined as output groups in the groups file of Zonation (see section 3.3.3.12 Groups file). If you do not use the groups file there will be a single box plot in this tab (including all features). Different groups may correspond for example to different taxonomic groups or otherwise different types of biodiversity features, such as habitats, species, etc. As with other interactive plots, it is possible to select a top fraction of the ranking, or a range of the ranking, for example all the areas of the ranking with level from 83% to 90%.



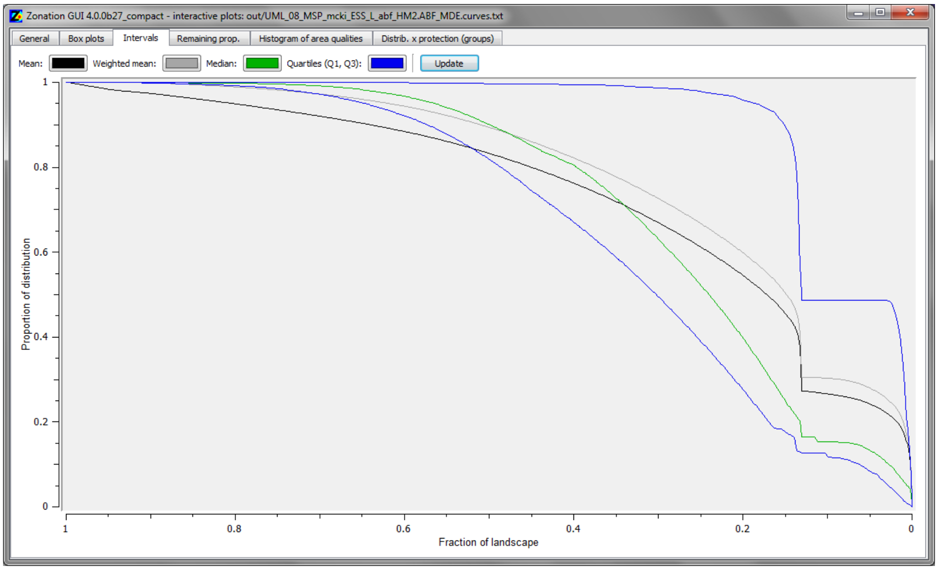
Zonation GUI 4.0.0b27_compact - interactive plots: out/UML_08_MSP_mcki_ESS_L_abf_HM2.ABF_MDE.curves.txt

Ranking from: 85.00% to: 100.00%

Example box plots.

Intervals

This tab displays, for the global results, the mean and weighted mean across biodiversity features (as in the general plots tab), and also the median and first and third quartiles (25th and 75th percentiles, respectively). This plot is a simple way to visualize to what extent protection levels are uniform or balanced across different features.

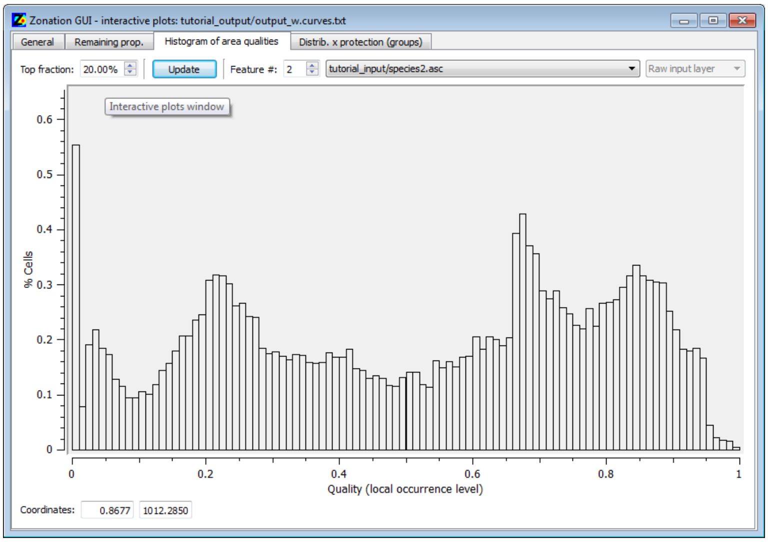


Zonation GUI 4.0.0b27_compact - interactive plots: out/UML_08_MSP_mcki_ESS_L_abf_HM2.ABF_MDE.curves.txt

Example plot of intervals.

Histogram of Area Qualities

This tab shows a histogram of cell qualities (scaled in [0,1]) for a given top fraction of the landscape for a particular feature/species.



Top fraction: 20.00%

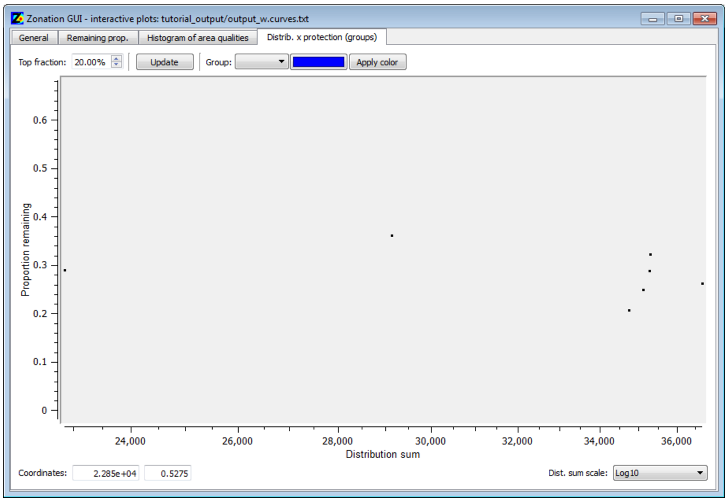
Feature #: 2 tutorial_input/species2.asc

Coordinates: 0.8677 1012.2850

Example showing the interactive plots window 'Histogram of area qualities' tab.

Proportion remaining by proportion protected (Distribution x protection (groups)

This tab displays a scatter plot of features for a given top fraction. The axes are the proportion remaining and the distribution sum fields in the .curves.txt. In this tab, specific groups of features (if groups have been defined) can be highlighted by applying a color.



Top fraction: 20.00%

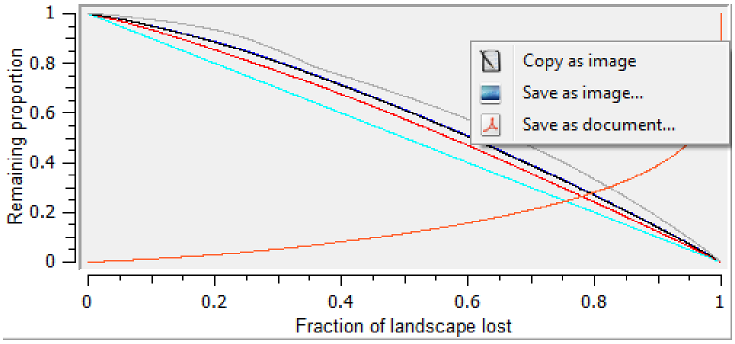
Group:

Coordinates: 2.285e+04 0.5275

Dist. sum scale: Log10

Example showing the interactive plots window 'Distribution x protection (groups)' tab.

In all the plots shown above, it is possible to zoom in/out using the left/right mouse button. Plot coordinates are shown in specific fields (location depending on the tab) and can also be obtained by clicking on the plots. It is also possible to copy these plots to the clipboard or save them as image and/or document by right clicking on them, as shown in the image below.

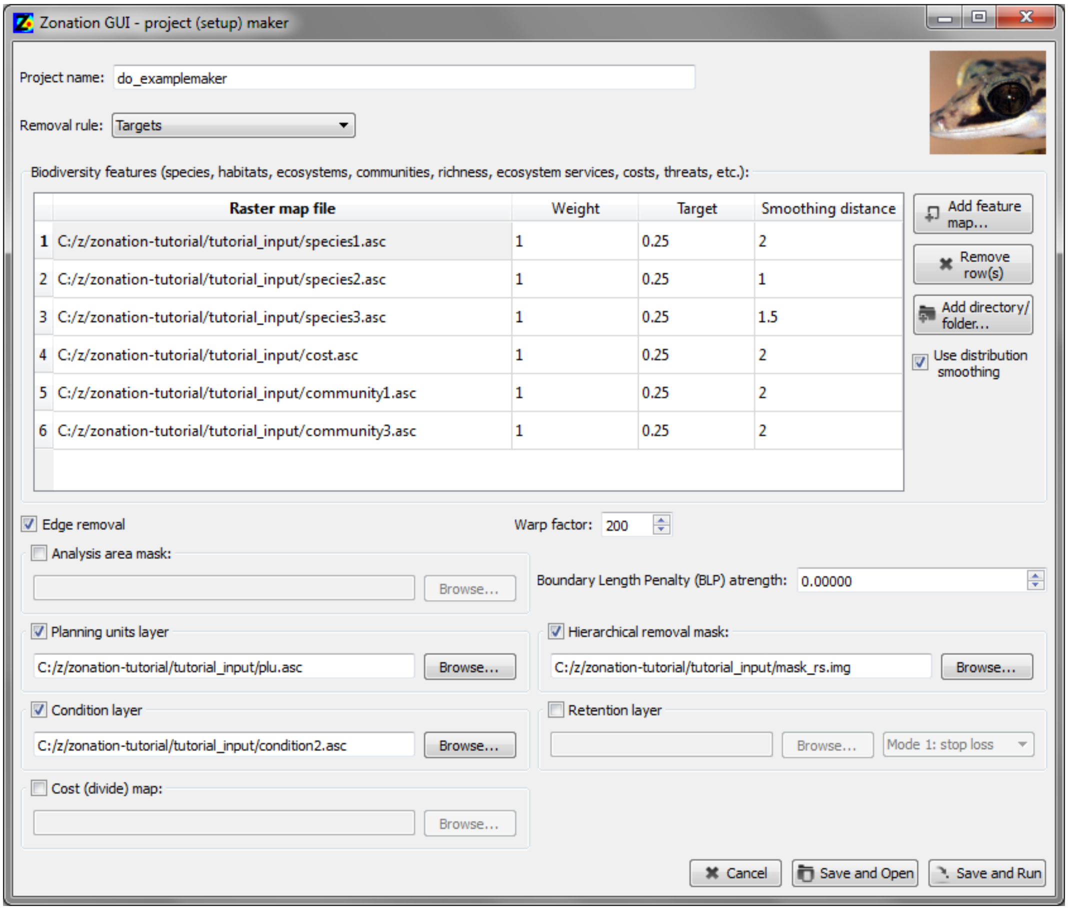


Example showing the options available for each plot by right-clicking on the plot.

4.5 Project maker

Zonation 4 introduces the project maker, a simple graphical interface for creating Zonation setups (projects) from scratch. With the project maker you can create new projects by filling in a simple form. From this form you can directly save and run newly created projects. We recommend using this interface especially if you are a new user of Zonation. With the project maker you can create projects with little or no knowledge about the structure of the Zonation input files, by simply selecting a list of biodiversity feature raster maps, setting a few parameters, and adding other layers such as masks or planning units.

You can start the project maker from the main menu, selecting "**Project",** and then "**Project maker"**, with the "**Ctrl-M"** keyboard shortcut or, alternatively by right-click on the project view and selecting "**Project maker...**". It should open the following window:



Zonation GUI - project (setup) maker

Project name: do_examplemaker

Removal rule: Targets

Biodiversity features (species, habitats, ecosystems, communities, richness, ecosystem services, costs, threats, etc.):

	Raster map file	Weight	Target	Smoothing distance
1	C:/z/zonation-tutorial/tutorial_input/species1.asc	1	0.25	2
2	C:/z/zonation-tutorial/tutorial_input/species2.asc	1	0.25	1
3	C:/z/zonation-tutorial/tutorial_input/species3.asc	1	0.25	1.5
4	C:/z/zonation-tutorial/tutorial_input/cost.asc	1	0.25	2
5	C:/z/zonation-tutorial/tutorial_input/community1.asc	1	0.25	2
6	C:/z/zonation-tutorial/tutorial_input/community3.asc	1	0.25	2

☑ Use distribution smoothing

☑ Edge removal

Warp factor: 200

☐ Analysis area mask:

Boundary Length Penalty (BLP) strength: 0.00000

☑ Planning units layer: C:/z/zonation-tutorial/tutorial_input/plu.asc

☑ Hierarchical removal mask: C:/z/zonation-tutorial/tutorial_input/mask_rs.img

☑ Condition layer: C:/z/zonation-tutorial/tutorial_input/condition2.asc

☐ Retention layer: Mode 1: stop loss

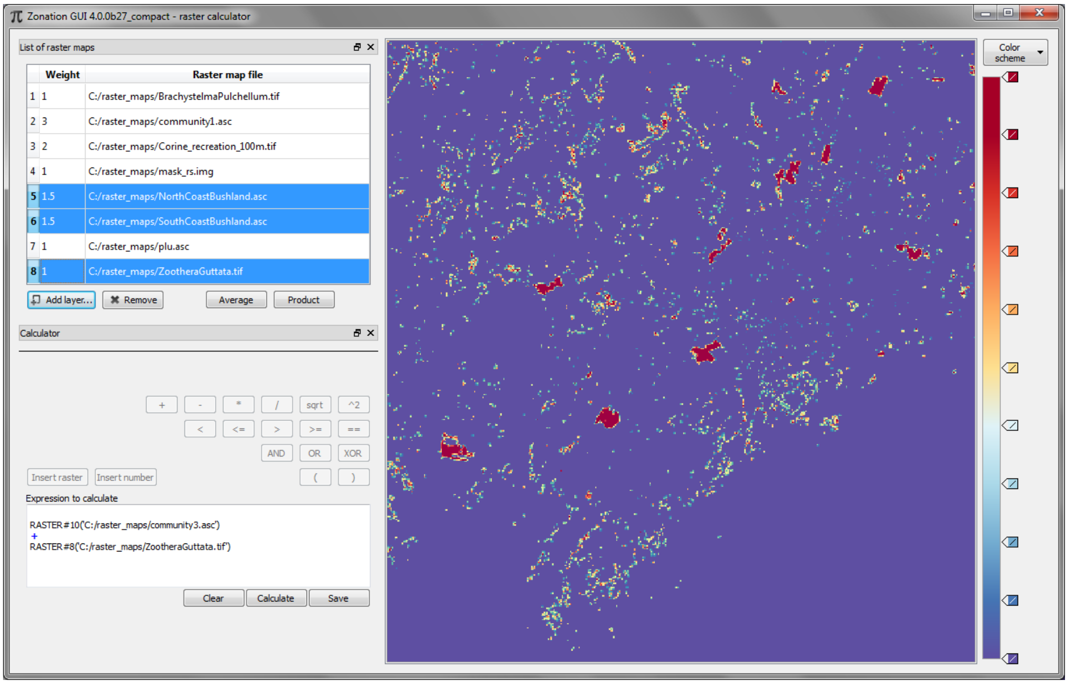
☐ Cost (divide) map:

The project maker form supports most if not all the commonly used analysis features of Zonation. This includes 4 different ranking (cell removal) rules: Core area zonation, additive benefit function, targets, and random, the BLP and distribution smoothing aggregation methods, the analysis area mask, hierarchical removal mask, planning units layer, condition layer, retention layer, cost layer, etc. The weights of different features and other parameters required for different rules can also be set from the project maker.

4.6 Raster calculator

Zonation 4 includes a new tool called raster calculator. It is not intended as a fully featured calculator, such as those found in GIS software, but it should suffice to perform basic operations on raster maps in order to prepare input maps for Zonation analyses, or to perform simple post-processing of output maps. This can include for example preparation of combinations of maps, application of masks or thresholds, calculation of average ranking maps for multiple ranking outputs, etc. The calculator makes sure that the size of all the raster maps is the same.

The **raster calculator** window (you can open it from **Tools** -> **Raster calculator**) consists of three parts. Normally, on the right there is a sub-window that displays the raster maps being manipulated, using a color bar similar to the one used in the main window of the Zonation GUI. On the left there are normally two sub-windows: the list of raster maps, and the calculator.



	Weight	Raster map file
1	1	C:/raster_maps/BrachystelmaPulchellum.tif
2	3	C:/raster_maps/community1.asc
3	2	C:/raster_maps/Corine_recreation_100m.tif
4	1	C:/raster_maps/mask_rs.img
5	1.5	C:/raster_maps/NorthCoastBushland.asc
6	1.5	C:/raster_maps/SouthCoastBushland.asc
7	1	C:/raster_maps/plu.asc
8	1	C:/raster_maps/ZootheraGuttata.tif

```
RASTER#10('C:/raster_maps/community3.asc')
+
RASTER#8('C:/raster_maps/ZootheraGuttata.tif')
```

In the list of raster maps you can load any raster map and give different weights. By default the weights are all set to 1. From the list of raster maps it is possible to calculate weighted average maps, or products of maps. This can be done by simply selecting several maps and using the buttons "average" or "product" in the rater map sub-window (top-left). In the calculator (normally bottom-left window), it is possible to perform different arithmetic and logic operations on raster maps. This is done by first inserting a number or a raster map (selected from the list) into the equation or expression to calculate, then selecting an operator, and then inserting one number or raster map.

Note that the way raster maps are added in the raster calculator differs from the merged map feature of the Zonation GUI. In the **Merged map** tab of the Visual output of the Zonation GUI you can combine maps and visualize overlaps between them, with an average of maps that is calculated by colors. In the raster calculator, average and other operations are calculated on the values of the raster maps, independently of the colors used to display them. If you want to inspect the values of the pixels of a map generated with the raster calculator you can do so by loading it as a raster in the **Map View**, and using the "**Map peek**" and other facilities of the main Zonation visual output.

Part

V

5 Zonation analysis setups for common planning needs

In the following sections we describe some analysis setups for Zonation that correspond to typical needs of managers and researchers. The idea is to facilitate the use of Zonation for practical needs, so this chapter identifies and links to relevant parts of the documentation. However, we recommend getting acquainted with the theoretical basis of the methods (section 2) and examples of analysis implementations from the literature. Some, but not all, setups have corresponding tutorials with example files.

5.1 Basic analysis components

This section describes the use of basic components that are often included in a Zonation prioritization. These components can be applied to any analysis, but please remember that including several connectivity considerations in a single analysis is not recommended as it can complicate interpretation of the output.

5.1.1 Simple Zonation and biodiversity feature weighting

Planning problem to be solved

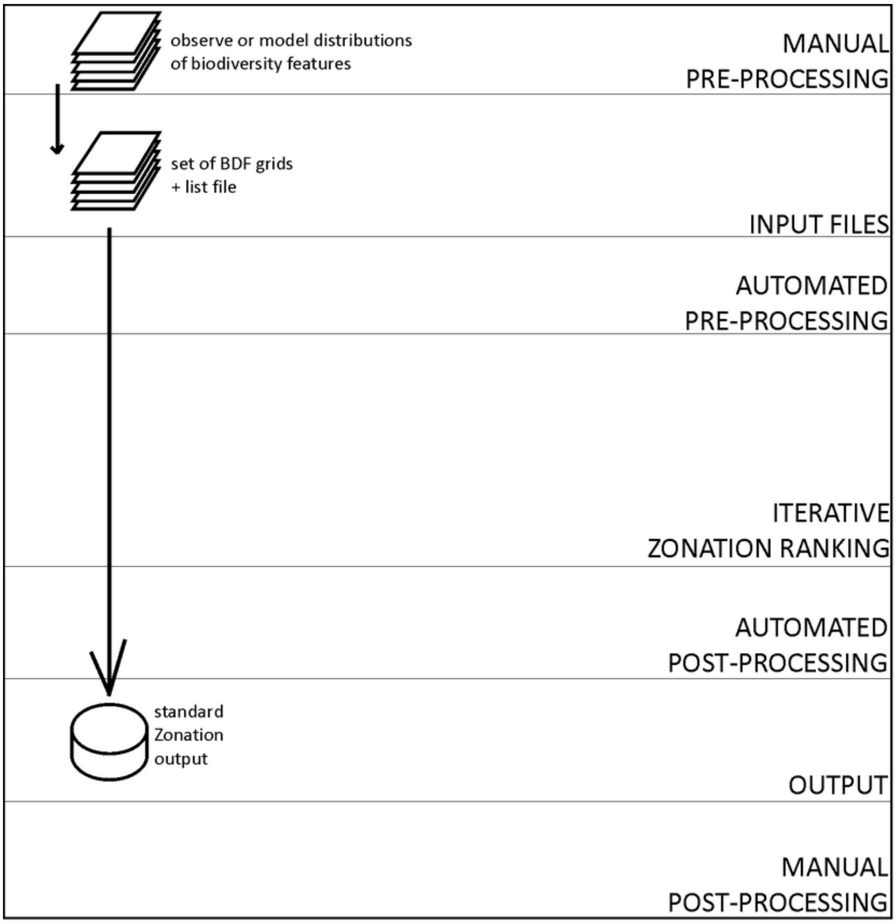
Simple Zonation is frequently the first analysis run for any planning situation. In most cases, this analysis is not intended to be a final product used in planning. Rather, this basic Zonation analysis can be used to ensure all input files function properly and to get a general idea about what the output might look like. Simple Zonation can also be useful as a reference analysis. For example, this solution can be compared to another that includes connectivity considerations or other more complex settings. Representation curves of the solutions can be compared to assess, for example, how much local quality is lost when connectivity is increased. As this analysis is simple and quick to run, it is often worthwhile to try different cell removal rules to see how they affect the outcome. Note that different assumptions about relative richness and rarity are implemented when conservation value is aggregated depending on whether you choose the core-area Zonation, the additive benefit function, or the generalized benefit function as your cell removal rule. Computing the analysis using both core-area Zonation and the additive benefit function can help you find out whether there are areas that support narrow-range species or other biodiversity features in an area with otherwise low richness in biodiversity features.

Examples from literature

Moilanen, A., and B.A. Wintle. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

The theory and algorithms behind basic Zonation are explained in sections 2.2 and 2.3.

Process chart for the analysis



A process chart for a basic Zonation analysis.

Pre-processing of inputs

Processing of inputs often includes fitting habitat suitability models to existing species data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

Input files

To run simple Zonation, you need:

- A set of biodiversity feature grid layers (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
- A run settings file with appropriate settings (section 3.3.2.3)

Analysis stages and settings

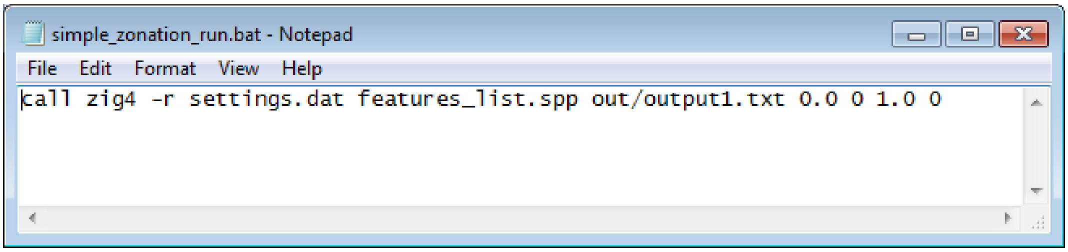
The basic setup for running Zonation is described below. The skeleton of the simple setup can and should be complemented with spatial and ecological considerations to add realism to the analysis. More complex setups follow the same basic scheme and are described in the next sections.

To set up a basic Zonation prioritization run, you need to:

1. Set all options of **additional analyses** (e.g., BQP, uncertainty analysis, etc.) to zero in your **run settings file** AND in the **command line call** to indicate that no additional analyses are used.
2. Adjust your settings in the run settings file for the following options if necessary:
    - cell removal rule
    - warp factor
    - edge removal
    - add edge points
    - logit space

    Note that if you do NOT select the edge removal, the computation times will increase significantly with large data sets.
3. Adjust the feature weights in the biodiversity feature list file if you wish to stress the conservation of certain features (e.g., rare species of high conservation value or commercially valuable species).
4. Type the call for Zonation in the command prompt and press enter to initiate the computation.

See section 3.2.1 for how to use batch files to call Zonation.



```
call zig4 -r settings.dat features_list.spp out/output1.txt 0.0 0 1.0 0
```

No special settings are needed for this analysis. Weighting of biodiversity features and choice of cell removal rule can affect the outcome.

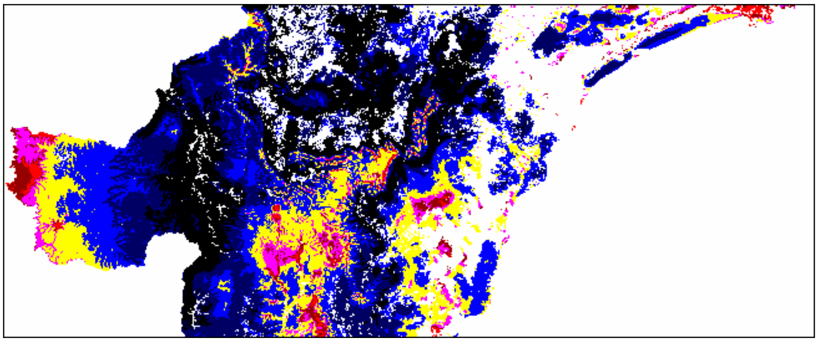
Weighting of biodiversity features is a critical component of the Zonation algorithm. Problems associated with different initial sizes of feature distributions are circumvented in Zonation by assigning a value for the full distribution of each feature. By default, these values are equal, but features can be assigned different weights in the biodiversity feature list file based on, for example, their taxonomic status, global rarity, economical value, or population trend. Weighting of features affects the order in which cells are removed from the landscape and influences the fraction of a feature's distribution retained at any point during the cell removal process. Cells that contain a valuable feature (high weight) are retained longer in the iterative cell removal process than cells containing only low-weight features, assuming everything else is equal between the occurrences. Features with high weight also retain a relatively higher proportion of their distribution than features with low weight. Also, note that the balance in representation levels developed by Zonation is such that narrow-range features typically have a larger fraction of their ranges protected compared to initially wide-ranging features.

Weights can also be used to test the efficiency of **surrogate features**. This is done by weighting the surrogate features normally (e.g., by 1) and giving a weight of 0 to all features that are NOT used as surrogates. By giving a weight of 0 to the non-surrogate features, Zonation will not use them in cell removal calculations but will still monitor the decrease in the distributions of these species. This can show how well a reserve network selected using surrogates will protect all features.

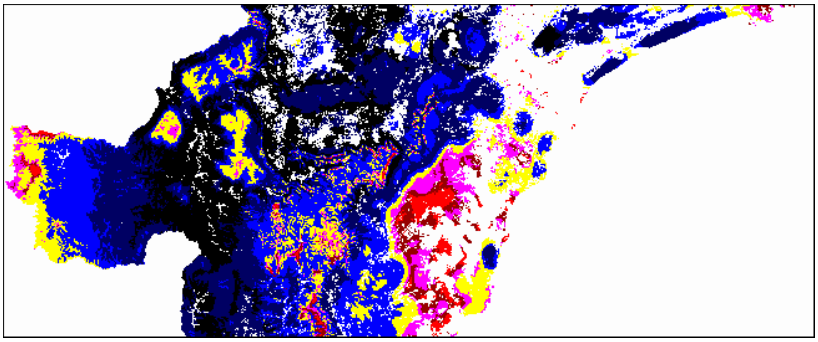
Output and its interpretation

The priority ranking map and representation curves of a basic Zonation analysis can be used as references for more complicated analyses.

Below is an example of how feature weighting can influence the final solution. The pictures show the results of two basic Zonation runs for seven species. In the first picture no species weights have been used whereas in the second picture, one of the species has been given a weight of 3.0. In the weighted map, the areas of high prioritization (pink and red areas) have shifted from the western and central region to the eastern region where the weighted species has a higher occurrence.



Example of a priority ranking map where all species receive the same weight.



Example of a priority ranking map where one species was assigned a weight of 3. The species in this map are the same as in the previous one.

Strengths, weaknesses, and further considerations

Strength: Simple Zonation is simple and quick to set up and run.

Weakness: The analysis is too simple to be realistic. Considerations of connectivity, costs, landscape condition and retention, and other such factors are missing even though they can have a significant influence on the conservation outcome at the implementation stage.

Link to tutorial

See Exercise 1 for a tutorial example of the basic use of Zonation. Exercise 2 provides an example of species weighting.

5.1.2 BLP Zonation

Planning problem to be solved

Boundary length penalty (BLP) is used to account for connectivity in conservation planning. BLP is a simple method to induce aggregation in the protected area network and is based on the assumption that increasing the total length of boundary will decrease the conservation value of the remaining cells. The algorithm minimizes the total boundary length of the reserve network by assigning a penalty to the boundary length and prioritizing cells that are aggregated.

Examples from literature

Moilanen, A., and B.A. Wintle. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

The theory and algorithm behind boundary length penalty is explained in section 2.4.1.

Input files

No special input files are needed in addition to the normal biodiversity feature layer grid sets and associated list file. The detailed combination of input files depends on the specific aims and components of your analysis.

Analysis stages and settings

To include BLP in the analysis, assign a value for the penalty of increased boundary length. To do this, type the following line in your run settings file (section 3.3.2.3):

```
BLP = [value]
```

The value of BLP should be a small decimal number. Try a small value first (e.g., 0.01) to perceive the effect of BLP on the solution. It is advisable to run the analysis with multiple BLP values to estimate the sensitivity of the solution to the value chosen.

When including BLP in the analysis, always use a warp factor of 1.

Strengths, weaknesses, and further considerations

Strength: Computationally quick and effective.

Weakness: BLP is not a feature-specific aggregation method and does not account for the actual effects of fragmentation on biodiversity features. Therefore, the results from a BLP analysis may not be the most biologically realistic.

Note potential difficulties in interpretation of results if multiple aggregation methods are used simultaneously.

Link to tutorial

See tutorial Exercise 4.

5.1.3 Distribution smoothing + Zonation

Planning problem to be solved

Distribution smoothing is used to account for connectivity of the protected area network from the perspective of feature-specific dispersal ability. Distribution smoothing is a two-dimensional kernel smoothing, where the width of the smoothing kernel is determined by the estimated dispersal ability or scale of landscape use of the feature in question. This option results in a much more compact solution where small isolated patches have been removed. Using distribution smoothing does not significantly increases computation times.

Examples from literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and C.D. Thomas. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

Moilanen, A. and B.A. Wintle. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation*, 129: 427-434.

The theory and algorithm behind distribution smoothing is explained in section 2.4.2.

Pre-processing of inputs

To include distribution smoothing in an analysis, determine the width of the dispersal kernel (a value for the distribution smoothing α) for each feature (i.e., to what extent can a species can move in the landscape). Insert the α values in the second column of the biodiversity feature list file.

Input files

- A set of biodiversity feature grid layers (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2). Enter the widths of the dispersal kernels for each feature in the second column of this file.
- A run settings file (section 3.3.2.3) with appropriate settings

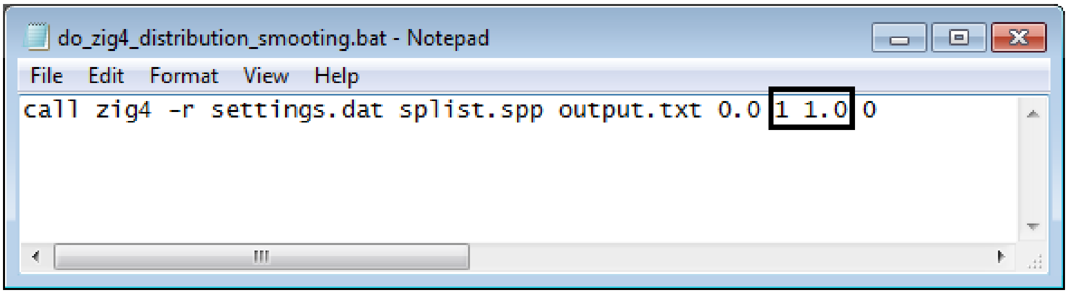
Distribution smoothing can be included in any analysis to induce aggregation in the protected area network. The detailed combination of input files depends on the specific aims and components of the analysis.

Analysis stages and settings

The parameters for distribution smoothing are entered in the command line (batch file), not the run settings file (see example below).

1. Set the third last parameter of the command line call to 1 to indicate that distribution smoothing will be done.

2. Give a factor for multiplying the **feature-specific α values** in the second last parameter in command line. α values (feature-specific kernel widths) are in the second column of the biodiversity feature list file. This factor is useful if you are interested in running multiple solutions that assume several levels of dispersal capabilities because it allows you to multiply all dispersal capabilities simultaneously. Thus you do not need to change the parameters manually in the biodiversity feature list file after each run. If you do not wish to multiply the α values, set this factor to 1.

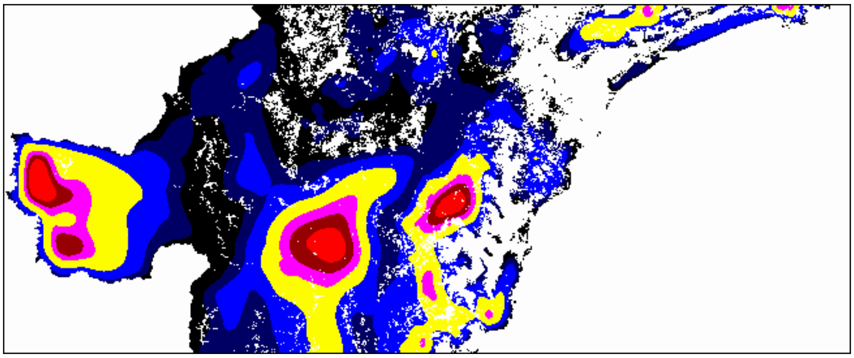


```
call zig4 -r settings.dat splist.spp output.txt 0.0 1 1.0 0
```

Example of a batch file for an analysis using distribution smoothing. Highlighted in the black box are the parameters that must be modified in order to use distribution smoothing.

Output and its interpretation

Distribution smoothing should result in a distinctively more aggregated solution compared to that of a basic Zonation analysis, and the two solutions should not have a large amount of overlap with each other. If your solution with distribution smoothing does not show clear aggregation, check the run settings file for possible errors (e.g., you could have α in different units than the cell size in the biodiversity feature distribution files).



Example of a typical output map when distribution smoothing has been included in the analysis.

Strengths, weaknesses, and further considerations

Strength: Distribution smoothing adds realism to the analysis as it turns feature-specific dispersal ability into a factor in the conservation planning process. Multiple levels of connectivity, such as within and between populations of a single species, can be accounted for in a single Zonation run (see section 5.3.3).

Weakness: Distribution smoothing always favors uniform areas over patchy ones, which could create problems for species that prefer a fragmented landscape.

Link to tutorial

See tutorial exercise 4.

5.1.4 BQP Zonation

Planning problem to be solved

Boundary quality penalty (BQP) is a quantitative method that induces aggregation into reserve networks according to the needs of individual features. It can be used to account for feature-specific effects of habitat fragmentation in conservation planning.

Using BQP decreases the biological quality of a land unit (grid cell) that is located close to the edge of the reserve, resulting in a more highly aggregated optimal reserve structure. As the change in local value is based on feature-specific responses to neighborhood habitat loss, the biological value of a cell may also increase if a site includes a feature that benefits from fragmentation.

Examples from literature

Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

The theory and algorithm behind BQP is explained in section 2.4.3.

Pre-processing of input files

To include BQP in an analysis, define set of penalty curves that represent feature-specific responses to habitat fragmentation (see section 2.4.3) in a BQP definition file (section 3.3.3.2). Adjust the biodiversity feature list file by entering the row number of the correct BQP curve (in the BQP definition file) for each biodiversity feature in the third column of the biodiversity feature list file. This links the biodiversity features to the BQP curves. Note that multiple features can be linked to the same penalty curve. In the fourth column of the biodiversity feature list file, indicate a suitable buffer size in number of cells for each feature.

Input files

- A BQP definition file (section 3.3.3.2) containing responses to habitat fragmentation as penalty curves
- A set of biodiversity features grid layers (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
- A run settings file with appropriate settings (section 3.3.2.3)

BQP can be included in any analysis to induce aggregation in the protected area network. The detailed combination of input files depends on the specific aims and components of your analysis.

Analysis stages and settings

To run a BQP analysis, type the following lines in the run settings file:

1. `use boundary quality penalty = 1`

2. `BQP profiles file = my_BQP_curves.txt`

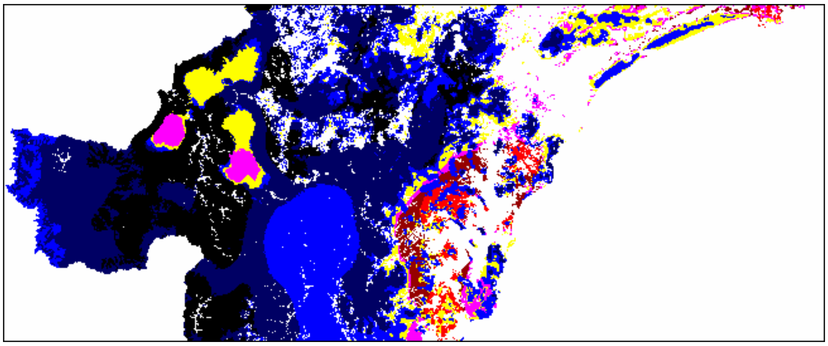
3. `BQP mode = 1` If the data/no data matrix in all biodiversity feature distribution map files are uniform and aligned and if there are no differences between species in terms of which cells are considered potential habitat and which are used in BQP buffer calculations. In other words, all features would be dependent on the same general habitat type, such as forest. OR

   `BQP mode = 2` If the data/no data matrices are not uniform and aligned and the program needs to calculate feature-specific buffers for each feature/cell separately. Mode 2 is more realistic in the sense that fragmentation/loss in habitats that are not suitable for the feature will not influence the value of the focal cell. But, mode 2 also requires longer computation times due to more complicated feature-specific calculations.

Output and its interpretation

Including BQP in your analysis will most likely increase computation times significantly, especially if the feature-specific buffer sizes are large (many cells). It might be wise to reduce the data resolution if computation times become too long. It is recommended to use a moderate to high warp factor for your preliminary runs and a low warp factor only for the final BQP analysis. A BQP analysis should result in a distinctly more aggregated solution compared to a basic Zonation analysis, at least in part of the landscape. Thus, these two solutions should not have a large amount of overlap unless your biodiversity features are not influenced by habitat loss (as defined in the BQP file). If your solution with BQP shows signs of high fragmentation, check the run settings file for possible errors. Note however, that a BQP solution may include fragmented areas if the data contains many features that are indifferent to fragmentation or even favor fragmented habitats.

The use of BQP will produce feature-specific curves that show lower proportions of distribution retained as compared to the basic (non-spatial) analysis. This is because BQP implies that habitat loss and fragmentation will have negative consequences for remaining areas as well. We emphasize that this does not mean that the solution developed by BQP is inferior to the basic analysis, but rather, the basic analysis gives an overly optimistic estimate of how much biodiversity value a fragmented landscape would retain.



Example of a typical output map when BQP has been included in the analysis.

Strengths, weaknesses, and further considerations

Strength: BQP accounts for feature-specific responses to habitat fragmentation, which makes it a rather realistic aggregation method.

Weakness: Individual considerations increase the need for biological data.

Please remember that even though there are no technical hindrances for using several aggregation methods in a single analysis, it may have unexpected effects and cause difficulty interpreting the results. Thus we recommend using only one aggregation method at a time.

Link to tutorial

See tutorial exercise 4.

5.1.5 Directed freshwater connectivity

Planning problem to be solved

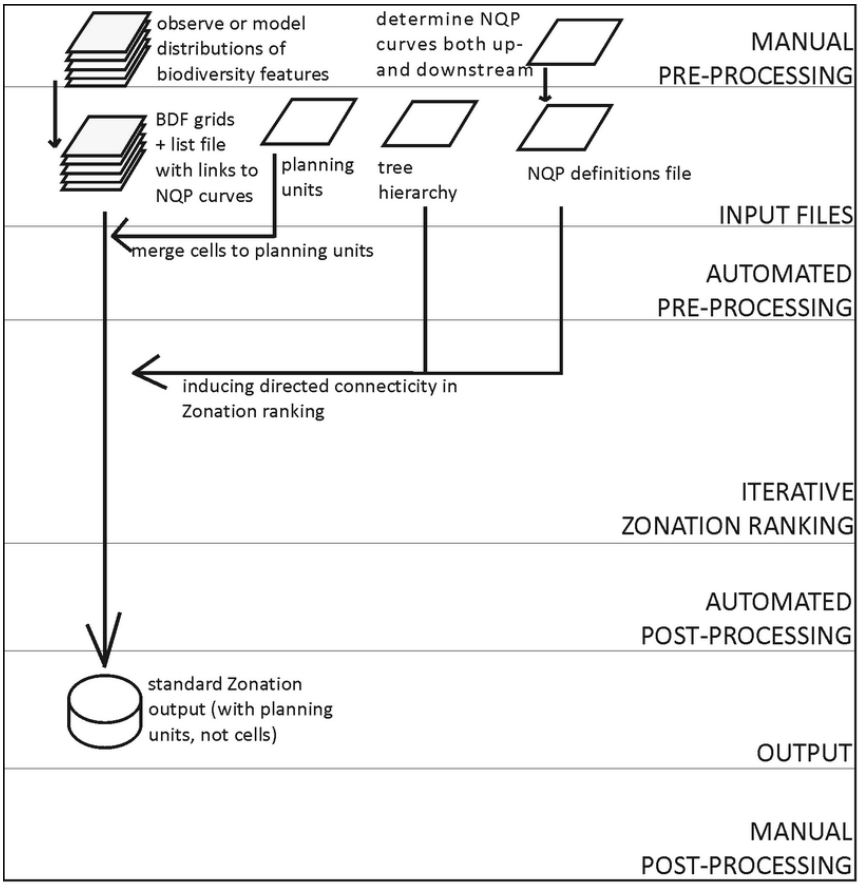
Directed freshwater connectivity is used to identify conservation priorities in freshwater systems where hydrological up- or downriver connectivity needs to be accounted for. Catchments are treated as planning units in this analysis.

Examples from literature

Moilanen, A., Leathwick, J. and J. Elith. 2008. A method for freshwater conservation prioritization. *Freshwater Biology*, 53: 577-592.

Leathwick, J.R., Moilanen, A., Ferrier, S. and K. Julian 2010. Community-level conservation prioritization and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

Process chart for the analysis



A process chart of an analysis with directed freshwater connectivity.
Please note that only the compulsory and most commonly used optional analysis components are presented. You can combine different components according to your specific needs.

Pre-processing of inputs

To induce directed freshwater connectivity in a prioritization, first merge the data into planning units that consist of catchment areas. The hierarchy in terms of direction of water flow also needs to be identified. Finally, quantify feature-specific responses to fragmentation of surrounding habitat and build sets of boundary quality penalty (BQP) curves. One BQP curve should represent responses to fragmentation upstream and one should represent responses downstream from the focal cell where the biodiversity feature occurs.

Input files

- A set of biodiversity feature grid layers (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2); Link all features to the correct penalty curves by entering the correct row number of the respective curve in the third (upstream) and fourth (downstream) columns.
- A planning unit layer, in which catchments are treated as planning units (section 3.3.3.11)
- A tree hierarchy file describing the linkage between planning units (section 3.3.3.3)
- A BQP definition file that describes both upstream and downstream connectivity responses; each feature is assigned two penalty curves instead of one (section 3.3.3.2)
- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, landscape condition and retention layers, etc., can be included.

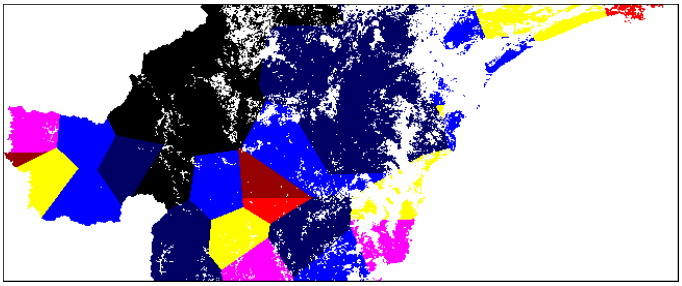
Analysis stages and settings

To run an analysis with directed connectivity, include the following lines in the run settings file:

1. Set "`use planning unit layer`" to 1 in your run settings file and give the name of your planning unit layer file
2. Set "`use tree connectivity`" to 1 to indicate that NQP will be used
3. Define the name of your NQP connectivity file in the run settings file.
4. Note that when planning units are used, the program will automatically set warp factor to 1, regardless what has been defined in the run settings file.

Output and its interpretation

As this analysis directs the connectivity between sites and uses planning units instead of singular grid cells, the output should look significantly different in comparison to a basic analysis. The aggregation of areas depends on the average size of your planning units: large planning units lead to a high level of aggregation but coarser resolution whereas a small planning unit size results in higher resolution solutions.



Example output map when planning units and directed connectivity have been included in the analysis.

Note that the use of large planning units will cause an apparent decrease in the quality of results because large planning units will likely contain areas that are both good and bad for conservation. Consequently, the performance curves will suggest lower protection levels than what can be obtained if selection is based on individual grid cells. However, this effect can be counteracted by the relatively high levels of connectivity gained with the use of planning units. Using planning units will likely shorten computation times with reductions directly proportional to the size of planning units (large planning units = larger reductions in computation times).

Strengths, weaknesses, and further considerations

Strength: In the freshwater context, a relatively realistic analysis is possible with directed connectivity.

Weakness: Developing the connectivity files is relatively complicated. The process has many steps where errors can occur (planning units may end up being linked in a circle, for example). Zonation fixes some problems automatically and raises warnings. Please check the memo after these analyses for any warnings or error notes.

Link to tutorial

See tutorial exercise 8 for an example.

5.1.6 Matrix connectivity

Planning problem to be solved

Matrix connectivity is used to account for the extent to which multiple features influence each other's connectivity. If the biodiversity features are species, this analysis feature can be used to include multiple ecological interactions. If community types are used as biodiversity features, matrix connectivity can account for the extent to which community types resemble each other and facilitate dispersal or other processes that benefit from connectivity.

Examples from literature

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

Pre-processing of inputs

To run an analysis that uses matrix connectivity to induce aggregation in the protected area network, define the pair-wise interactions between the biodiversity features in terms of how they influence each other's connectivity. Connectivity interactions can be limited to certain key features and not all feature pairs need to have connectivity interactions defined. See sections 2.4.5 and 3.3.3.4 and Lehtomäki *et al*. (2009) for details about the application of the matrix.

The features which are included in the connectivity analysis should be listed first in the biodiversity feature list file in the same order that they appear in the matrix. For this analysis, the connectivity requirements are defined by feature-specific dispersal kernels that are used to transform occurrence levels to connectivity. Define widths of the dispersal kernels (dispersal α) for each biodiversity feature and enter them in the second column of the biodiversity feature list file.

Input files

- A set of biodiversity feature grid layers (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
- A connectivity similarity matrix (section 3.3.3.4)
- A connectivity edge effect fix file (optional, section 3.3.3.5)
- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, landscape condition and retention layers, etc., can be included.

Analysis stages and settings

To run Zonation with matrix connectivity, include the following lines in the run settings file:

```
[Community analysis settings]
load similarity matrix = 1
connectivity similarity matrix file = my_connectivity_matrix.txt

apply to connectivity = 1
```

To include edge effect fix in the analysis, include the line:

```
connectivity edge effect fix file = fixfile_name.asc
```

To correct for the amount of habitat contained within cells as well, use a cost layer defining habitat proportions for each cell:

```
use cost = 1
cost file = habitatproportion_name.asc
```

See section 2.4.6 for more information about using edge adjustment and cost in matrix connectivity.

Note that the parameter in the Zonation call multiplying the dispersal a values modifies the distances used in matrix connectivity as well. See section 5.1.3 for more detailed instructions.

Strengths, weaknesses, and further considerations

Weakness: Developing the similarity matrices requires effort.

Link to tutorial

No tutorial example is provided for this analysis.

5.1.7 Cost efficiency analysis

Planning problem to be solved

Cost efficiency analysis is used to maximize cost efficiency in terms of conservation value per cost unit. Cost can be integrated into any analysis by inputting a grid layer with cost information. Cost can refer to actual financial cost of acquiring the area for conservation, or it can represent the opportunity cost of not using the area for production of fish, timber, agricultural goods, etc. In Zonation, conservation value in each grid cell is divided by the cost of protecting that cell so that conservation value per cost unit is maximized in the solution.

Examples from literature

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters* 1: 91-102.

For discussion about cost considerations in conservation planning, see
Arponen, A., Cabeza, M., Eklund, J., Kujala, H. and J. Lehtomäki. 2010. Costs of integrating economics and conservation planning. *Conservation Biology* 24: 1198–1204.

Pre-processing of inputs

It is possible to modify the effect to which cost influences the solution by altering the relative cost differences between cells. An example is provided by Leathwick *et al*. (2008), where a cost layer was first created and then three modified cost layers were computed by introducing *a* parameter a as in:

$$C_i^{\mathrm{modified}} = \begin{cases} a\,, \\ a + 1 + \log_{10}(C_i)\,, \end{cases}$$

with *a* assigned to cells in which conservation would imply no cost, and *a* + 1 + $\log_{10}(C_i)$ to the cells where a cost C_i would follow from conserving them. The larger the value of a, the smaller the differences between cells become, reducing the influence of cost in the solution. Using several values of a can give an idea about how sensitive the conservation solution is to the consideration of cost. In practice, this would mean computing several cost layers and then running the analysis a number of times with a different cost layer each time. The most convenient way to do this is to compile all the different analysis calls into a single batch file (see section 3.2.1).

Input files

- A set of biodiversity feature grids (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
- A cost layer (section 3.3.3.6)
- A run settings file with appropriate settings (section 3.3.2.3)

Other input files depend on the details and other considerations of the analysis.

Analysis stages and settings

Consideration of cost can be integrated into any analysis. To do this, include the following lines in the run settings file:

```
use cost = 1

cost file = my_cost_layer.asc
```

An advisable strategy is to run the same analysis twice: first without and then with the cost layer.

Output and its interpretation

After running the analysis both with and without cost, it is possible to compare the representation curves (section 3.4.1) from each analysis to see how consideration of cost-efficiency affects the biodiversity value of the conservation solution.

Strengths, weaknesses, further considerations

Strength: Taking cost into consideration is important in a real conservation situation. Targeting conservation action to areas with low opportunity cost can help alleviate conflicts of interest.

Weakness: If conservation priorities are chosen using only the cost-efficiency analysis, there is a risk that cost is actually driving the solution and that priorities are strongly based on low cost. This is the case especially if variation in cost is larger than variation of biodiversity value. See Arponen *et al*. (2010) for a discussion on the topic.

Note: With Zonation, it is also possible to balance multiple land uses and related opportunity costs. This analysis is described in section 5.3.4.

Link to tutorial

See tutorial Exercise 6 for an example.

5.1.8 Analysis with uncertain inputs

Planning problem to be solved

This analysis is used to account for uncertainty in the occurrence data of biodiversity features. Uncertainty in the data can arise either from statistical distribution modeling or from threats to the continued persistence of the features. In both cases, uncertainty is related to whether the areas are actually suitable for the biodiversity features in a way that can guarantee their persistence in the long term. Information-gap theory (uncertainty analysis) in Zonation gives lower conservation values to areas where uncertainty is high. Consideration of uncertainty can be included in most analyses.

Examples from literature

Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Y. Ben-Haim. 2006a. Planning for robust reserve networks using uncertainty analysis. *Ecological Modelling*, 199 (1): 115-124.

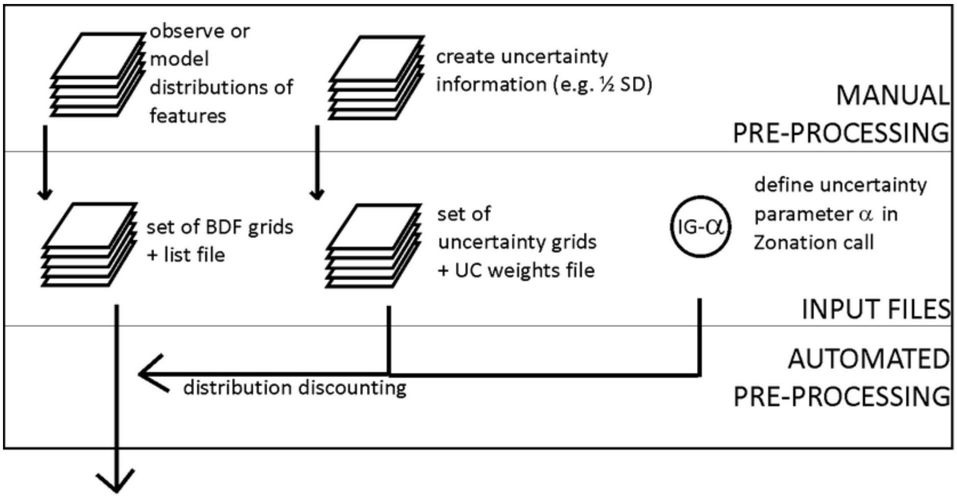
Moilanen, A., Wintle, B. A., Elith, J. and M. Burgman. 2006b. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology*, 20: 1688-1697.

Moilanen, A. and B.A. Wintle. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation*, 129: 427-434.

Carroll, C., Moilanen, A., and J. Dunk. 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. *Global Change Biology*, 16: 891-904.

The theory and algorithms behind distributional uncertainty analysis are explained in section 2.5.

Process chart for the analysis



Process chart showing the necessary input files and pre-processing to conduct an analysis with uncertainty.

Pre-processing of inputs

A pre-processing step you would often take in this analysis is fitting habitat suitability models to existing biodiversity feature data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

To include distribution uncertainty in your analysis, first define the scale of uncertainty. As explained in section 2.5.1, this can be done by giving values to (1) the uncertainty parameter α and (2) the feature- and cell-specific relative error measure w_{sc}. Feature- and cell-specific relative errors are given in the uncertainty map layers. **Using feature- and cell-specific errors (i.e., the uncertainty map layers) is optional, but a value must always be given to the uncertainty parameter**.

The uncertainty parameter α determines the horizon of uncertainty in the data and is usually unknown. Generate solutions with several a values to determine how the spatial pattern behaves with increasing uncertainty. α can be zero (no uncertainty), any positive value (distribution discounting), or any negative value (opportunity analysis).

If α is the only measure of error used, it is important that the value of α is determined in relation to the data. For example, if the form of biodiversity feature data is probability of occurrence [0,1], the uncertainty parameter should be set to a reasonably small scale (e.g., $\alpha < 0.4$) to avoid complications which may arise if all cells receive an effective discounted value of zero, resulting in a full loss of information from the distribution of the feature.

To use feature- and cell-specific errors, include:

1. Distribution uncertainty map layers for each biodiversity feature. These layers show the relative magnitude of error (uncertainty) of feature occurrence in each cell. Remember that the meaning of α must be interpreted with respect to the error measure you use. For example, if your error measure is the standard error of statistical prediction, then $\alpha=1$ essentially means subtracting one SD from the value of each cell.

2. Uncertainty analysis weights file containing a list of the feature-specific **error weights** and the name of the distributional uncertainty map layer for each feature. With error weights you can stress the data accuracy for certain (e.g., very rare) species the same way as in the biodiversity feature list file. If no feature uncertainty analysis weights are used, these parameters should be set to 1.0. Note that you can not use feature-specific error weights without the distribution uncertainty map layers

Input files

- A set of biodiversity feature grids (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
- A set of uncertainty map layers, one for each biodiversity feature (section 3.3.3.7)
- An uncertainty map weights file (section 3.3.3.7)
- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input data can be included as well.

Analysis stages and settings

To set up an analysis that accounts for uncertainty, include the following information in the run settings file:

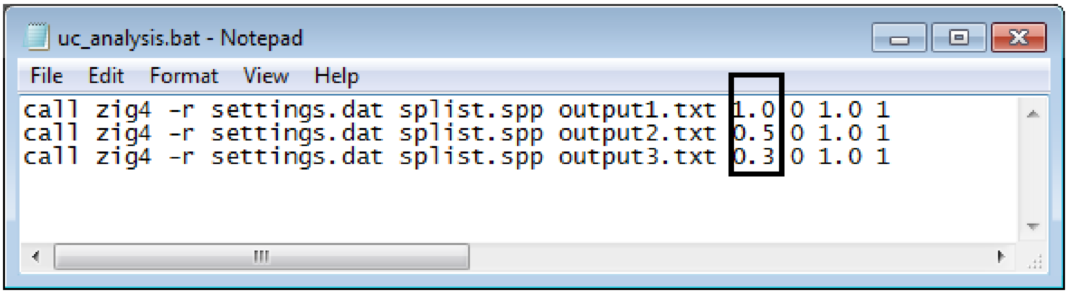
```
Info-gap proportional = 0 OR = 1
```

This parameter depends on whether errors in feature occurrences are uniform (value = 0) or proportional (value = 1). Uniform error is the default setting and works for most data sets, but in some cases it is more appropriate to use proportional errors (see e.g. Ben-Haim 2001).

```
use info-gap weights = 1
```

```
Info-gap weights file = my_IG_weights_file.txt
```

Assign a value to the uncertainty parameter α in the batch file (the fourth parameter from the end; indicated with the arrow in the image below). The α value determines the degree of uncertainty in the data and is usually unknown. An advisable strategy is to generate solutions with several α values to determine how the spatial pattern behaves with increasing uncertainty. α can be zero (no uncertainty), any positive value (distribution discounting), or any negative value (opportunity analysis). One option is to write the calls for several analyses in a single batch file, remembering to give different names for the output files of each analysis. Remember to also set the last parameter of each command line to "1" to close the command prompt application and start over after each run. Sensitivity of the solution to the uncertainty parameters can be analyzed by checking the selection frequency of priority areas after conducting several analyses with different parameters.



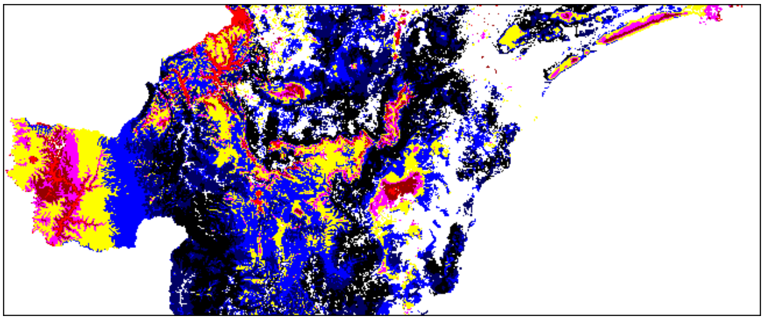
```
call zig4 -r settings.dat splist.spp output1.txt 1.0 0 1.0 1
call zig4 -r settings.dat splist.spp output2.txt 0.5 0 1.0 1
call zig4 -r settings.dat splist.spp output3.txt 0.3 0 1.0 1
```

Example of batch file for an uncertainty analysis. The alpha value for uncertainty is highlighted.

Other settings will depend on the specific aims and components of the analysis.

Output and its interpretation

This analysis produces a standard map of the landscape where red, pink, and yellow indicate sites that have both high biodiversity feature occurrence and high certainty. Depending on the amount of error in the data, the differences between the basic Zonation solution and distribution discounting can be either small or significant. The Text output window shows more detailed information about the analysis. Note that as the program starts to run the analysis, it recalculates the feature occurrence values based on the amount of uncertainty. Thus, for each feature, the program first displays the absolute value in the whole landscape (sum over all cells) and then calculates what fraction of this value can be expected to occur in the landscape with certainty. This value depends on the degree of uncertainty parameter and on the level of uncertainty in the data.



Example output map when uncertainty in species occurrences is included in the analysis.

Strengths, weaknesses, and further considerations

Strength: This analysis explicitly acknowledges uncertainty and assigns high priorities to areas with low uncertainty about feature occurrence or suitability.

Weakness: Parameter settings are subjective, which means that several analyses and consideration of solution sensitivity may be necessary.

Link to tutorial

See tutorial Exercise 6 for an example of the analysis.

5.1.9 Ecological interactions

Planning problem to be solved

This analysis is used for incorporating ecological interactions into conservation prioritization. This method makes it possible to value cells not only by the presence of a biodiversity feature within the cell, but also by its connectivity to important resources (such as a population of prey species) or to avoidable features (such as a population of a competitor species). This feature can also be used to induce connectivity of good quality habitats in the reserve network across different time steps to account for distribution shifts due to habitat management or climate change.

Zonation includes interactions by considering connectivity of suitable sites for the interacting features as a component of conservation value of each cell. For example, a cell in which an owl breeds can get a higher conservation value if there are cells with vole occurrences close by. The interactions can be either positive or negative.

Examples from literature

Ecological interactions have been utilized to account for species interactions by:

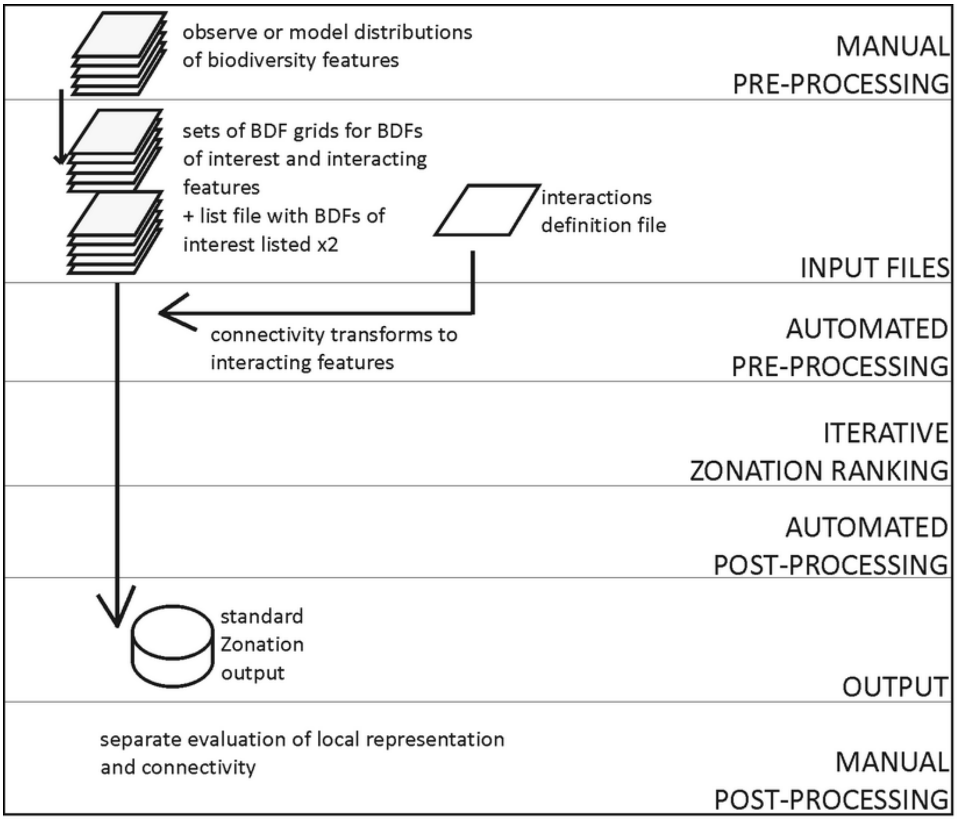
Rayfield, B., Moilanen, A. and M.-J. Fortin. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

Interaction to enhance connectivity of good quality habitat between time steps have been utilized by:

Carroll, C., Moilanen, A., and J. Dunk. 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. *Global Change Biology*, 16: 891-904.

The theory and algorithm behind ecological interactions is explained in section 2.6.

Process chart



A process chart of an analysis with ecological interactions. Please note that only the compulsory and most commonly used optional analysis components are presented. You can combine different components according to your specific needs.

Pre-processing of inputs

To run an analysis with ecological interactions included, determine the connectivity requirements between interacting biodiversity features. These are defined in the interactions definition file with parameter β.
When compiling the biodiversity feature list file, please keep in mind the following points:

- **Every interaction changes the loaded, original distribution layer**. For example, assume two features (A and B) are included in an analysis where the two original distributions plus the connection of feature A's distribution to feature B's distribution. To do this, list feature A twice in the biodiversity feature list file. The first layer will be used as it is (original distribution) while the second layer will be transformed based on its connectivity to feature B. If feature A were listed only once, the landscape ranking would be done based on the distribution of feature B and the connectivity of feature A to feature B only. The original distribution of feature A would not be included.

- **Make sure that an already transformed layer is not being used to transform other layers**. In this case, the interpretation of results is outside the scope of this manual.
- **Note that every file listed in the biodiversity feature list file will be used for landscape ranking**. If a layer is used to transform another layer but not meant to be included in the analysis itself, set the weight of that layer to zero.

Input files

- A set of biodiversity feature grid layers (section 3.3.2.1); These should include layers for the features of interest and the interacting features or the observed present suitability and predicted future suitability of habitats for each feature, depending on the analysis.
- A biodiversity feature list file (section 3.3.2.2)
- An interactions definition file (section 3.3.3.8) indicating which features interact and to what extent
- A run settings file with appropriate settings (section 3.3.2.3)

Analysis stages and settings

To run Zonation with ecological interactions, adjust the run settings file to include:

```
use interactions = 1

interaction file = my_interactions.txt
```

Output and its interpretation

In addition to the standard Zonation output, the program shows all calculations related to using the ecological interactions option in the text output (run_info.txt file). It is recommended to check this to be sure that all interactions have been loaded correctly in the beginning of the analysis to avoid false results. To understand interpretation of the curves, assume a feature has been entered in the same run as a raw distribution, a connectivity distribution, and a transformed layer showing a spatial interaction to a source. The curve for the raw quality layer tells the fraction of local habitat quality retained, the curve for the smoothed layer tells about the fraction of connectivity retained, and the curve for the interaction tells about the fraction of the potential for spatial interaction retained.

Link to tutorial

No tutorial example is provided for this analysis.

5.2 Setups and interpretations for basic planning problems

This section describes different ways to interpret the output of an analysis. The interpretations can be applied to any analysis setup.

5.2.1 Selecting conservation areas

Planning problem: Identifying the most valuable areas of the landscape to form a conservation area network.

An optimal network displays a balance of high quality areas and connectivity for conservation features. Maximizing representation of biodiversity features is a central goal in systematic conservation planning. With this approach, it is also possible to select areas for forms of conservation action other than assigning protected areas. This analysis can be adjusted to account for connectivity in various ways. It can also account for cost, uncertainty, landscape condition and retention, suitability for alternative land uses, and other factors that are present in real life conservation planning.

Examples from literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and C.D. Thomas. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

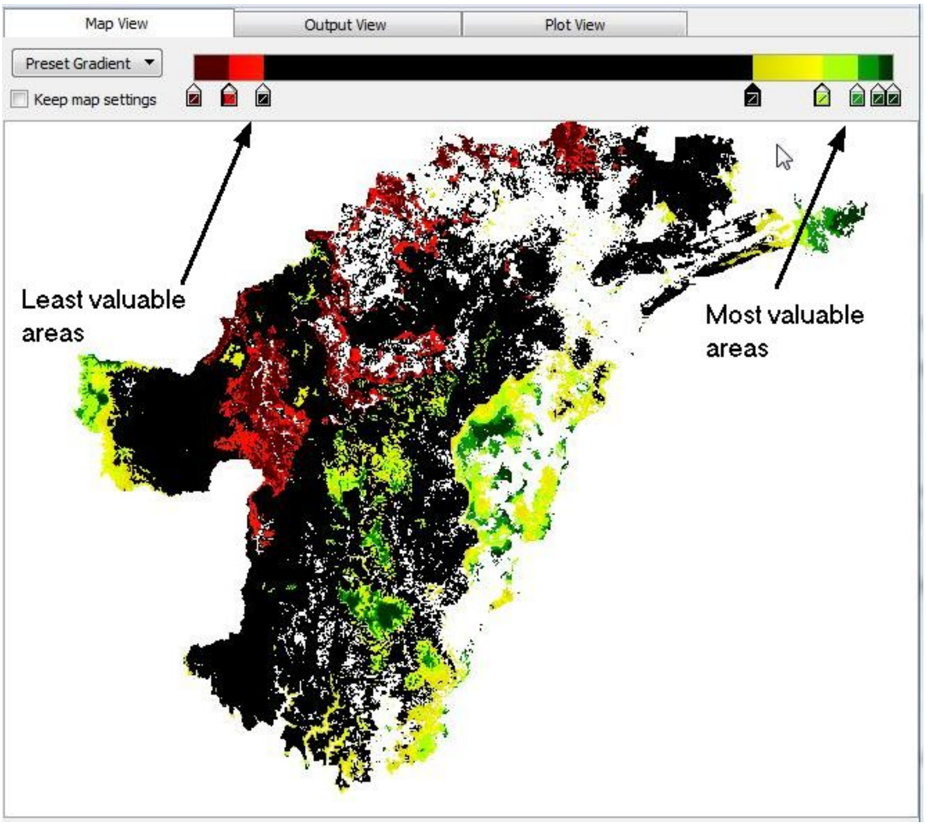
Analysis stages and settings

The analysis to identify the most valuable areas for conservation can be performed with various types of input data, setup combinations, and degrees of complexity. See simple Zonation (section 5.1.1) for the necessary input files and settings for a basic analysis. Input files, pre-processing, additional features used, and interpretation of output depend on the specific goals and data in each case. Analysis setups to include spatial and ecological factors are explained in the later sections.

Output and its interpretation

A central stage in an analysis to select conservation areas is to look for the most valuable areas in the top fraction of a Zonation priority rank map. Zonation also produces curves showing the representation level of biodiversity features as cells are removed; these curves can be used to verify the quality of the top priority areas for conservation.

The Zonation GUI contains an adjustable color gradient, making it possible to adjust the colors on the output map to match the specific planning situation. It is possible to show the most meaningful proportions of the most valuable areas as well as the least valuable areas on a single map (see figure below for an example and section 4.3.1 for advice).



Example showing a Zonation priority rank map with only the most and least valuable areas highlighted.

Strengths, weaknesses, and further considerations

Strengths: Using this approach can provide a valuable reference baseline for replacement cost analysis.

Weaknesses: In most cases, the conservation area network cannot be created from a pristine landscape. Selecting the most valuable areas across the entire landscape may not provide the most appropriate solution when previous limitations, land use plans, and existing protected areas are not accounted for.

Link to tutorial

See Exercise 1 for a tutorial example of the basic use of Zonation. Exercise 2 provides an example of basic Zonation with species weighting.

5.2.2 Identifying least valuable areas for conservation

Planning problem: Identifying areas with low conservation value and that are most suitable for other land uses.

This approach can be applied when land needs to be assigned for purposes that do not support biodiversity. By identifying the least valuable areas, it is possible to guide land use so that biodiversity is adversely affected as little as possible.

Examples from literature

Gordon, A., Simondson, D., White, M., Moilanen, A., and S.A. Bekessy. 2009. Integrating conservation planning and land-use planning in urban landscapes. *Landscape and Urban Planning*, 91: 183-194.

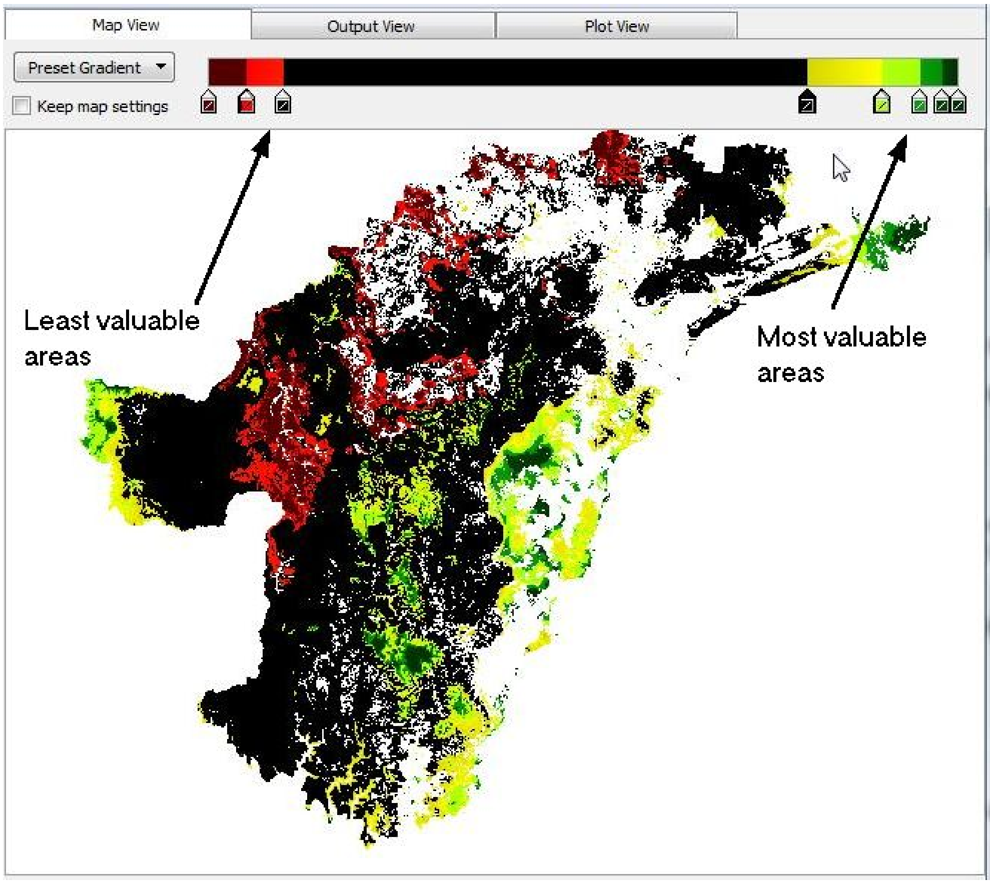
Analysis stages and settings

Identifying the least valuable areas is performed in exactly the same way as selecting conservation areas. The only difference is in the interpretation of the output. In this analysis, look for the lowest fraction of priority ranking rather than the highest. The lowest priorities are assigned to sites that contribute the least conservation value to a network. This interpretation can be applied to both simple and complex Zonation analyses. See simple Zonation (section 5.1.1) for the necessary input files and settings to run a basic analysis. You can add spatial and ecological considerations according to your needs. Setups for more complex analyses are described in the following sections.

Output and its interpretation

When identifying areas that are not that important for biodiversity, it is important to keep in mind that the results reflect only those aspects of biodiversity that were included in the analysis.

The Zonation GUI contains an adjustable color gradient, making it possible to adjust the colors on the output map to match the specific planning situation. It is possible to show the most meaningful proportions of the most valuable areas as well as the least valuable areas on a single map (see figure below for an example and section 4.3.1 for advice).



Example showing a Zonation priority rank map with only the most and least valuable areas highlighted.

Strengths, weaknesses, and further considerations

An alternative approach to identify areas for other land uses is to optimize multiple land use purposes simultaneously (see section 5.3.4).

5.2.3 Expanding conservation areas

Planning problem: Identifying areas that would best complement the already existing conservation area network (i.e., the areas that would add the most conservation value).

The primary focus of this analysis is on the quality of the new areas and the overall connectivity of the conservation area network.

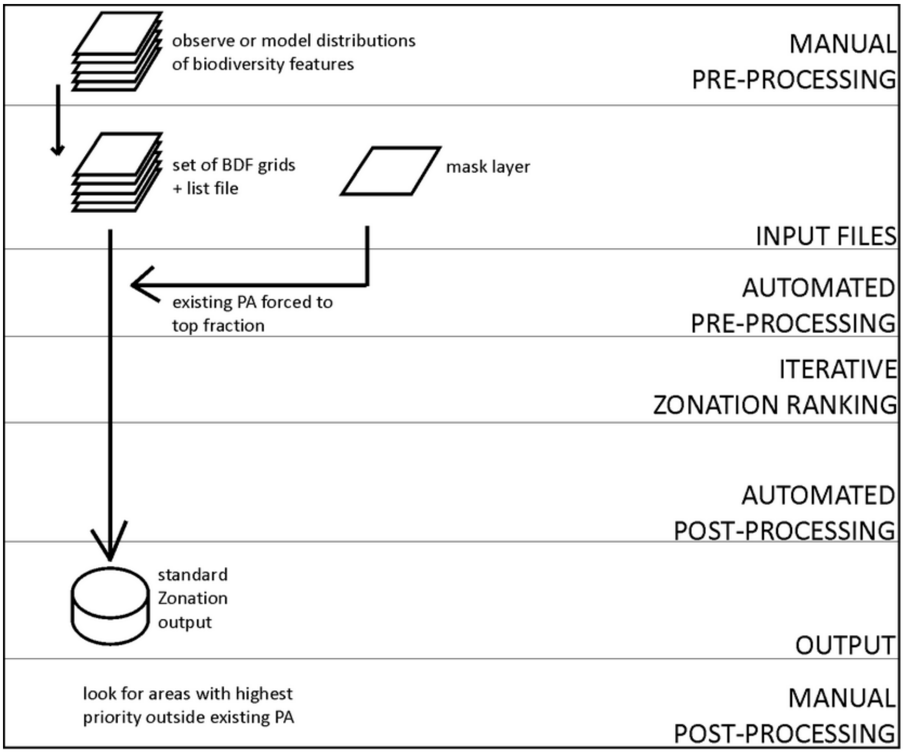
This variant of selecting conservation areas can be applied in any of the more complicated analyses. The key component of this variant is the area inclusion mask that forces the existing protected areas into the top fraction of the solution.

Examples from literature

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

Kremen, C., Cameron, A., Moilanen, A., Phillips, S., Thomas, C.D., Beentje, H., Dransfeld, J., Fisher, B.L., Glaw, F., Good, T., Harper, G., Hijmans, R.J., Lees, D.C., Louis Jr., E., Nussbaum, R.A., Razafimpahanana, A., Raxworthy, C., Schatz, G., Vences, M., Vieites, D.R., Wright, P.C. and M.L. Zjhra. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science*, 320: 222-226.

Process chart for the analysis



A process chart of the analysis for expanding a conservation area network. Please note that only the compulsory analysis components are presented. Different components can be combined according to the specific analysis.

Input files

Compulsory files to perform this analysis include:

- A set of biodiversity feature grids (section 3.3.2.1)
- Associated biodiversity feature list file (section 3.3.2.2)
- A run settings file with appropriate settings (section 3.3.2.3)
- Removal mask layer to mask in existing conservation areas (section 3.3.3.9)

Optional input files, depending on the specific setting and considerations of each case, include:

- Cost layer (section 3.3.3.6)
- Uncertainty layers and weights file (section 3.3.3.7)
- Set of landscape condition layers (section 3.3.3.14)
- Set of landscape retention layers (section 3.3.3.15)
- Matrix of community similarity, if your biodiversity features are communities (section 3.3.3.4)

Additionally, it is possible to define connectivity transforms of responses using some of the feature-specific connectivity methods available in Zonation (sections from 5.1.3 to 5.1.6).

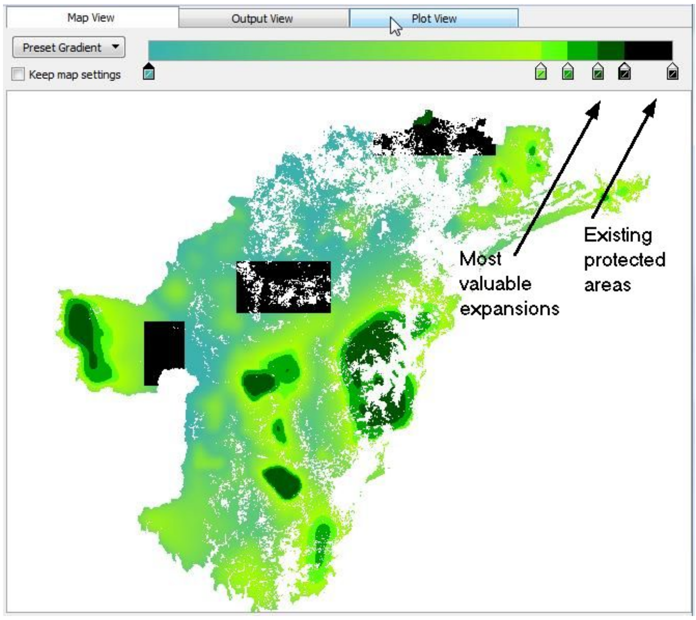
Analysis stages and settings

The specific analysis and settings depends on the specific aims of the analysis. However, at least connectivity responses would typically be included to model connectivity to existing protected areas.

Output and its interpretation

The area inclusion mask forces existing protected areas into the top fraction of the priority ranking. To identify areas to complement the existing protected area network most efficiently, look for areas with the highest ranking outside the existing protected areas.

With the Zonation GUI, the color gradient can be adjusted to easily visualize the location of the best areas outside the existing conservation network (see section 4.3.1 for advice).



Example of a Zonation priority rank map showing the output when an area inclusion mask is used to force the already existing reserve network into the highest ranking fraction. The color gradient represents the least to most valuable expansion areas.

Strengths, weaknesses, and further considerations

Strength: Provides a clear answer to a clear question.

Weakness: Set-up may be relatively complicated.

Link to tutorial

See Exercise 7 for an example.

5.2.4 Evaluating existing/proposed conservation areas

Planning problem: Evaluating the conservation value of existing or proposed networks of protected areas in terms of representation and/or connectivity.

An essential component of this analysis is replacement cost (see section 2.7) to determine how much of the maximum possible conservation value is lost if a given network is protected. The question can be considered another way, as in what additional funds or land is needed to achieve the same conservation value as in the unconstrained top fraction.

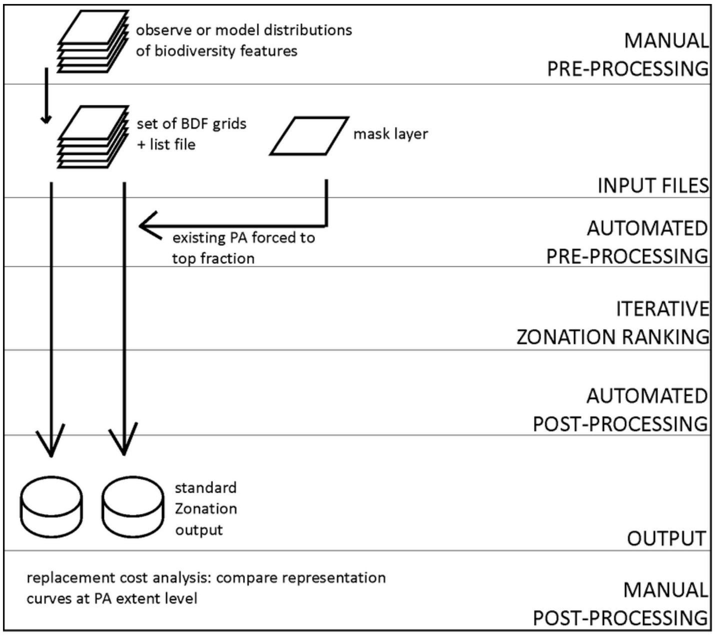
Examples from literature

Cabeza, M. and A. Moilanen. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation*, 132: 336-342.

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

Kremen, C., Cameron, A., Moilanen, A., Phillips, S., Thomas, C.D., Beentje, H., Dransfeld, J., Fisher, B.L., Glaw, F., Good, T., Harper, G., Hijmans, R.J., Lees, D.C., Louis Jr., E., Nussbaum, R.A., Razafimpahanana, A., Raxworthy, C., Schatz, G., Vences, M., Vieites, D.R., Wright, P.C. and M.L. Zjhra. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science*, 320: 222-226.

Process chart for the analysis



A process chart of the analysis for evaluating a conservation area network. Please note that only the compulsory analysis components are presented. Different components can be combined according to the specific analysis.

Input files

Necessary input files for the analysis include:

- A set of biodiversity feature grids (section 3.3.2.1)
- A feature list file (section 3.3.2.2)
- A run settings file with appropriate settings (section 3.3.2.3)
- A removal mask layer identifying existing or proposed protected areas (section 3.3.3.9)

Analysis stages and settings

In a replacement cost analysis, the priority ranking of the same biodiversity feature data set must be computed twice.

First, run a normal Zonation prioritization without a mask file, as when selecting conservation areas. The output from this run is the unconstrained solution that serves as a reference for replacement cost analysis.

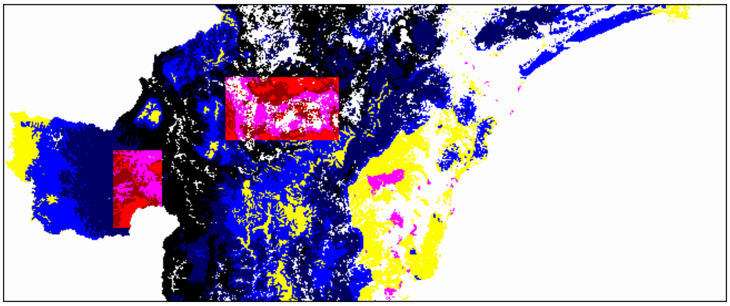
Second, run Zonation with otherwise the same settings as the first time, but now including the area inclusion mask file with areas in the current or proposed network forced into the top fraction. To include this, adjust the run settings file to include:

```
use mask = 1

mask file = my_current_network.asc
```

The detailed settings of the analysis depend on the specific combination of analysis components.

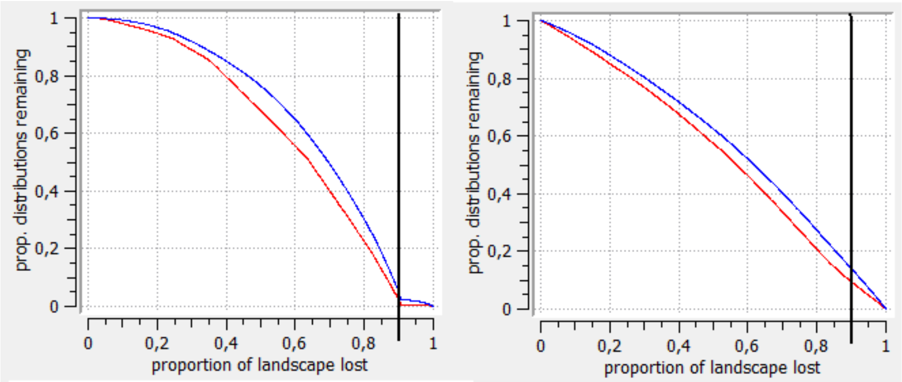
Output and its interpretation

The effects of running the analysis with a mask file should be clearly seen from the output map. The included areas should receive the highest values in landscape ranking whereas the excluded areas receive the lowest values.



Example of a priority rank map when an existing or proposed conservation network has been forcibly included in the highest ranked fraction of the solution using a mask file. Forcibly included areas are indicated by two large, red squarish areas in the map.

The distribution curves below (**.curves.txt** files, section 3.4.1) show an example comparing the performance curves between a solution where certain areas are forced into the top fraction (left) and an unconstrained solution (right). By comparing the constrained and unconstrained curves at the same proportion of landscape lost, it is possible to evaluate the costs induced by the usage of inclusion/exclusion masks (see also section 2.7). The difference between the proportion of distributions remaining at a given fraction of landscape can be considered the biological inclusion cost. Note that in many cases the use of a mask file results in a suboptimal solution. In other words, the program cannot select the best possible solution because it is forced to either exclude biologically valuable areas from or to include poor quality areas in the top fraction. This effect can be seen at the end of the curves below.



Example of distribution curves when the program has been forced to include low quality areas in the solution above (left) and an unrestricted solution (right). Note the changes at the end of the curves, which clearly demonstrate that the forcibly included areas were not what one would have ideally chosen. In the restricted solution (left), only about 0.2% of the species distributions are within the forced area (black line), whereas in the unrestricted solution, over 10% of the distribution would be protected with similar area.

5.2.5 Targeting of incentive funding

Planning problem: Obtaining information about (1) how to target conservation marketing or (2) which site to choose amongst multiple sites offered for conservation or other similar decisions.

Here, as the decision maker is not in search of the optimal solution considering the whole landscape, Zonation ranking outputs are interpreted in a different manner than usual.

Analysis stages and settings

This approach can be utilized with any set-up combination. The settings depend on the specific aims and details of your analysis.

Output and its interpretation

The relevant output for this approach is the weighted range size corrected richness (wrscr) raster (see section 3.4.1), which is one of the maps Zonation automatically produces from each analysis. The wrscr value is a sum over the weighted fraction of biodiversity feature distributions as measured from the original input files, and it illustrates the intrinsic conservation value of sites regardless of complementarity or connectivity in relation to other sites. Choices between a certain number of sites is possible by comparing their wrscr values. Range sizes and weighting of biodiversity features and the combination of analysis components will affect the wrscr scoring.

As the wrscr scoring does not account for the relative values of the site in the landscape or network context, it is advisable to also examine the basic priority rank map and base decisions on both maps.

5.2.6 Offsetting and targeting of compensation measures

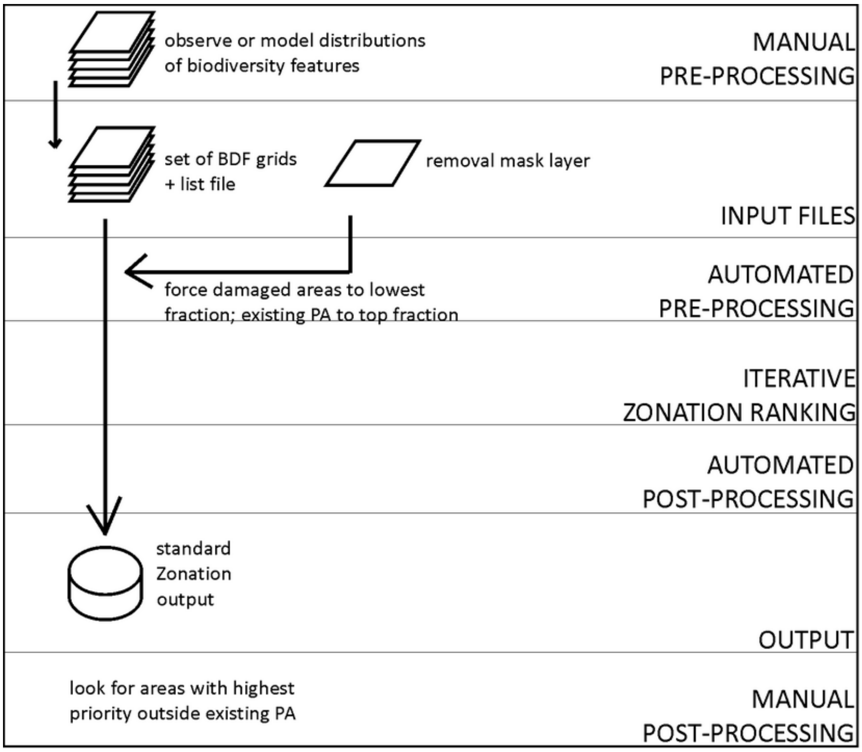
Planning problem: Identifying targets to compensate (offset) a biodiversity loss (due to economic activity) in one location by conservation action at another location.

Examples from literature

This analysis has not been published. For discussion on aspects to consider when planning offsetting, see:

Moilanen, A., van Teeffelen, A.J.A., Ben-Haim, Y. and S. Ferrier. 2009. How much compensation is enough? A framework for incorporating uncertainty and time discounting when calculating offset ratios for impacted habitat. *Restoration Ecology* 17: 470-478.

Process chart for the analysis



A process chart of the analysis for targeting compensatory measures. Please note that only the compulsory analysis components are presented. Different components can be combined according to specific analysis.

Input files

The following input files are needed to identify areas to target compensation measures after habitat conversion due to economic activity:

- A set of biodiversity feature grids (section 3.3.2.1)
- A removal mask layer (section 3.3.3.9)
    - Assign the lowest mask values (e.g. 1) to cells that are adversely effected by the planned activity to exclude them from the solution. The highest mask values (e.g. 3) should be assigned to cells that are already protected so that they will be included in the top rank. The rest of the landscape should be assigned intermediate mask values (e.g. 2).
- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, considerations of cost, alternative land uses, landscape condition and retention, etc., can also be included.

Analysis stages and settings

Adjust the run settings file in the following way:

```
use mask = 1
```

`mask file = my_mask_file.asc` (name of your removal mask layer file)

The detailed set-up will depend on the combination of components included in the analysis.

Output and its interpretation

Look for sites with high ranking outside the areas that are already protected. These would be the best sites for targeting compensation measures in terms of how they complement the existing protected area network.

Strengths, weaknesses, and further considerations

This analysis corresponds to the weak sustainability assumption (see Moilanen *et al*. 2008), which does not require that compensation is near the damaged area nor targeted to exactly the same habitats. If strong sustainability is required, then the analysis needs to be limited to damaged habitats and connectivity to damaged areas must be included, but this could result in an inefficient use of resources.

Link to tutorial

There is no tutorial example for offsetting provided. Please refer to Exercise 7 for a tutorial example for using mask files.

5.2.7 Target-based planning

Planning problem: Target-based planning is about identifying the part of the landscape that satisfies given feature-specific targets with minimum cost (or area, if cost information is missing). This is the typical planning problem in Systematic Conservation Planning. See section 2.3.3 for additional technical information.

In extended target-based planning the idea is that targets are reduced and the saved resources are used in a manner that provides high return on (conservation) investment (ROI). At the same time targets are not fully given up, implying that some minimal representation is guaranteed for all features. As a positive, Zonation actually by default does this type of analysis. When above targets, resources are allocated as with ABF, which implies high efficiency.

Also, note that the basic implementation of targets in Zonation effectively lets the feature go after it fails the target. A more stable method could be to use steep sigmoids instead of the target. Steep sigmoids can be set up using the generalized benefit function (Section 2.3.4).

Examples from literature

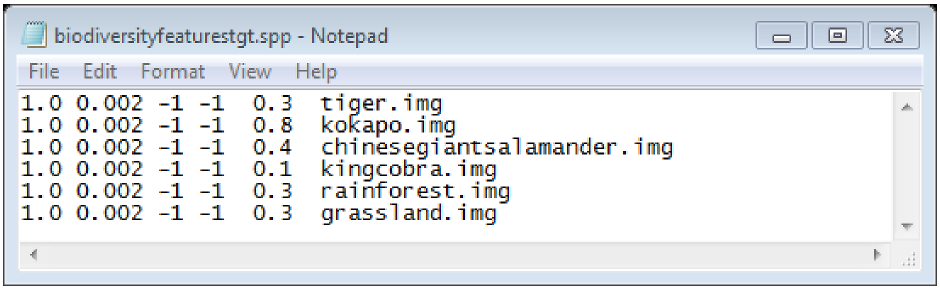
How Zonation does targets is described in
Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

For the logic of extended target-based planning see
Laitila, J. and A. Moilanen. 2012. Use of many low-level conservation targets reduces high-level conservation performance. *Ecological Modelling,* 247: 40-47.

The following paper describes inverse spatial conservation planning, which is similar to having targets for negatively weighted features.
Kareksela, S., Moilanen, A., Tuominen, S. and J.S. Kotiaho. 2013. Use of Inverse spatial conservation prioritization to avoid biodiversity loss outside protected areas. *Conservation Biology*, 27: 1294–1303.

Analysis stages and settings

This analysis is simple to implement and is just like any other analysis you can find in previous sections. The feature-specific targets are defined in the 5th column of the biodiversity feature list file (just before the file name on each row). In the example below, the features have been given targets between 0.1 and 0.8.



```
1.0 0.002 -1 -1  0.3  tiger.img
1.0 0.002 -1 -1  0.8  kokapo.img
1.0 0.002 -1 -1  0.4  chinesegiantsalamander.img
1.0 0.002 -1 -1  0.1  kingcobra.img
1.0 0.002 -1 -1  0.3  rainforest.img
1.0 0.002 -1 -1  0.3  grassland.img
```

Note: even if a feature has a negative weight (in the first column), the target is still interpreted as the fraction of the feature distribution that Zonation tries to keep. However, Zonation will remove the remainder of the distribution as soon as possible. For example, if you wish to remove 80% of a negatively weighted feature, you should set the target as 0.2. In other words, negatively weighted features are like opportunity costs of conservation. A target on such would imply limiting the cost of the solution to (1-target) of the cost of the full landscape (for that cost layer). See also Kareksela et al. (2013), above, for inverse prioritization for impact avoidance.

Output and its interpretation

This analysis produces a full priority ranking, as always. However, it is important to observe that the ranking is most relevant for the part of the ranking where no targets have failed (Section 2.3.5). The minimum set coverage solution is the highest top fraction of the landscape that does not violate a single target. The fractions of the landscape when the each feature violates target can be found from the features_info.txt file output by Zonation, see section 3.4.1, automated file output. Figure out the fraction from that file, and then use GUI to display respective top fraction of the landscape.

Strengths, weaknesses, and further considerations

Strengths: Target based planning is a traditional method in SCP, and it is easy to understand.

Weaknesses: Setting possibly thousands or tens of thousands of targets is difficult and prone to producing inefficient solutions, as described in Laitila and Moilanen (2012) and references therein. A midway method is the extended target-based planning described in the same work.

Link to tutorial

Ready-made files for this functionality are not distributed, but you could modify the feature list file of exercise 1 to have targets, and then set the removal rule to target-based planning in the settings file, that is, removal rule = 3.

5.3 Full analysis setups

The following sections provide examples of full analysis setups in different types of planning contexts. The examples presented here can be combined, but proceeding with care is advisable. Section 5.3.10 provides some guidelines for creative combinations of analysis components.

5.3.1 Community level analysis

Planning problem to be solved

In the examples so far, as often in conservation planning, biodiversity data has been a set of individual biodiversity features. The community level analysis, however, operates on habitat types or community classes. The aim in the community level priority setting is to maximize complementary representation of habitat or community types, weighted by their richness. The complementarity approach utilizes information about similarity (i.e., shared features between community classes).

Examples from literature

The theoretical basics of community level conservation prioritization is explained by:

Arponen, A., Moilanen, A., and S. Ferrier. 2008. A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology*, 45: 1436-1445.

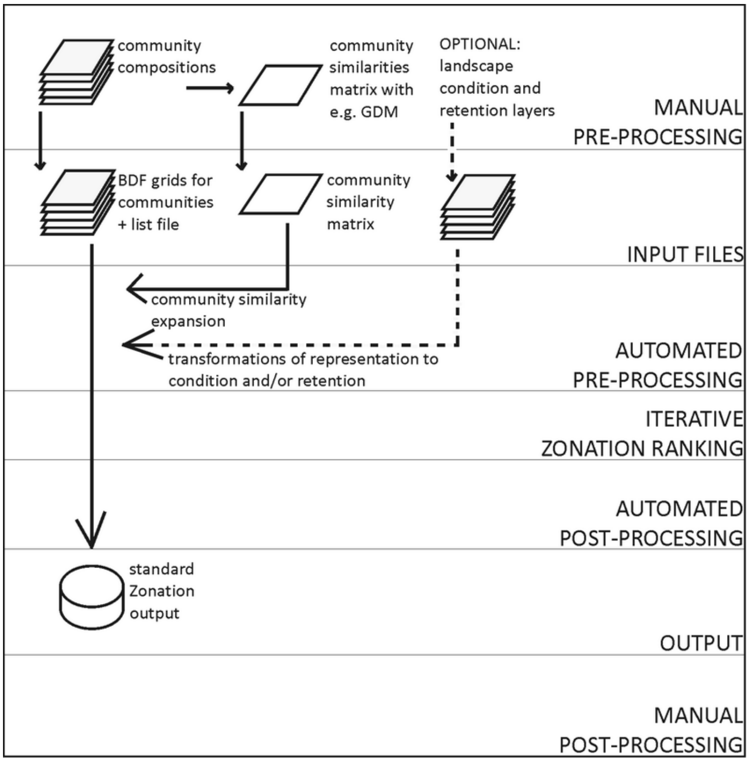
An example of implementation in Zonation can be found in:

Leathwick, J.R., A. Moilanen, S. Ferrier and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

Please note that Leathwick *et al*. (2010) utilize Zonation v.2.0. for which the data must be pre-processed before running the analysis. Zonation v.3.1 and higher does the pre-processing by itself.

The theory and algorithm behind the community level analysis is explained in section 2.8.

Process chart for the analysis



A process chart of a community level analysis. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

A common pre-processing step is fitting habitat suitability models to existing species data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

A necessary input for a community level analysis is a community similarity matrix. This assigns a similarity value to all possible pairs of community types and can be compiled utilizing a generalized dissimilarity model (GDM; see Ferrier 2002) in R with the package 'gdm' (http://www.biomaps.net.au/gdm/).

A meaningful method for assigning weights to community types is to weight them by their species richness. Weights are assigned in the first column of the biodiversity feature list file.

See section 5.3.5 for necessary pre-processing steps if landscape condition and retention is included.

Input files

- A set of community type grids (section 3.3.2.1)
- Associated biodiversity feature list file (section 3.3.2.2); Weight community types according to their species richness
- A community similarity matrix (section 3.3.3.4) describing the overlap in species composition between community types
- A run settings file with appropriate settings (section 3.3.2.3)

It is possible to include considerations of cost, landscape condition and retention, or other analysis components depending on the specific aims and considerations of the analysis.

Analysis stages and settings

To run a community level analysis, adjust the run settings file to include:

```
[Community analysis settings]
load similarity matrix = 1
community similarity matrix file =
  my_community_similarity_matrix.txt
apply to representation = 1
```

Output and its interpretation

When examining the representation of community types, it is worth noting that the coverage for communities after similarity transformation may not be equivalent to the coverage of the original community types. Use one of these alternative methods to obtain representation curves for the original, unexpanded community types:

- Use solution loading (section 3.5.2) to load the output rank file from the community level analysis. Define the original community types as biodiversity features, and do not include the community similarity matrix.

- List the original community types twice in the biodiversity feature list file, but assign zero weights for one set. With this method, Zonation will produce the representation curves for the original, unexpanded community types in the same run that the similarity expansion is used. With this analysis, it is convenient to use a groups file (section 3.3.3.12), where the layers with zero weights are assigned one output group and the similarity-transformed layers are assigned separate output group. This will also automatically produce a summary of representation levels for the original data.

Strengths, weaknesses, and further considerations

Strengths: Fewer features are needed for the community level analysis than when working with individual species.

Weakness: Consideration of connectivity in the community context is more abstract than in the individual species context.

In many countries, conservation legislation is based on habitat types, which can be seen as surrogates for community types. Note, however, that accounting for variation in richness and overlap in community composition greatly increases the accuracy of the analysis compared to setting priorities for habitat types only. Consideration of landscape condition and retention is a natural component in this analysis.

Link to tutorial

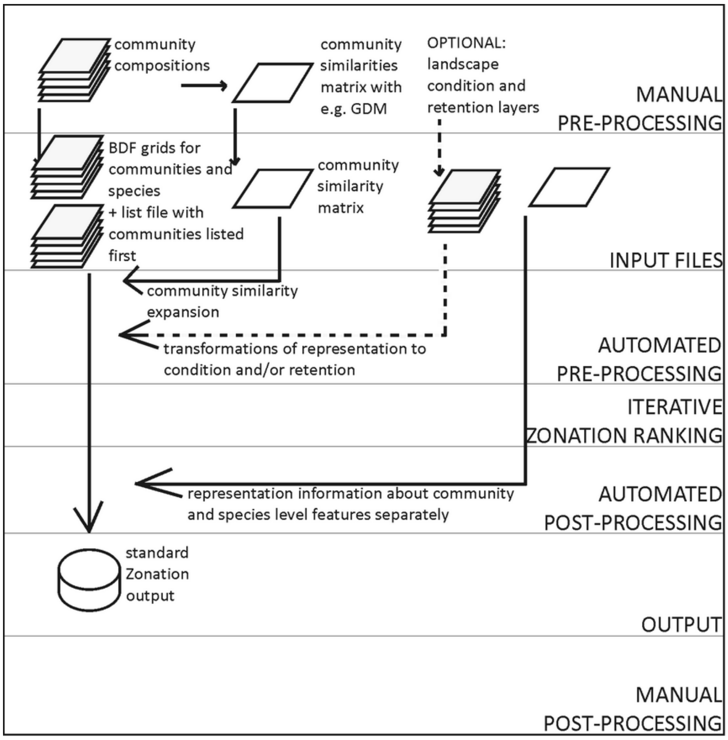
See Exercise 10 for an example of a community level analysis.

5.3.2 Combined community and species level analysis

Planning problem: Identifying conservation priorities based on a mixture of community types and occurrences of individual biodiversity features.

This approach may be useful when conservation decisions are mostly based on community types but representation of some individual biodiversity feature also needs to be considered.

Process chart of the analysis



A process chart of a combined community and species level analysis. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Input data

- A set of community type grids (section 3.3.2.1)
- A set of feature occurrence grids (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
  - List community features first followed by individual biodiversity features. This is necessary as there is a direct link between the rows in the biodiversity feature list and the community similarity matrix files. In other words, row 1 in the community similarity matrix file is linked to row 1 in the biodiversity feature list file.
- A community similarity matrix (section 3.3.3.4) describing the overlap in feature composition between community types
  - Be sure that the order that communities are listed here is the same as in the biodiversity feature list file.
- A groups file (section 3.3.3.12)
  - Places representation curves for community and biodiversity features in separate group output files. Assign community level features to output group 1 and biodiversity features to output group 2 by typing the respective values in column 1 of your groups file.
- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, landscape condition and retention layers, etc. can be included.

Analysis stages and settings

Please refer to the previous section (5.3.1) for necessary pre-processing and community level analysis settings.

To assign different types of biodiversity features to output groups, adjust the run settings file by including the following lines:

```
use groups = 1
groups file = my_groups_file.txt
```

under the `[Settings]` header (whereas the settings for a community level analysis would be under the `[Community level analysis]` header).

Link to tutorial

An example of combined community and species level analysis is provided in Exercise 10.

5.3.3 Single-species prioritization

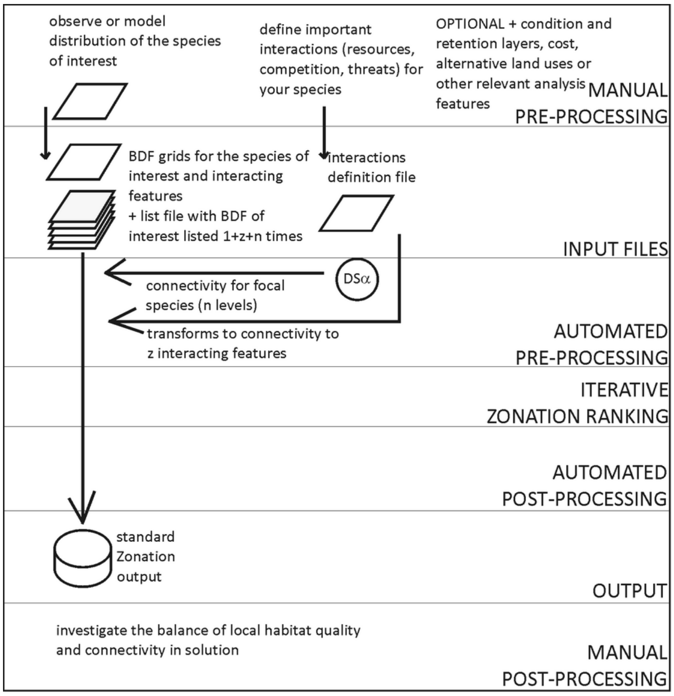
Planning problem: Identifying conservation priorities for one (or a few) biodiversity features with ecologically sound considerations of multiple aspects.

The aim is to optimize the protected area network to cover well-connected high quality patches for a single features (e.g., to protect foraging and breeding areas) while avoiding harmful features such as human habitation or invasive species. Availability of resources and other vital ecological interactions can also be included, but this analysis requires ecological information about the feature in addition to the occurrence data. Connectivity of multiple features and on multiple spatial scales can also be accounted for.

Examples from literature

Rayfield, B., Moilanen, A. and M.-J. Fortin. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

Process chart for the analysis



A process chart of a single-feature prioritization analysis. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

To be able to consider multiple ecological aspects of a biodiversity feature in a conservation planning context, all the aspects need to be quantified and set into a spatial context. When considering connectivity within and between populations, two (or more) widths of dispersal kernels need to be defined for each feature. Distributions of resources and other vital interactions, as well as threatening features, need to be transformed into raster files.

Input files

To account for connectivity at multiple levels in the prioritization, include:

- A biodiversity feature grid layer (section 3.3.2.1) describing habitat quality (or probability of occurrence) for the biodiversity feature
- A biodiversity feature list file (section 3.3.2.2)
  - The number of times the feature layer must be listed is the number of connectivity considerations and interactions included + 1. Assign a different value of the dispersal a for each level of connectivity (see section 5.1.3 for details).
- A run settings file with appropriate settings (section 3.3.2.3)

To account for connectivity to resources or other vital ecological interactions, or to favor sites that are not connected to threatening features such as human habitation or invasive species, include:

- Occurrence grid layers for the interacting features
  - Add these layers to the biodiversity feature list file. If one or several interacting features should not be accounted for in the prioritization, give each one a weight of zero in the biodiversity feature list file. This will ensure that the interaction occurrence will only be used to estimate the conservation value of any cell in terms of how well it is connected to cells with interacting features.
- An ecological interactions definition file (section 3.3.3.8)
  - Define resources as positive interactions (type 1) and threatening features as negative interactions (type 2).

It is also possible to include considerations of cost, landscape condition and retention, or other components.

Analysis stages and settings

To include ecological interactions, adjust the run settings file to include the following lines:

```
use interactions = 1
interaction file = my_interactions.txt
```

Adjust the batch file to indicate that distribution smoothing is applied in the analysis (see section 5.1.3):

- Set the third last parameter of the command line call to 1 to indicate that distribution smoothing will be done.

- Give a factor for multiplying the feature-specific α -values as the second last parameter in the command line. α -values (feature-specific widths of kernel) are in the second column of the biodiversity feature list file. This factor is useful if you are interested in running multiple solutions assuming several levels of dispersal capabilities because it allows you to multiply all dispersal capabilities simultaneously. Thus you do not need to change the parameters manually in the biodiversity list file after each run. If you do not wish to multiply the α -values, set this factor to 1.

Output and its interpretation

To assess the conservation solution, investigate the balance between habitat quality and connectivity (see Rayfield *et al*. 2009). An ideal solution is one where only little local quality is lost for the sake of improved connectivity. Compare the balance of features and connectivity by examining the respective performance curves for the features and their connectivity transformed versions.

Strengths, weaknesses, and further considerations

Strength: By including multiple ecologically realistic spatial considerations, it is possible to (partially) account for the processes that support persistence of the biodiversity feature in the long term.

Weakness: Systematic conservation planning is, in general, a multi-feature enterprise. Placing a lot of emphasis on a single feature is questionable in that sense.

It is possible to include basic occurrence data for a large set of biodiversity features and more detailed considerations for a few key features.

Link to tutorial

A tutorial example is not provided for this analysis. Please refer to Rayfield *et al*. (2009) for a case example.

5.3.4 Balancing alternative land uses considering multiple costs

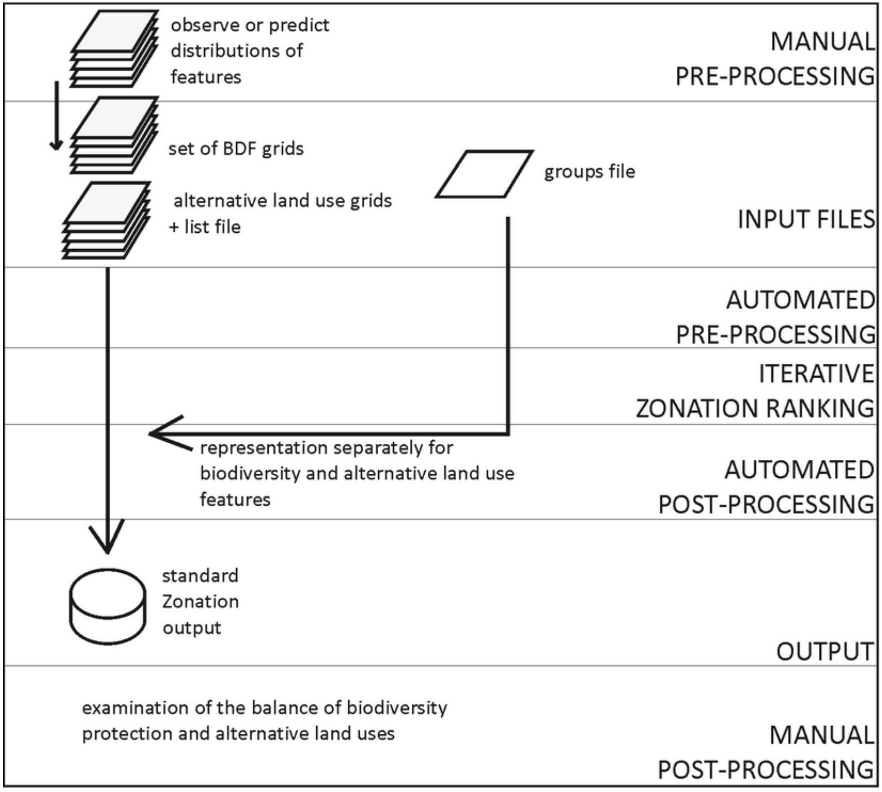
Planning problem: Finding a solution where land use is balanced between biodiversity conservation and other purposes.

The aim is to separate conservation priorities from competing land uses. Areas with ongoing or intended use for other purposes are given negative weights in feature weighting to ensure that those areas will be removed first in the ranking process. This analysis can help to identify multiple land use priorities and alleviate conflicts of interests.

Examples from literature

Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J. and C.D. Thomas. 2011. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

Process chart for the analysis



A process chart of an analysis that balances multiple land uses. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

No pre-processing is needed for the alternative land uses analysis itself. Other necessary pre-processing depends on the details of the analysis (e.g., the kind of connectivity consideration you include).

Input files

- A set of biodiversity feature grids (section 3.3.2.1)
- A set of alternative land uses grids (section 3.3.3.13)
- A biodiversity feature list file(section 3.3.2.2)
  - Please note that both biodiversity features and alternative land uses are listed in the same file. The difference is that the alternative land use features are assigned negative weights in the biodiversity feature list file.
- A groups file (section 3.3.3.12)
  - Generates representation curves separately for the biodiversity features and for the negative features. Assign biodiversity features to output group 1 and

alternative land use features to output group 2 by typing the respective values to column 1 of the groups file.

- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, landscape condition and retention layers, etc., can be included in the analysis.

Analysis stages and settings

Other than including a groups file, there are no specific settings needed to run a simple balancing land uses analysis. See simple Zonation (section 5.1.1) for the basic settings. The detailed combination of settings for the analysis depends on the features that are included.

To include the groups file, adjust the run settings file to include:

```
use groups = 1
groups file = my_groups_file.txt
```

Output and its interpretation

An optimal solution for balancing land uses would be one in which a small low-ranked fraction of the landscape includes most of the negative features, whereas a small top fraction would have high representation of the positive features (biodiversity). If such a solution can be obtained, then competing land uses and biodiversity conservation can be optimized simultaneously without conflict of interest.

Strengths, weaknesses, and further considerations

Please note that the cost efficiency analysis and alternative land uses analysis are fundamentally different in the technical sense. In the cost efficiency analysis, biodiversity value of a cell is divided by cost while in the alternative land uses analysis, the values of negative features are subtracted from the biodiversity value. Both approaches can be utilized in a single analysis in which the overall structure becomes [(weighted aggregate of benefits) - (weighted aggregate of opportunity costs)] / (direct costs).

Link to tutorial

See Exercise 2 for an example of an alternative land use analysis.

5.3.5 Balancing representation and retention

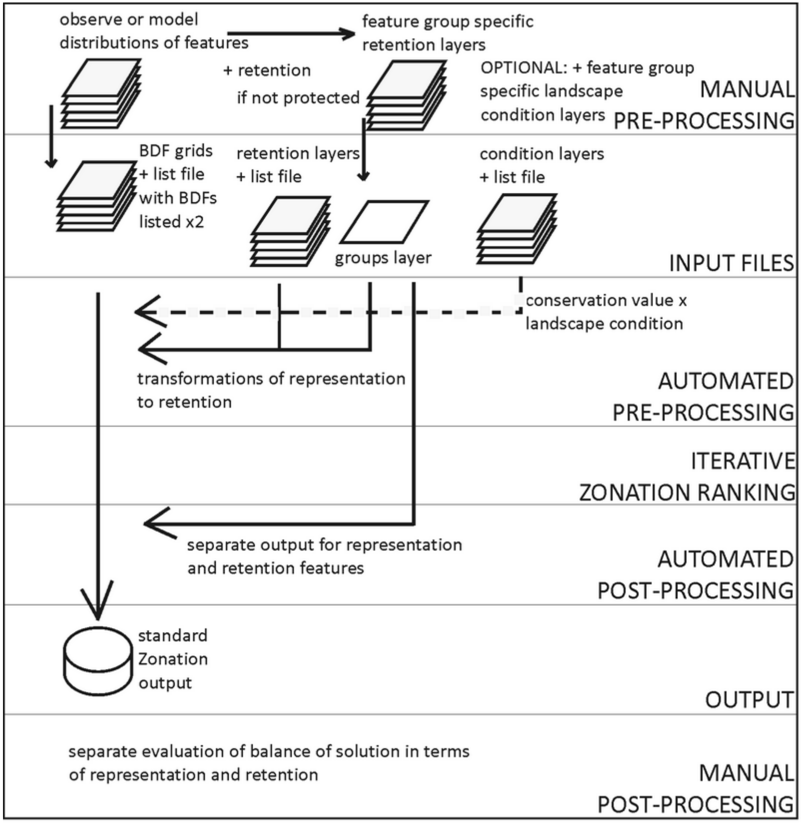
Planning problem: Identifying conservation priorities when the non-protected landscape can also support some biodiversity value or when conservation value (population sizes, habitat condition) can be increased with appropriate management.

This analysis balances representation of features in the protected area network and their retention across the full landscape. This approach also indirectly acknowledges that some biodiversity features or communities are in need of protection through protected area assignment more than others. For example, features that are confined to old-growth forests may need protected areas more than generalist features of semi-open areas.

Examples from literature

Moilanen, A., Leathwick, J.R. and J.M. Quinn. 2011. Spatial prioritization of conservation management. *Conservation Letters*, 4: 383-393.

Process chart for the analysis



A process chart of an analysis that balances representation and retention. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

If the analysis focuses on community features, see section 2.8 and section 3.3.3.4 for information about the community similarity matrix.

For running an analysis with retention, it is necessary to duplicate the list of features in the biodiversity feature list file. The first set of layers are used to model conservation value with representation in reserves while the second set is used for modeling retention in the landscape without conservation action (or through a management intervention, using mode 2).

Several groupings of biodiversity features are necessary for including landscape (condition and) retention. First, it is advisable to divide the two copies of the biodiversity list into separate output groups. This will allow comparison of the difference between priorities for representation and retention. Next, divide the biodiversity features into groups according to how past and projected changes in the landscape will affect condition and retention for each feature. For example, for forest species that prefer an undisturbed environment, the landscape condition would be decreased as a result of past clear-cutting of a patch of forest. For species that benefit from openings in the landscape, however, a moderate clear-cut would have no effect on the condition. The same applies to retention. The groups do not need to be the same for condition and retention, although it would often make the most sense. Prepare a groups file (section 3.3.3.12) to indicate which features belong to which output groups.

To include condition, create one condition layer for each condition group. Assign each grid cell a value between 0 and 1 according to how past changes have affected the condition of that cell. Make a list of the layers in a condition layer list file. Make sure that the rows in the condition layer list file match those in the groups file (see section 3.3.3.14).

To create the retention layers, an estimate of how changes in the landscape will reduce (or improve) the conservation value of each grid cell for each retention group is needed. Create retention layers and a list file (see section 3.3.3.15 for details). Make sure that the rows in your retention layer list file match those in the groups file (see section 3.3.3.12). Define the retention mode (whether the conservation value is expected to increase or decrease) for each retention group in the groups file as well.

Decide how to balance representation in selected reserves and retention in the landscape. To set the balance, assign a value for the parameter β, which is a decimal multiplier for the retention layer weights to balance between representation and retention.

Input files

- A set of biodiversity feature grid layers (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
  - List the biodiversity features twice. The first copy of layers models representation while the second copy will be modified to model loss if the cell is not selected.
- A community similarity matrix, if features are communities (section 3.3.3.4)
- A groups file (section 3.3.3.12)
  - Assigns features to retention groups (column 3 in groups file) and determines the retention mode (column 4). It is advisable to assign the layers used to model representation and the ones used to model retention to separate output groups. Insert suitable output group numbers to layers in the first column of this file.
- A set of retention layers, one for each retention group (section 3.3.3.15)
- A retention layer list file (section 3.3.3.15)

- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, alternative land use layers, etc. can be included in the analysis.

Analysis stages and settings

To run an analysis to balance representation and retention, adjust the run settings file to include:

```
use groups = 1
groups file = my_groups_file.txt
use retention layer = 1
retention file = my_retention_file.txt
retention layers relative weight = value for β, a decimal multiplier for the
                                   retention layer weights to balance between representation and
                                   retention
```

Output and its interpretation

When interpreting the output from retention analysis, it is important to note that Zonation tries to cover proportions of distributions, not absolute amounts. Retention layers are also treated as normalized distributions. A proportion of a small absolute amount (absolute fraction lost) may not have the same relevance as the same proportion from a large absolute loss from another feature.

Strengths, weaknesses, and further considerations

Strength: An approach where landscape retention is considered alongside representation in the protected area network is more realistic than considering the protected areas as the only fractions of landscape supporting biodiversity values. It is relevant to account for what would be lost in the absence of action.

Weakness: In the analysis, landscape dynamics and retention are dealt with in a relatively simplistic manner. Nevertheless, it is still a decent first approximation.

Link to tutorial

See Exercise 11 for a tutorial example.

5.3.6 Habitat restoration and dynamic landscapes

Planning problem: Accounting for time and habitat dynamics in conservation prioritization.

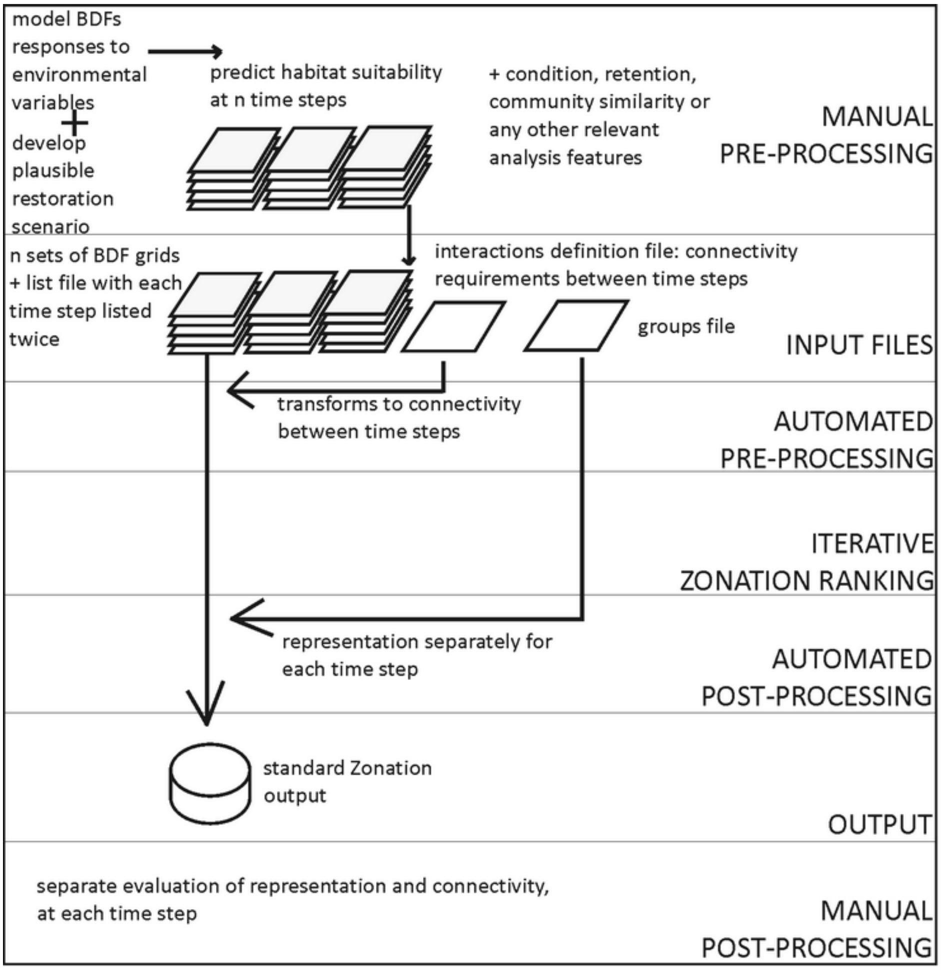
In the habitat restoration and dynamic landscapes analysis, distributions are assumed to be dynamic, yet stationary. Predictions of habitat suitability or biodiversity feature distributions are produced for multiple time steps. The prevailing conditions at each time step are extracted from a scenario of habitat restoration actions at appropriate locations across the landscape. The operational key to the analysis is inputting habitat suitability models for multiple features in a single analysis. Using core-area Zonation as the cell removal rule ensures that the solution will include suitable sites for all biodiversity features at every time step.

Examples from literature

Thomson, J.R., Moilanen, A., McNally, R., and P. Vesk. 2009. Where and when to revegetate: A quantitative method for scheduling landscape reconstruction. *Ecological Applications*, 19: 817-828.

For more information about an analysis with dynamic landscapes, see section 2.11.

Process chart for the analysis



A process chart of an analysis that considers landscape dynamics. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

A necessary pre-processing step in a dynamic landscape analysis is fitting habitat suitability models to existing biodiversity feature data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region. A scenario for habitat restoration or other landscape dynamics must also be developed. In the scenario, actions should be applied in locations where they are most feasible. Data about the cost-efficiency of different actions can be utilized if it is available. The scenario should be expressed as quantitative habitat properties at discrete time steps for each grid cell where the properties are the most relevant ones explaining variation in biodiversity feature occurrences. Any distribution modeling approach can be used to generate predictions of habitat suitability (usually expressed as probability of occurrence) for each biodiversity feature at each time step.

Input files

- A set of biodiversity feature grids for each time step (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
  - Feature grids for all time steps are listed in the same file
- A run settings file with appropriate settings (section 3.3.2.3)
- A groups file (section 3.3.3.12)
  - Generates representation curves separately for each time step. Insert suitable values to column 1 of the groups file.

While not necessary, it is possible to include a removal mask layer (section 3.3.3.9) in the analysis to mask in existing good quality habitats that are not likely to be affected by habitat restoration. Using the area inclusion mask ensures that good quality habitats are removed from the landscape last. It is advisable to use this in situations where good quality habitat is currently existing in the landscape.

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, landscape condition and retention layers, etc., can also be included.

Analysis stages and settings

Use the predicted probabilities of occurrence for multiple time steps as a basis for conservation prioritization. Grid layers for all time steps are listed in a single biodiversity feature list file. Adjust the run settings to include the files used in running simple Zonation (section 5.1.1) to evaluate existing conservation areas or a removal mask layer when expanding conservation areas (section 5.2.3).

Adjust the run settings to also include a groups file:

```
use groups = 1
groups file = my_groups_file.txt
```

Applying distribution smoothing, boundary quality penalty, or other connectivity considerations in the analysis is advisable. Please refer to the corresponding sections for necessary input files and settings.

Output and its interpretation

The ranking in this analysis should be interpreted as a prioritization order that maximizes the availability of suitable habitat across all time steps given that the habitat restoration scenario created during the pre-processing stage is successfully executed.

Strengths, weaknesses, and further considerations

Weakness: This analysis is data intensive and complicated to set up properly as it requires quantitative estimates of the habitat quality development. The underlying scenario is susceptible to uncertainties on at least two levels: 1) the full implementation of the scenario and 2) the real effects of action on habitat quality.

Connectivity between time steps could be set up in many ways. For example, the interlinkedness of the different time steps for one feature can be emphasized by inputting the time steps as ecological interactions (see section 5.1.9).

Link to tutorial

There is no tutorial provided for this analysis. Please refer to Exercise 1 for selecting conservation areas and Exercise 7 for using an area inclusion mask.

5.3.7 Setups for climate change

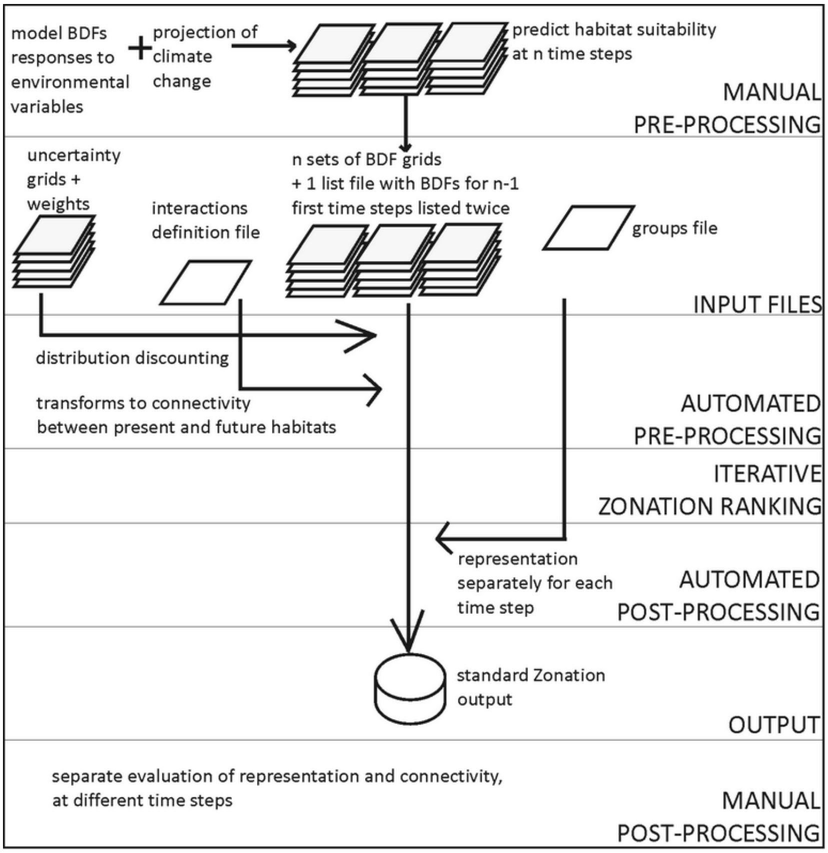
Planning problem: Identifying priorities for conservation to account for changing distributions due to climate change.

The analysis for a climate change scenario is similar to that of habitat restoration and landscape dynamics. In this analysis, however, biodiversity feature distributions are assumed to be dynamic and non-stationary. Connectivity of suitable habitats at different time steps needs special attention in this type of analysis. An uncertainty analysis also needs to be included as a high amount of uncertainty is related to considerations of climate change.

Examples from literature

Carroll, C., Dunk, J. and A. Moilanen. 2010. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. *Global Change Biology*, 16: 891-904.

Process chart for the analysis



A process chart of an analysis that considers habitat quality and connectivity between time steps in the course of climate change. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

Obtain a projection of climate development under a given emission scenario for one or more time steps. It is advisable to consider multiple climate models and emission scenarios as there is large variation in climate projections depending on the climate model used and the underlying emission scenario. Create spatial predictions of future habitat suitability (or probability of occurrence) for the biodiversity features. Again, as variation between different models is high and model validation is difficult or impossible, considering a number of modeling strategies using ensemble forecasting or a similar strategy is advisable.

The connectivity between present and future suitable habitats are applied as ecological interactions (type 1; see section 3.3.3.8). Connectivity requirements for each biodiversity feature are defined as widths of dispersal kernels (see section 5.1.3).

To account for uncertainty, uncertainty maps for each biodiversity feature must be produced. The standard deviation of the model predictions can be utilized to produce the maps (see section 2.5.1).

Please refer to section 5.3.5 for the necessary pre-processing and other settings if landscape condition and retention will be included in the analysis.

Input files

- A set of biodiversity feature grid layers describing present distributions (section 3.3.2.1)
- A set of biodiversity feature grid layers describing future distributions (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
    - List both present and future layers in a single file. List present biodiversity features twice. One set will account for present distributions while the other will be transformed to represent connectivity between present and future distributions.
- An ecological interactions definitions file (section 3.3.3.8)
    - Note that here the interaction is not actually between features, but between present and future distributions of a single feature.
- A set of uncertainty map layers (section 3.3.3.7)
- An uncertainty weights file (section 3.3.3.7)
- A groups file (section 3.3.3.12)
    - Produces representation curves separately for each time step. Insert suitable values to column 1 of the groups file.
- A run settings file with appropriate settings (section 3.3.2.3)

Optional input files include a removal mask layer (section 3.3.3.9) to force existing protected areas into the top fraction.

Depending on the specific aims and details of the analysis, other components such as considerations of cost, landscape condition and retention, etc., can also be included.

Analysis stages and settings

To identify conservation priorities that are likely to have high conservation value both in the present and the future, adjust the run settings as follows:

Include interactions (section section 5.2.9)

```
use interactions = 1
interaction file = my_interactions_definitions_file.txt
```

Include a groups file

```
use groups = 1
groups file = my_groups_file.txt
```

Include distribution discounting (section 5.2.8)

```
[Info-gap settings]
use info-gap weights = 1
Info-gap weights file = my_IG_weights_file.txt
```

Determine the value of the uncertainty parameter α in the command line (the fourth last parameter). The a value determines the horizon of uncertainty in the data and is usually unknown. Thus, it is necessary to generate several solutions using different α values to determine how the spatial pattern behaves with increasing uncertainty.

To include current reserves (expanding conservation areas), landscape condition and retention, cost, or other components, please refer to the corresponding sections for setting details.

Output and its interpretation

As the network should maintain both present and future biodiversity, examination of the output to evaluate representation should be done separately for present and future occurrences of biodiversity features.

Strengths, weaknesses, and further considerations

Strength: Allows accounting for the ongoing change in climate, which is bound to affect biodiversity feature's distributions.

Weakness: Pronounced uncertainty prevents the solutions from being very robust. It is advisable to consider multiple scenarios of climate development and responses. As validation of distribution models for the future is difficult or impossible, a strategy to reduce sensitivity of the solution to uncertainty is highly recommended.

Link to tutorial

There is no tutorial example provided for this analysis. Please refer to Carroll *et al*. (2010) for an example.

5.3.8 Administrative units analysis

Planning problem: Setting conservation priorities over multiple administrative regions.

Conservation decisions are often taken at the regional level, but the range and population dynamics of species and other biodiversity features extend beyond regions. This analysis allows for balancing global and local conservation needs.

There are two variants of this analysis that emphasize global and local considerations in different ways.

Mode 1 (weak local administrative priorities) allows regional priorities (weights) for features to be given. While weights vary, representation is taken as global. This analysis is done in the same way as the basic Zonation analysis but with spatially variable weights.

Mode 2 (strong local administrative priorities) is fundamentally different in that it requires representation of all features in each area. There are different weights globally and locally, and representation is tracked separately both globally and in each subregion. Loss of conservation value is computed from 1+N components, where 1 is for global analysis and N is the number of subregions.

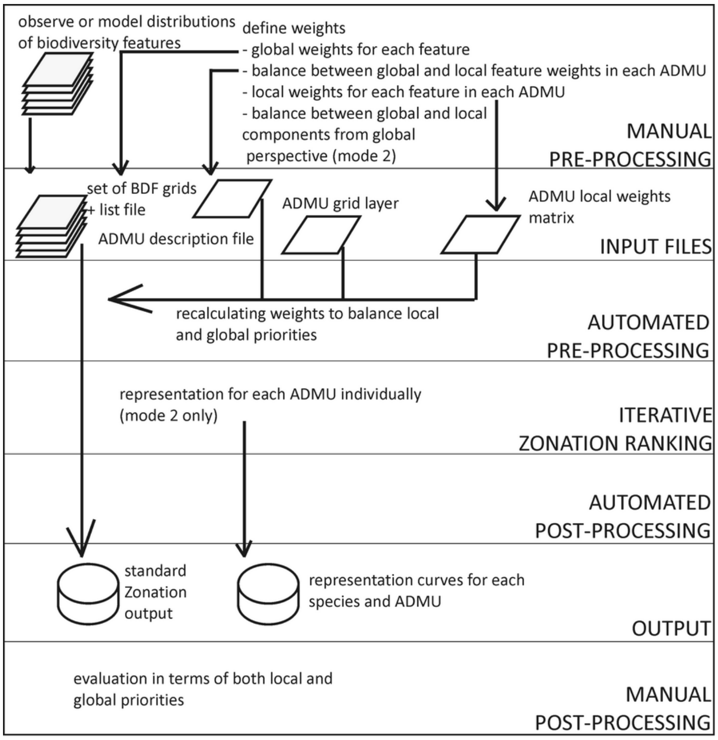
See section 2.12 for more information about the modes for an administrative units analysis.

Examples from literature

Moilanen, A. and A. Arponen. 2011. Administrative regions in conservation: Balancing local priorities with regional to global preferences in spatial planning. *Biological Conservation*, 144: 1719-1725.

Moilanen, A. Anderson, B.J., Arponen, A., Pouzols, F.M., Thomas, C.D. 2012. Edge artefacts and lost performance in national versus continental conservation priority area. *Diversity and Distributions*, 19: 171-183.

Process chart for the analysis



A process chart of an analysis that accounts for different priorities across several administrative regions. Please note that only the compulsory and most commonly used optional analysis components are presented. Different components can be combined according to the specific analysis.

Pre-processing of inputs

There are several weighting schemes that take place in the administrative units analysis. Prior to conducting an analysis, the following must be assigned:

1. Global feature weights for biodiversity features. Global feature weights go in the weights column in the biodiversity feature list file.
2. Region-specific weights and parameters. Regional weights (G_A) determine a region's weight with respect to the other regions. The balancing parameter $ß_A$ describes how local and global feature weights are balanced in each administrative region. These are entered in the second and third column of the administrative units descriptions file.
3. Region-specific local weights for each biodiversity feature. These weights reflect purely local conservation priorities and are compiled in a matrix where biodiversity features are in rows and administrative regions in columns. The order of the biodiversity features must be the same as in the biodiversity feature list file.

Input files

- A set of biodiversity feature grids (section 3.3.2.1)
- A biodiversity feature list file (section 3.3.2.2)
  - Assign the global weights for the biodiversity features in the first column of this file
- An administrative unit map grid (section 3.3.3.16)
- An administrative unit description file (section 3.3.3.16)
  - The region-specific weight should be in the second column of this file and the balancing parameter between local and global feature weights in the third column
- A features by administrative units weights matrix (section 3.3.3.16)
  - Defines local weights for biodiversity features in each administrative region
- A run settings file with appropriate settings (section 3.3.2.3)

Depending on the specific aims and details of the analysis, other input files such as uncertainty layers, a cost layer, landscape condition and retention layers, etc., can also be included.

Analysis stages and settings

For an administrative units analysis, include the following lines in the run settings file:

```
[Administrative units]
use ADMUs = 1
ADMU descriptions file = my_ADMU_descriptions.txt
ADMU layer file = ADMUs_distribution_map.asc
ADMU weight matrix = ADMU_weights_matrix.txt
calculate local weights from condition = 1
```

- Presently non-operational dummy parameter

`ADMU mode = 1 OR 2`

- Mode depends on whether the analysis should enforce local representation for all features (mode 2) or not (mode 1); section 2.12

`Mode 2 global weight = 0.5`

- A parameter specifying the balancing of global and local conservation value when using ADMU mode 2. This parameter is bounded between 0 (only local considerations) and 1 (only global considerations influence value)

Output and its interpretation

All basic Zonation output, the priority rank map, and the performance curves are produced. Several additional files are also generated by the administrative units analysis (see section 3.4.3):

- Rank raster file with values redistributed in the range [0, 1] within each administrative unit. In this rank raster file it is easy to identify a given top fraction for every administrative unit.
- Group-specific curves files for every administrative unit.
- A file that describes the occurrence levels of features in each administrative unit during landscape ranking (optional).

Strengths, weaknesses, and further considerations

Strength: This analysis can help especially in a regional conservation priority setting to account for the global conservation status of species or other biodiversity features. Alternatively, it can be utilized to identify regional conservation priorities that contribute to the systematic conservation over a larger region.

Weakness: As multiple weighting schemes take place simultaneously, the solution becomes sensitive to subjective weighting decisions. Avoiding this is difficult and it is not necessarily a problem, but it is important to keep in mind while interpreting the result. Also, the use of strong local priorities (mode 2) may result in a loss of global efficiency because features may become protected in regions where it is not (globally) efficient to do so (i.e., extra resources may need to be expended where a globally common feature is locally rare).

Link to tutorial

See tutorial Exercise 9 for an example.

5.3.9 Setup combinations

There are many different analysis setups, and the ones described in this section cover only a small fraction of what is possible. We encourage the creative use of Zonation analysis components, but keeping a few considerations in mind will help you develop successful and realistic analyses.

- Combine elements from different studies: species, communities, connectivity, costs, alternative land uses, condition, retention, etc. Choose the elements with respect to what kind of circumstances and limitations exist in your planning context.
- Develop complex analyses in stages. Start with the simplest analysis. Each time a new component is added to the analysis, verify that the changes to the solution make sense. This helps to trace back to data problems or bugs in computations.
- When there are multiple plausible analysis setups, a selection frequency analysis can be done across the priority rankings to find areas that (1) are always good, (2) are never good, or (3) are either good or not depending on the settings. Areas of class (1) are the most relevant for conservation.

Part VI

6 Tutorial & Examples

Development of Zonation setups

The tutorials here include individual sample setups that help you get familiar with Zonation. If planning a real Zonation case study, there are additional considerations of work flow that you should be aware of. There are two topics to pay attention to, but which are not illustrated in an ideal manner in the following setups: (i) running a Zonation project, and (ii) stepwise development of setups. Both of these topics are addressed in the following open online Zonation core reference:

Lehtomäki, J. and A. Moilanen. 2013. Methods and workflow for spatial conservation prioritization using Zonation. *Environmental Modelling & Software*, 47: 128-137.

Running a Zonation project

Running a Zonation project includes many things other than running a spatial prioritization. It may include development of objectives together with stakeholders; it includes acquisition, preparation, and verification of data; it involves preparation of analyses that investigate the problem under different sets of sensible assumptions (sensitivity analysis); it involves interpretation and communication of results. Zonation can also be used inside the broader operational model of systematic conservation planning (see work by R.L. Pressey, C. Margules, S. Sarkar, A. T. Knight, and many others). Often systematic conservation planning uses target-based analysis, but there is no reason why other types of analyses could not be used in it.

Because of the steps mentioned above, running a proper spatial prioritization project requires time, money, and people with skills in GIS, possibly habitat modeling, Zonation project leadership, the spatial analysis itself, etc. See Lehtomäki and Moilanen (2013) for more detailed discussion. Frequently data acquisition and preparation is the most time-consuming part of the project. Prepare accordingly.

Stepwise development of setups

Complicated analyses can include a variety of biodiversity data, connectivity, costs, and additional information about constraints, such as the present protected area network. It is important to realize that a complicated analysis should not be developed in one step. More sensibly, the analysis is developed in stages such as:

1. Simple analysis with limited data and no complicated factors (this is just to verify that mechanics of data preparation are working).
2. Add weights; for example, give higher weights to species that have high red list status.
3. Add more data, like data blocks for species, habitats, ecosystem services, and costs. Note that costs represent negatively weighted features.
4. Add simple connectivity considerations for species and habitats.
5. Add habitat condition layers, when relevant.

6. Add more complicated connectivity considerations, including for example connectivity interactions to an existing protected area network.
7. Add complicated considerations such as administrative units analysis, retention analysis, etc.
8. And so on. Some projects we have been involved in have developed the ultimate analysis in as many as 50 steps, which naturally takes time.

There are at least three reasons for developing a complex analysis in steps. First, it makes it more likely that problems with data or analysis settings are caught. If the solution does not change in a sensible manner between steps, it pays to check the additional data and setup features. Second, developing the ultimate analysis in steps gives improved appreciation about the relative effects of different data blocks and assumptions. Third, stepwise development and verification of analyses is a way of engaging stakeholders and experts, and facilitating their appreciation of the planning process. Thus, with stepwise planning overall confidence in the analysis is improved and interpretation is helped. In contrast, if a complicated analysis is implemented in one go, catching faults in data preparation is difficult (because the result integrates so many factors), and overall understanding of the prioritization remains poor.

The present tutorial

The purpose of this tutorial is to illustrate the use and function of different Zonation analyses. They also help you to familiarize yourself with the program and its settings. Later on, when working with your own data, you can use the example files to create your own input files.

The installation package includes the following example files (note that the extension of the raster map files may vary depending on the GIS format used):

- For running basic Zonation
  - species1.asc
  - species2.asc
  - species3.asc
  - species4.asc
  - species5.asc
  - species6.asc
  - species7.asc
  - splist.spp
  - splist_abf.spp
  - splist_tbp.spp
  - cost.asc
  - set.dat
  - set_cost.dat
  - do_zig3.bat
  - do_load.bat
- For weighting species
  - splist_w.spp
  - splist_w_alu.spp
  - do_w.bat
  - do_w_alu.bat
- For using SSI species
  - SSI_sp8.txt
  - SSI_sp9.txt
  - SSI_list.txt
  - set_ssi.dat
  - do_ssi.bat
- Distribution smoothing
  - do_ds.bat
- For including BLP
  - set_blp.dat
  - do_blp.bat
- For including BQP
  - BQPcurves.txt
  - splist_bqp.spp
  - set_bqp.dat
  - do_bqp.bat
- Uncertainty analysis
  - sp1_UC.asc
  - sp2_UC.asc
  - sp3_UC.asc
  - sp4_UC.asc
  - sp5_UC.asc
  - sp6_UC.asc
  - sp7_UC.asc
  - UCweights.spp
  - set_uc.dat
  - do_uc.bat

For running the replacement cost analysis
- mask_rs.asc
- mask_towns_z3.asc
- set_maski.dat
- set_maske_z3.dat
- do_rs.bat
- do_towns.bat

For including directed connectivity (NQP)
- plu.asc
- rivers.txt
- NQPcurves.txt
- splist_nqp.spp
- set_nqp.dat
- do_nqp.bat

For community level analyses
- community1.asc
- community2.asc
- community3.asc
- community_similarity.txt
- groups_community_sp.txt
- bdlist_community.spp
- sp_community_list.spp
- set_community_similarity.dat
- set_community_sp.dat
- do_comm_nosimilarity.bat
- do_comm_similarity.bat
- do_comm_sp.bat

For considering administrative units
- splist_ADMU.spp
- ADMU_map.asc
- ADMU_descriptions.txt
- ADMU_weights.txt
- set_ADMU.dat
- do_ADMU.bat

For condition and retention analyses
- condition1.asc
- condition2.asc
- cond_list.txt
- groups_condition.txt
- set_condition.dat
- do_condition.bat
- bdlist_retention.spp
- retention1.asc
- ret_list.txt
- groups_retention_m1.txt
- set_retention_m1.dat
- do_retention_m1.bat

Other files needed for tutorial and exercises
- set_costds.dat
- set_costbqp.dat
- load_costds.bat
- load_costbqp.bat
- do_cost_ds.bat

As a final word, this tutorial does not include all variants of everything. When working with your own data, remember to carefully consider the cell removal rule you use. Analysis options that you might wish to modify include edge removal, the warp factor, and adding of fake edge points.

6.1 Exercise 1

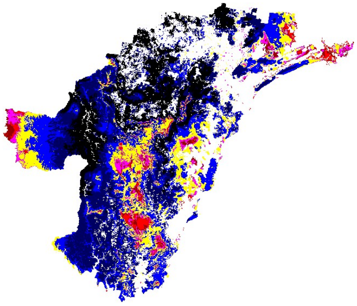
Getting started with the basic Zonation

Before starting see section 3.2 for how to operate the program either from the command prompt or the Zonation GUI. Also, please refer to the information at the beginning of the tutorial section as there is relevant information about the development of Zonation setups and about the running of a Zonation project.

We start with a simple exercise conducting a basic Zonation analysis. Assume that there is an area in a remote country which is the home of seven rare species. We have been given the task to create a conservation network that will help to protect them. However, due to cost restraints the proposed conservation areas cannot be larger than 15% of the landscape. We decide to use core-area Zonation to identify areas that have high priority for conservation as this method variant corresponds best to our planning objectives (see section 2.3).

The first species list file (**splist.spp**) contains a list of seven species distribution maps and species-specific parameters. Here all species have been given an equal weight (column 1), but they have different dispersal α values (column 2) as the use of surrounding landscape (e.g., home ranges) differs between the species. Because no other features are used in this first exercise, the last three parameters are given a dummy value of 1. All the necessary parameters to run the basic core-area Zonation are defined in the first run settings file (**set.dat**). The edge removal option has been selected, but the analysis can also be run without this option to see the effect it has on the result. The warp factor (here 100) has been adjusted so that the analysis runs smoothly but with fair accuracy. Test different warp factors to see how much of a difference it makes in the result and the running time. In our experience the difference in running time is substantial, but the effects on accuracy are from minute to non-existent. The batch file (**do_zig3.bat**) can be run from Windows by double clicking on the file name. Remember that the batch file is a text file that can be edited using notepad. If the Zonation GUI is used, open the Zonation project from the Project View window and queue it to run. Please see the section about the Zonation GUI for detailed instructions. If the "classic Zonation" color scale is selected, the outcome of this analysis should look like this:

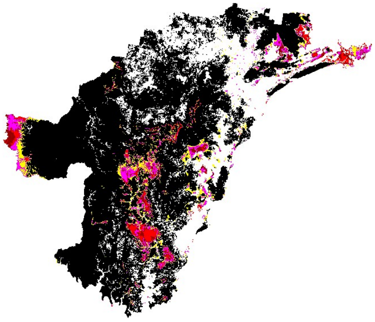
Batch file: do_zig3.bat



Output map of Exercise 1 using the 'Classic Zonation' color scale.

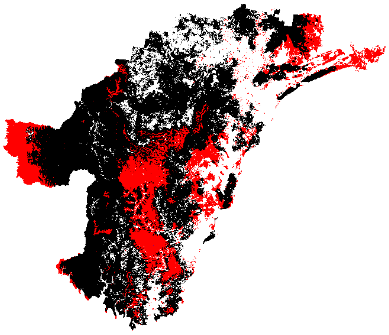
In the map above, the biologically most valuable areas are shown in red and least valuable are shown in black. Since we are limited in the amount of area we can protect, we now need to find the areas that represent the most valuable 15% of the landscape. In the Zonation GUI, these areas can be identified by adjusting the color scale settings. From the classic Zonation color scale (used above), right click on the yellow color tab representing 75% and adjust so that the percent value is 85. You can then change the color of the less valuable areas to black so that the top 15% stand out (as in the image below, see GUI instructions for further explanation). Another way to highlight the top 15% of the landscape from the GUI is to first select a binary scale and then adjust the color tab to 85. This map will contain only two colors: one for the top 15% and one for the remaining landscape fraction. Zonation automatically saves the original output map of the analysis (above) as a raster file that can be loaded into any GIS program for further processing. Any image can be saved as a graphics file by right-clicking on the image. Please see descriptions of the .rank.asc and .prop.asc files.

Batch file: do_zig3.bat
Top 15% areas:



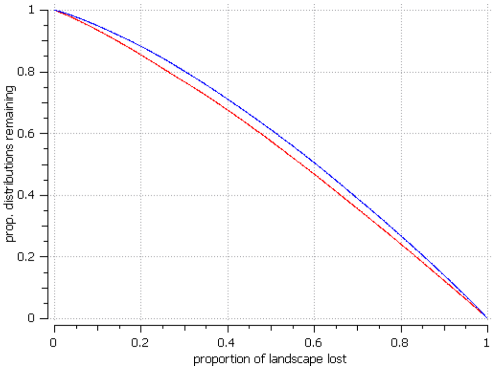
Output map of Exercise 1 showing only the top 15% of the landscape.

The Zonation GUI can also identify areas that include a specific percentage of the distributions of all species. This corresponds to a target-based proportional coverage solution and can be visualized for any method of aggregating conservation value, ABF, CAZ, TBF or GBF. To do this, first open the "proportion lost" (.prop.asc) file. This can be found in the file hierarchy in the Project View under Output files-->Proportional loss layers. Alternatively, the file can be opened by right clicking on the output file in the Process View window and selecting 'Open Proportional Loss Raster.' Use the color sliders to show the top areas that include the desired proportion for each species. For example, to identify areas that include at least 30% of each species, adjust a color tab so that it will have percent value 70.



Example showing the areas that contain at least 30% of all species.

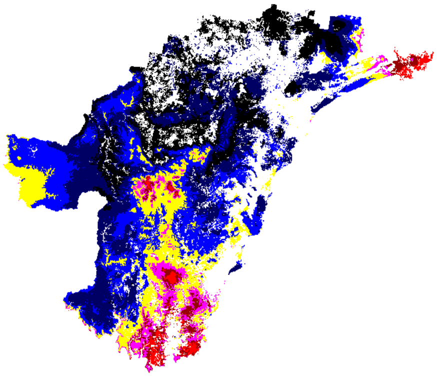
In addition to the cell ranking, Zonation produces a set a curves describing the absolute performance levels of species. The figure below shows the minimum (red line) and average (blue line) performance across all species for the basic Zonation analysis in this exercise. Please note that the minimum performance curve is linked to the proportional loss file. From these curves we can see that a fraction of the species' distributions is lost when only a small fraction of the landscape has been removed. This indicates that the species in this exercise are relatively widespread and have mostly non-overlapping distributions. These curves are shown in the Runtime plot tab of the Visual Output in the Zonation GUI, and the data for generating these curves, as well as the species-specific curves, is saved in a text file in a **.curves.txt** file.



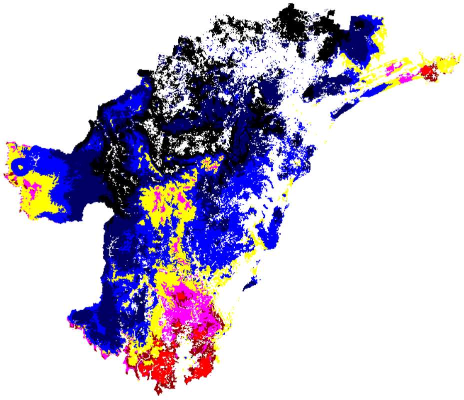
Performance curves from a basic Zonation analysis. The curves show that the average (blue line) and minimum (red line) proportions of species distributions are declining as landscape is removed.

The above example and the remaining examples use Core-area Zonation as the cell removal rule. However, an example of a basic Zonation analysis using the additive benefit function (**do_abf.bat**) and target based planning (**do_tbp.bat**) are included as part of the tutorial package. Compare the images below with the core-area Zonation analysis (above) too see how the cell removal rules differ.

Additive Benefit Function:



Target Based Planning:



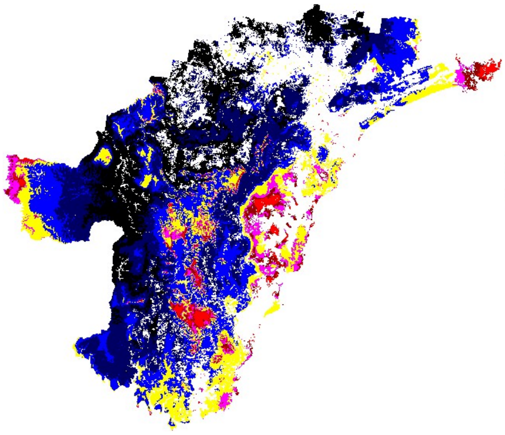
6.2 Exercise 2

Weighting of species or other biodiversity features

The default in Zonation is that all features are weighted equally with weights of one. It is encouraging that the analysis frequently works quite acceptably even with this default weighting. The reason for this is that the prioritization is driven by principles that make general sense as they are. Nevertheless, weights of features are frequently varied, to account for factors such as red list status, economical value, taxonomic uniqueness, etc. (see Lehtomäki and Moilanen 2013). Changing the weight of a feature changes its performance through prioritization perhaps less than one might expect. For example, doubling the weight of a species from 1 to 2 sounds like a lot, but if the analysis has another 1000 features with an aggregate weight of 1000, it is obvious that the the single feature does not start to dominate the priorities. Its weight still is only 1/500 of the total aggregate weight allocated to all features.

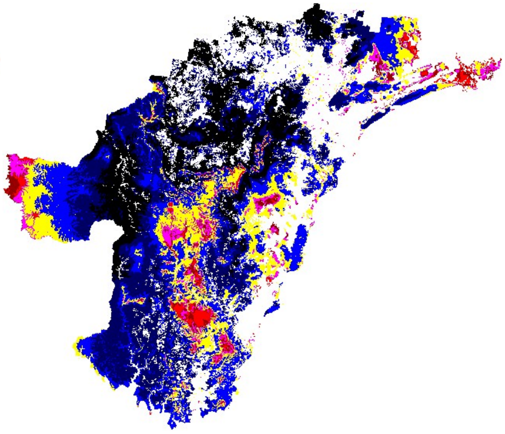
In this exercise, assume that two of the target species (species 2 and 3) are endemic and can not be found anywhere else in the world. We can enhance the protection of these species by giving each a weight of 2. Remember that the weights are defined in the species list file, so in this exercise we will use a different file (**splist_w.spp**) where the weights have been changed. Alternatively, the same species list file used in Exercise 1 can be used here by changing the weights manually. With this option, please remember to also change the name of the output file so that the program will not overwrite the solution from Exercise 1. The batch file used in this exercise (**do_w.bat**) already gives the output files a different name from Exercise 1 (**output_w.txt**). The settings file used in this exercise is the same as in the previous exercise. Use the Zonation GUI to open the batch file or call the program from the command line as in Exercise 1.

Batch file: do_w.bat



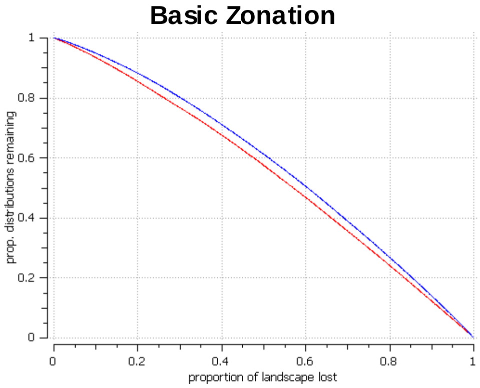
Output map with species 2 and 3 having a weight of 2.

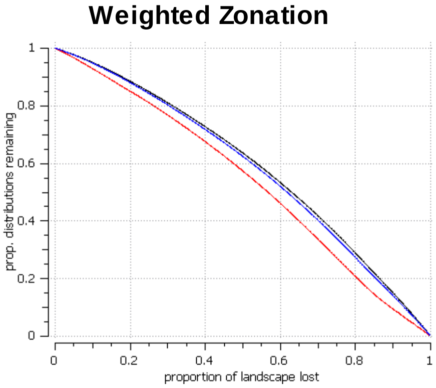
Batch file: do_zig3.bat



Output map where all species have an equal weight (from Exercise 1).

The images above show that the weighting of species alters the spatial distribution of the highest value cells. In this exercise (left map) more importance has been given to the areas where the two endemic species occur (west and south coast) and less to those areas that have a high representation for the other species (the peninsula on the east and the north-east and central regions). A difference between the basic and weighted Zonation runs can be seen also in the species distributions curves in the Runtime plot of the visual output window.





The graphs above show the proportion of species' distributions retained during cell removal. Comparing the two graphs shows that the relationship between landscape loss and **average** biodiversity protection for all species is quite similar in both cases (blue line). However, when species 2 and 3 are weighted, they retain a relatively higher proportion of their distributions through the cell removal process compared to the basic Zonation run where no weights are used (black line). In turn the minimum proportion retained (red line, species with the lowest distribution remaining) is smaller when using species weighting, meaning that some species lose more of their distribution in a weighted analysis when compared to an analysis where all species are weighted equally.

Solution comparison (see section 3.5.1.3) is a useful feature that can be used to visually compare two solutions. The figure below shows the overlap and differences in the best 15% of landscape between the basic and weighted Zonation runs (Exercises 1 and 2). Areas that appear in the top 15% in both solutions appear in yellow, areas that are included in the top 15% in the non-weighted solution only are shown in dark green, and areas that appear in the weighted solution only are shown in light green. Please see section 4.3.4 for information on conducting solution comparison in the Zonation GUI.

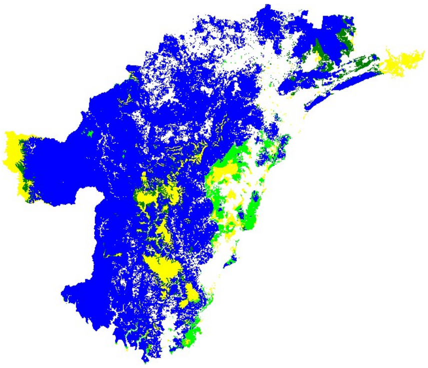


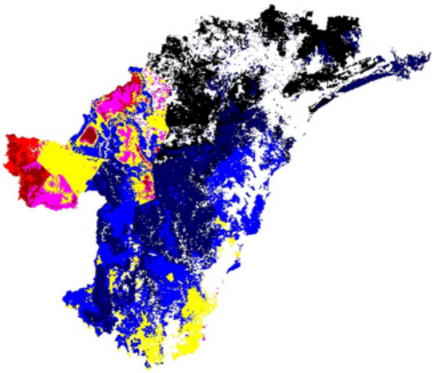
Figure showing differences in the top 15% fraction between the non-weighted (dark green) and weighted (light green) solution. Overlapping areas are shown in yellow.

Negative weights: Accounting for alternative land uses or negative effects to biodiversity

In Zonation v.3.1 and higher, it is possible to include features that have a negative effect for conservation value. These can be areas that have high priority for urban construction, farmland, or other purposes. The analysis is implemented by treating the alternative land use layer as any biodiversity feature layer but assigning it a negative weight in the biodiversity feature list file.

For example, assume that the values in file **plu.asc** represent the suitability of the grid cells for the expansion of urban development. The most suitable areas are likely to be zoned for urban construction in the near future, so we do not want to give those areas too much weight to alleviate the conflicting interests. We assign the alternative land use layer a weight of -2.0 in a new species list file (**splist_w_alu.spp**). Otherwise, this species list file is weighted the same as the previous weighted Zonation analysis example. Run the analysis with the same settings file as before and using the **do_w_alu.bat** batch file.

Batch file: do_w_alu.bat



Ranking output map when an alternative landuse layer is included in the analysis.

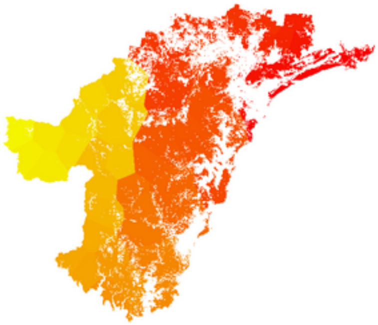


Figure showing suitability of expansion of urban development. Suitability increases with the intensity of red color.

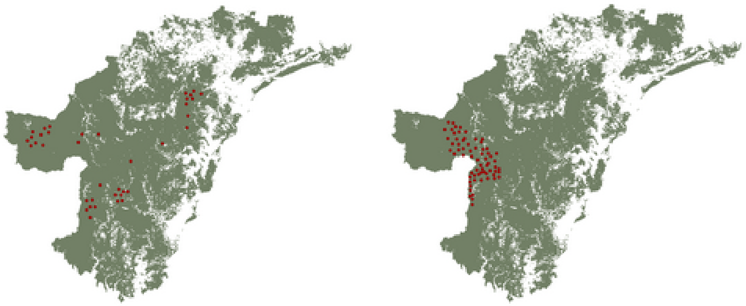
Compare the ranking map from this analysis to the one without the alternative land use layer. Sites that would have been high priorities for biodiversity are now removed from the solution early as they were assigned high values of a negative feature. Look at the representation curves in the **.curves.txt** file or in the Runtime of the Zonation GUI to see how the negative weights affected species representation. When looking at the minimum representation curve, it is worthwhile to remember that it is the negative feature and not an actual biodiversity feature layer.

6.3 Exercise 3

Species of special interest

Note: analysis options and output for SSI features are limited, and unless there are computer memory considerations, it is advisable to enter species (feature) distributions as standard grids. For example, connectivity computations do not operate on SSI features.

Assume that in addition to the seven species considered so far, there are two additional species that should be protected. However, these species are very rare and poorly known, and a set of occurrence points indicating where the species have been observed is the only information available. Unfortunately this data is not sufficient to model the species distribution to cover the entire study area. The species can be included in the analysis by using the point occurrence data to enter them as species of special interest (SSI). Remember that an SSI file is a text file containing the exact coordinates of the records for one species (files **SSI_sp8.txt** and **SSI_sp9.txt** in this exercise). The species-specific parameters are defined in an additional SSI species list file (**SSI_list.txt**).



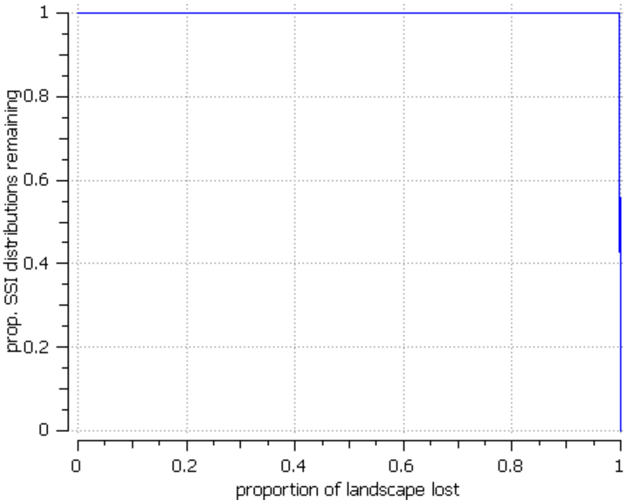
i) SSI species 1

ii) SSI species 2

Maps showing the point occurrence data of the two SSI species.

The analysis is conducted with the same original species data used in Exercise 2. However, to include the SSI, the run settings file must be modified to activate the SSI species option and to include the name of the SSI species list file. This can be done manually by editing the run settings text file (from the GUI or notepad) or the program can be run using the modified run settings file (**set_SSI.dat**) provided with this tutorial. Use the **do_ssi.bat** file to run the analysis from either the command line or the GUI.

The results of this analysis show that the inclusion of the two extra species does not change the results much. This is because the occurrence points cover a very small fraction of the study area, and it is easy for the program to include them in the top fraction without altering the spatial distribution of high value cells in the solution. Also, because the area covered by the points is so small, the cells with SSI species receive very high values. Thus it follows that the full distribution (i.e., all the points) of SSI species are almost always included in the top fraction. This is evident also from the SSI species graph in the Runtime plot window (below) where we can see that the entire distribution for both SSI species are retained until the very end of the cell removal process.



Graph of SSI species' landscape retention curves in the Runtime plots window.

Note that the same information displayed in the graph also appears in a **.SSI_curves.txt** file that the program automatically produces during the analysis.

As the inclusion of the two SSI species contributes very little to the final result, they will be kept separate from the main analysis and only the original seven species will be considered in the remaining exercises.

6.4 Exercise 4

Adding aggregation into the analyses

The previous exercises have identified sites that have a high occurrence of the target species by giving a higher weight to the two endemic species. However, the high occurrence areas are quite fragmented from each other. This is a poor quality in a reserve network, but there are ways to correct this and induce aggregation in the solution.

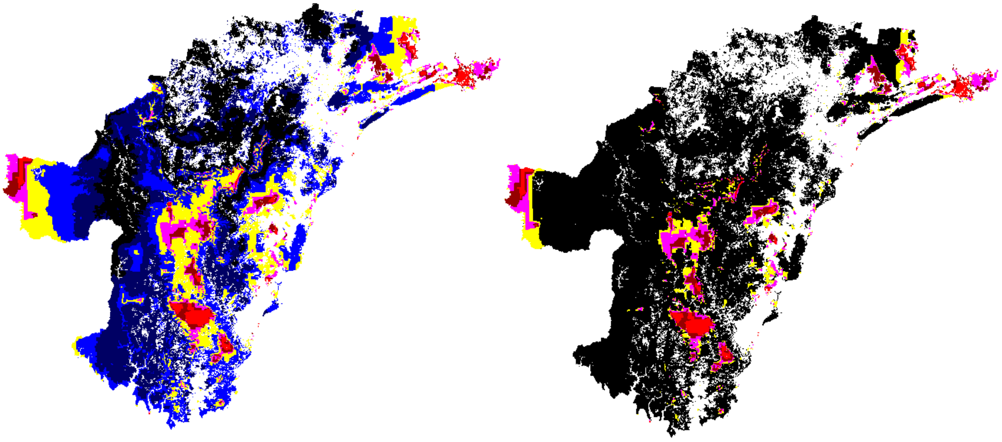
Three such methods are used in this analaysis: **boundary length penalty, distribution smoothing,** and **boundary quality penalty**. These methods favor the selection of contiguous cell groups, rather than selecting more fragmented sets of cells. This in turn offers advantages in terms of greater connectivity and can also promote more practical and cost-effective management. Note that it is not recommended to use several aggregation methods simultaneously due to difficulties in interpreting the results.

Boundary Length Penalty

Boundary length penalty (BLP) is a non-species-specific method to induce aggregation in a reserve network. The order of cell removal is based on the species occurrence level and the increase/decrease of boundary length resulting from cell removal. This analysis minimizes the boundary length, and aggregation is induced via a penalty for the boundary length of the reserve. For more information about BLP see section 2.4.1. Note that while BLP is a common and quick way to induce aggregation in a reserve network, it may not be the most biologically realistic method as it does not account for species-specific responses to fragmentation.

Use the batch file **do_blp.bat** to run this analysis. This analysis uses the unweighted species list file. In the run settings file (**set_blp.dat**) the BLP value is assigned as 0.05. See section 5.1.2 for more information about setting up an analysis using BLP.

Batch file: do_blp.bat

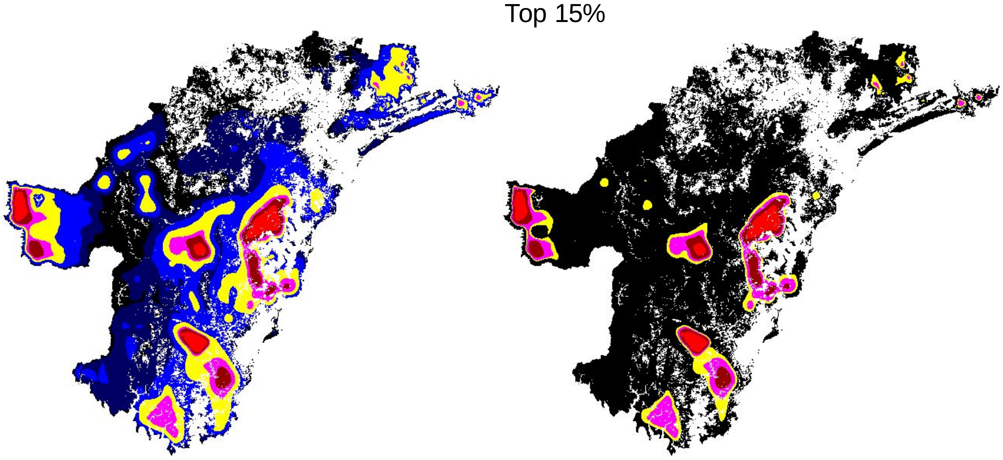


Top 15%

Distribution smoothing

Distribution smoothing results in more aggregated solutions based on the connectivity of sites. See section 2.4.2 for detailed information about distribution smoothing. This analysis is conducted with the same species input files as in Exercise 2, but a new batch file (**do_ds.bat**) is used to call the program. In the new batch file, the third parameter from the end (value = 1) indicates that distribution smoothing will be used. The second column in the species list file defines the species-specific smoothing kernel.

Batch file: do_ds.bat

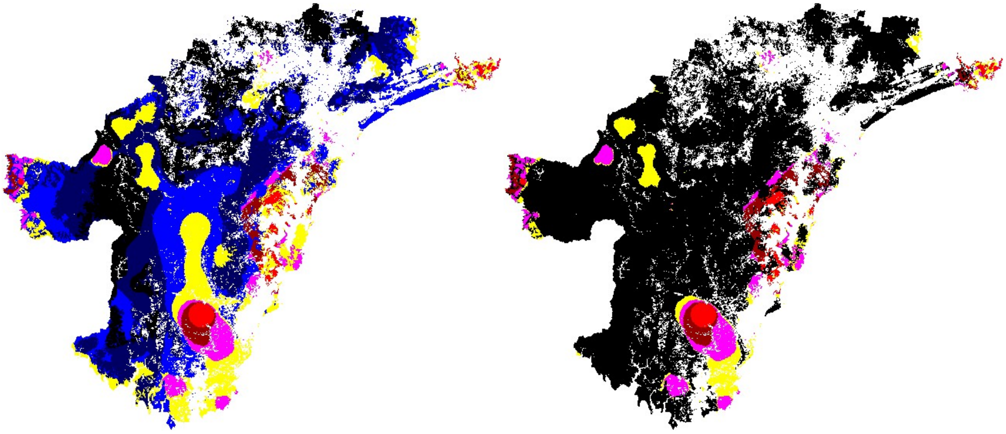


Boundary quality penalty

The boundary quality penalty (BQP) calculates the most valuable sites based on both the value of the cell and the effects of habitat loss in the surrounding cells. The effects of fragmentation (loss of neighborhood cells) are species-specific and thus BQP analysis also takes into account how different species are influenced by fragmentation and habitat loss. See section 2.4.3 for more information about BQP analysis.

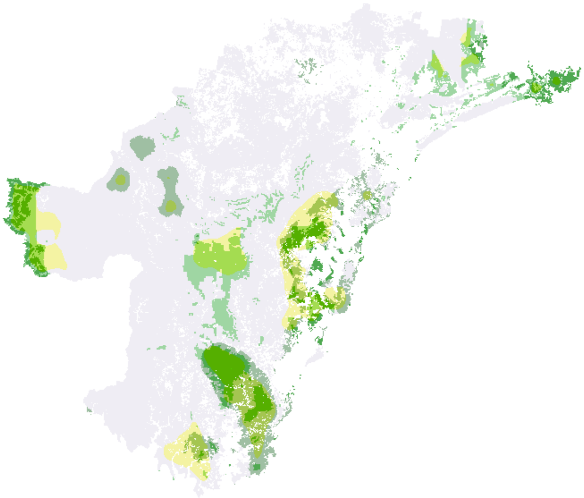
BQP analysis requires prolonged computation times compared to distribution smoothing because the loss of a cell effects occurrence levels in nearby cells and must be taken into account in calculations. To run the BQP analysis in this exercise, use the batch file **do_bqp.bat** to load the project to the Zonation GUI or to call the program yourself. This batch file calls the run settings file **set_bqp.dat** which selects the BQP option (value = 1) and sets the warp factor to 500. For this exercise, the warp factor has been increased to 500 because BQP substantially increases computation time, but a warp factor of 100 or lower is recommended when conducting an analysis for real planning purposes. The species list file **splist_bqp.spp** included in this project file defines all species-specific responses to fragmentation (columns 3 and 4, see section 3.3.2.2 for more information about these parameters).

Batch file: do_bqp.bat



Top 15%

It is clear from the images above that each method produces a quite different solution. The solution comparison feature can be used to see where the main differences are. The figure below shows the comparison between the three solutions using the highest ranked 15% of each solution. In this figure, the light green areas indicate BLP results, blue-green shows BQP results, and yellow shows distribution smoothing results. Light yellow-green areas indicate locations where the distribution smoothing and BLP results overlap. Dark yellow-green areas show where distribution smoothing and BQP results overlap. There are few areas where results from all analyses overlap. The most prominent are the fragmented dark green patches in the west and the dark green patches in the southeast.



Merged map showing the overlap in the 15% highest ranked areas of BLP, BQP, and distribution smoothing analyses.

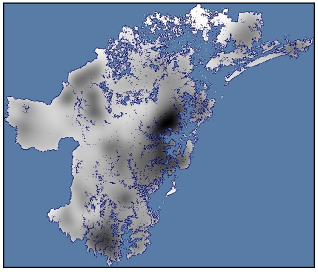
6.5 Exercise 5

Uncertainty in species information

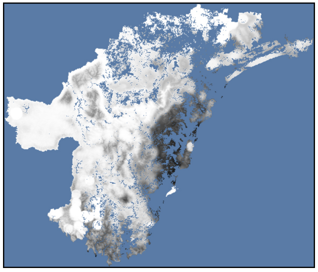
Due to lack of data, uncertainties in species distributions are a common problem encountered in ecological studies. With Zonation, these uncertainties can be accounted for in reserve selection. Uncertainty analysis is used to focus on sites where the uncertainty is low compared to the predicted representation levels, thus prioritizing sites that have both high abundance and low uncertainty.

The simplest way of accounting for an uncertain feature is to reduce the weight given to it (or conversely, increase the weight given to a feature with very reliable distribution data). However, in this exercise, assume that the species distribution data comes from statistical species distribution models. Models contain many uncertainties in their predictions which can be taken into account when selecting the best sites to be protected. To conduct an analysis that accounts for uncertainty, an uncertainty layer for each species (**sp1_UC.asc**, **sp2_UC.asc**, etc.) showing the spatial distribution of uncertainty in the data must be included. The pair of figures below shows an example of the distribution (left) and the error surface (right) for species 2. The figure on the left shows the distribution layer with white indicating areas with low occupancy and black indicating areas with high occupancy. In the uncertainty layer on the right, white indicates areas with low error and black indicates areas with high error. The uncertainty layers can be viewed in the Zonation GUI by selecting the map layer from the tree in the Project view.

Distribution layer



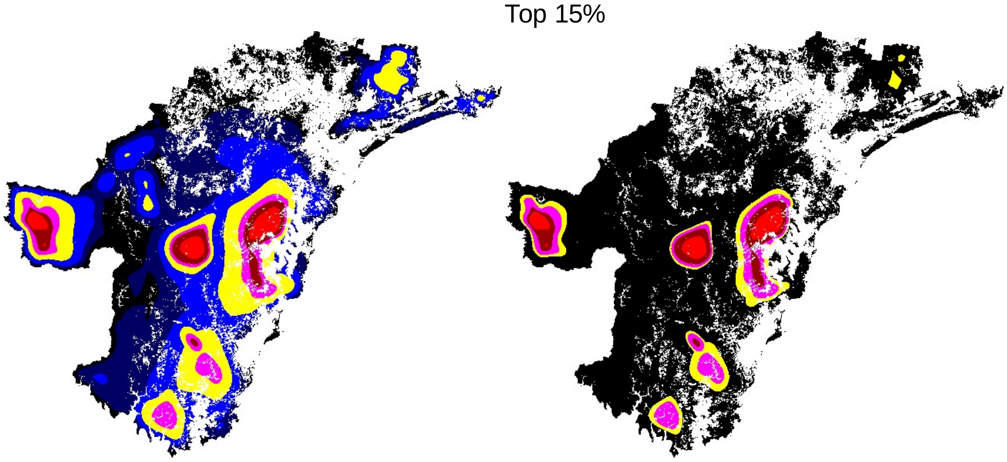
Uncertainty layer



In this exercise, uncertainty analysis is done together with distribution smoothing. While running the analysis, the program first calculates a new value for each cell based on the uncertainties in the data. The program then uses these values in the aggregation analysis. To run this analysis, use the batch file **do_uc.bat** to load the project in the Zonation GUI or to call the program from the command prompt. This batch file calls the run settings file **set_uc.dat** and the same species list file that was used for distribution smoothing in Exercise 4. In the run settings file, the parameter for `use info-gap weights` has been changed to 1, indicating that an uncertainty analysis will occur. Listing the file **UCweights.spp** as the `Info-gap weights file` name calls the uncertainty species list file , which lists the species-specific uncertainty weights and map layers used in the analysis. In this exercise, each species has been given a weight of 1 in the uncertainty species list file, stressing the certainty of data equally for all species.

For more information about the theory and algorithms associated with uncertainty analysis, see section 2.5. For more information about how the files and settings needed to run an uncertainty analysis, see section 3.3.3.7.

Batch file: do_uc.bat



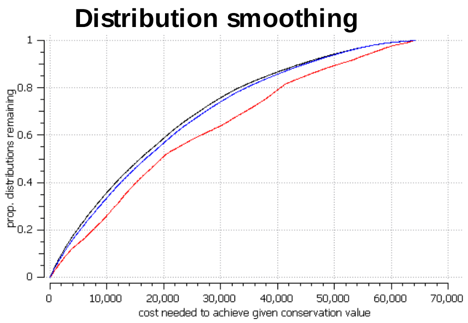
6.6 Exercise 6

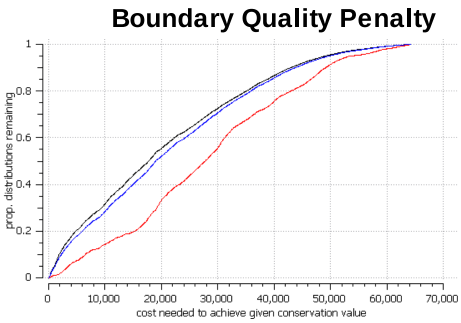
What about the costs?

There are two main ways of including cost in Zonation: (i) use of a cost layer and (ii) using negatively weighted features, which represent multiple costs, alternative land uses, or other such considerations. This example is primarily about the use of a cost layer. The other technique, negatively weighted features, is easy to use and convenient, and can account for multiple cost-like elements simultaneously.

As in many cases in conservation biology, costs play a vital role in reserve planning. The previous exercises have prioritized sites based solely on their biological value and have ignored the possible costs (except in terms of land area constraints). This exercise illustrates two ways to use a cost layer to determine how much it will cost (in terms of land cost) to protect the proposed areas. Assume that land in the study area is the most expensive in the south-west and cheapest in the north-east region.

The first method, solution cross comparison using solution loading (section 3.5.2), can be used to evaluate an old solution using a new set of criteria. In this exercise, the solutions calculated in Exercise 4 (aggregation with distribution smoothing and boundary quality penalty) are loaded with a cost layer (**cost.asc**). The program then calculates the increase in cost as cells are removed in the exact same order as they were in the solutions from Exercise 4. Use batch files **load_costds.bat** and **load_costbqp.bat** to run this analysis from the Zonation GUI or the command prompt. The cost curves (below) show the differences in land cost between the two solutions.

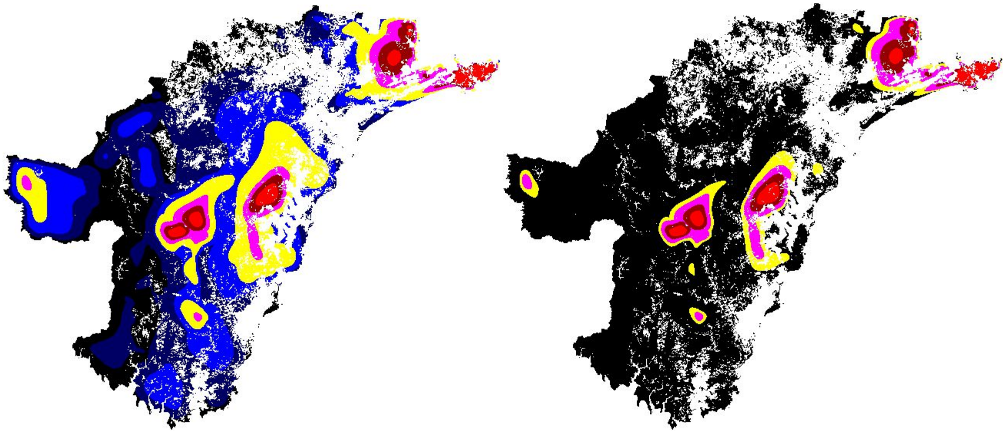

Distribution smoothing


Boundary Quality Penalty

An alternative method is to include cost during the cell removal process. This method produces a solution that has both high conservation value and low demands for resources by selecting cells that have a high conservation value/cost ratio. This exercise uses the same files as in distribution smoothing in Exercise 4 with the cost layer (**cost.asc**) included. Update the run settings file to include cost (`use cost = 1`) and define a cost layer file (`cost file = tutorial_input/cost.asc`). Alternatively, load the batch file **do_cost_ds.bat** in the GUI or call the program from the command prompt. This batch file calls the run settings file **set_costds.dat** where the necessary settings for this analysis are defined. See section 3.3.3.6 for more information about settings and files needed when including a cost analysis.

The images below show the output maps obtained from this analysis. How do the most important areas change compared to the solution from Exercise 4? Compare the species distribution curves of the two solutions. What are the differences between the two solutions?

Batch file: do_cost_ds.bat



Top 15%

6.7 Exercise 7

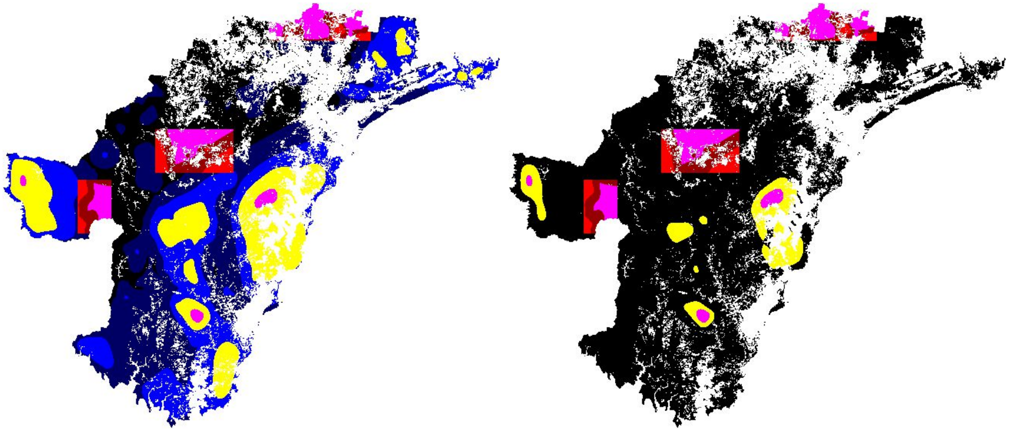
What about the already existing reserves?

The technique described in this section is among the most useful and commonly used ones in Zonation. When it comes to reserve network planning, land managers seldom have the chance to start from an area where no land management has taken place. In many cases the target area already includes reserves or areas that are ear-marked for other land uses such as agriculture, forestry, development, etc. Thus, an analysis often has to take into consideration areas that either can not be included or must be included in the reserve network.

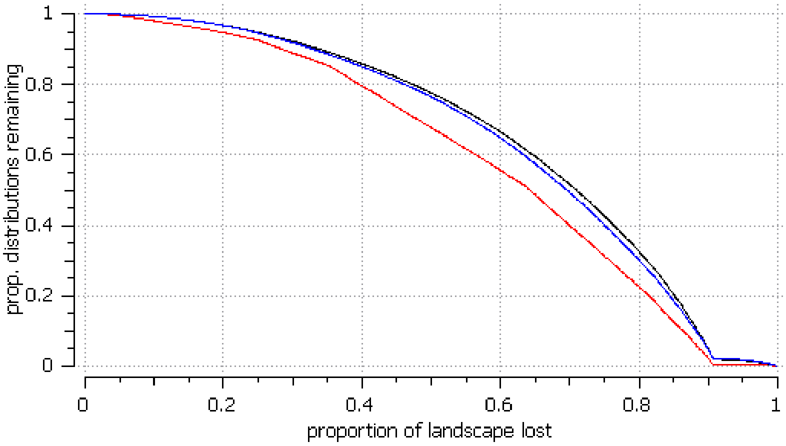
In this exercise, assume that there are three reserves already present in the landscape and that these reserves must be included in the final solution. This analysis uses the same species list files as before. A removal mask layer (**mask_rs.asc**) is included to classify cells as either reserve (indicated with 1) or non-reserve (indicated with 0). This ensures that the existing reserve areas are removed last. The settings for this exercise are defined in the run settings file **set_maski.dat**. Call the program with the batch file **do_rs.bat**. The images below show the ranked maps from this analysis. Please see section 3.3.3.9 for more information about how to set up an analysis using the removal mask layer.

Batch file: do_rs.bat

Top 15%



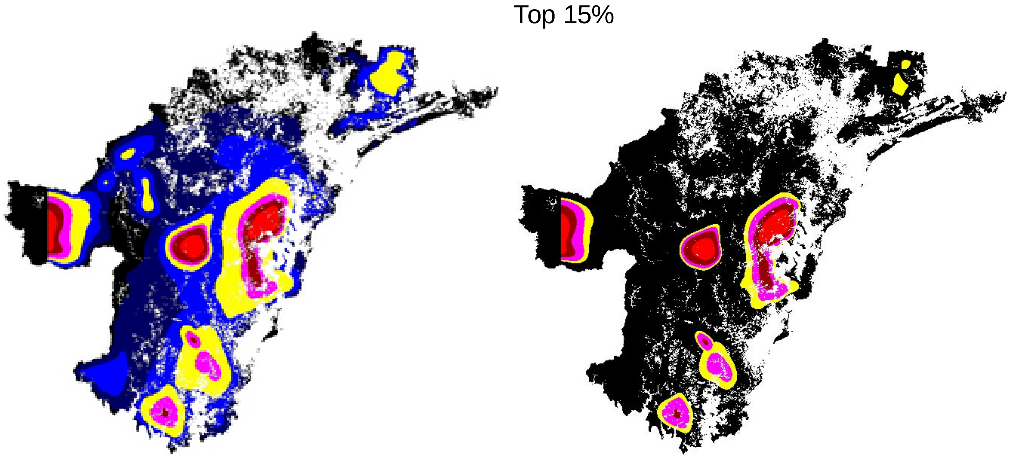
Compare the solution to the results from Exercise 5. How has the representation of species altered now that the old reserves are included in the top 15% fraction? Note the differences at the end of the species distribution curves (below). In this exercise, including the existing reserves results in a suboptimal solution since the program is forced to include poor quality areas in the top fraction.



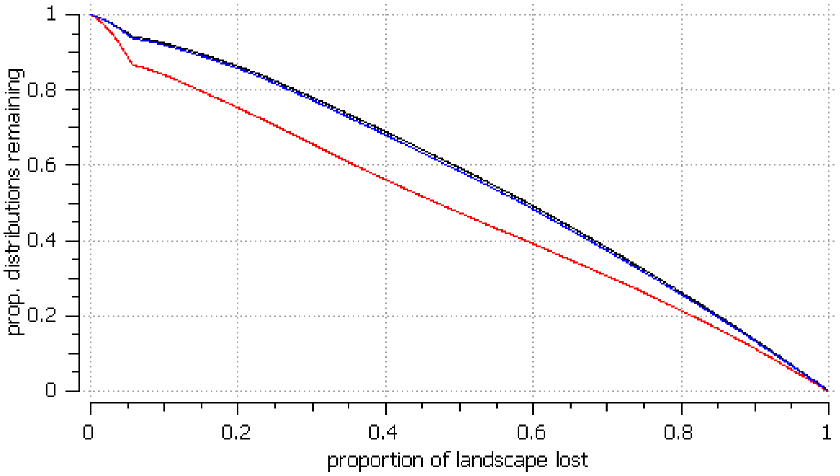
Average and minimum performances when old reserves are included in the top fraction.

The removal mask layer can also be used to exclude undesirable areas such as towns, industrial areas, mining sites, etc. from the reserve network. Cells earmarked for a land use other than conservation are assigned a 0 in the removal mask file while all other cells are assigned a 1. This will ensure that these areas are removed from the landscape first before prioritizing the rest of the area. In this exercise, assume that two residential areas are present in the study area and must be excluded from the reserve network. These areas are defined in the mask layer **mask_towns_z3.asc**. Call the program with the batch file **do_towns.bat**. The run settings are defined in the **set_maske_z3.dat** file. All input files other than the mask layers are the same as in Exercise 5.

Batch file: do_towns.bat



In this analysis, the biggest changes are seen in the beginning of the species distribution curves (below). Since the two excluded areas have high biological value, the proportions of species distributions decrease more steeply in the beginning when these areas are removed. However, since other valuable areas can be included in the solution, this masking has less effect on the final 15% top fraction compared to the previous example when existing reserves are included, as the average and minimum proportions of species distributions curves show.



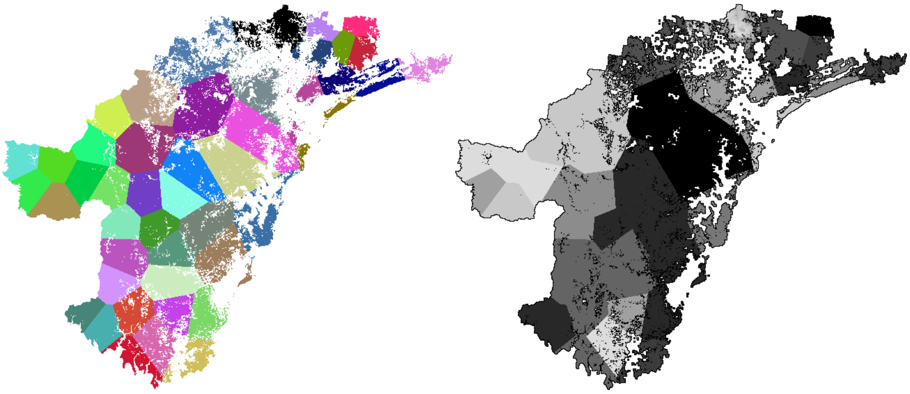
Average and minimum performances when residential areas are excluded from the top fraction.

6.8 Exercise 8

Working with directed connectivity

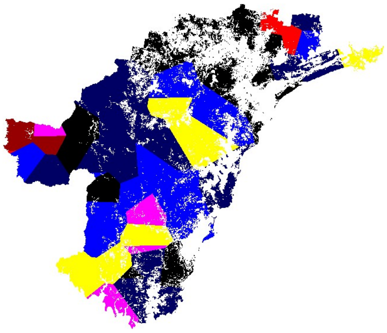
This exercise introduces the directed connectivity feature Neighborhood Quality Penalty (NQP). This is a variant of BQP, and the basic principles for landscape ranking are the same. The essential difference is that connectivity of sites in NQP is directed and defined as a set of linked planning units rather than a buffer around the focal cell as in BQP. Note that setting up the files required by the NQP can be complicated if there are many thousands of catchments: there may be special conditions in the description of the river network that need to be resolved manually, which may make data preparation time consuming.

To illustrate NQP, assume that the seven rare species used in previous exercises are freshwater species present in the numerous rivers in the study area.

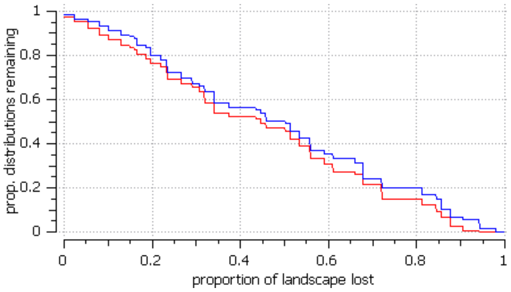


Picture of the study area divided into planning units (color) and river basins (gray scale).

This analysis uses the original, unweighted species list file **splist.spp**. Because NQP uses planning units rather than grid cells, a planing unit layer, **plu_file.asc**, is also included. Planning units can be defined by any criteria, but in this analysis they describe the smaller water catchment areas of riverine systems. Each planning unit is linked to the next downstream unit by direction of water flow, as defined in the file **rivers.txt**. To account for connectivity in NQP, both upstream and downstream connectivity responses must be considered. These responses are defined by two penalty curves in the **NQPcurves.txt** file: one curve defines upstream connectivity responses (curve 1) and one curve defines downstream connectivity (curve 2). Settings for running this analysis are defined in **set_nqp.dat**. Use the batch file **do_nqp.bat** to call the program. For more information about the theory and how to set-up an NQP analysis, see sections 2.4.4. and 3.3.3.3. For more information about the set-up of a planning unit layer, see section 3.3.3.11.



Priority ranking of an analysis using planning units.



Performance curves from a planning units analysis.

The use of planning units results in a considerably different priority ranking (shown above) compared to when planning units have not been used. Using planning units also changes the species performance curves from smooth to staggered lines. Compared to the basic analysis in Exercise 1, all species are performing notably worse in this analysis. This is due to the fact that the planning units used here are relatively large, and as the program removes units from the landscape, the risk of losing valuable areas inside the units is very high. It is important to understand that the use of large planning units will automatically cause a decrease in the quality of results and that the sizes of planning units should be selected carefully.

6.9 Exercise 9

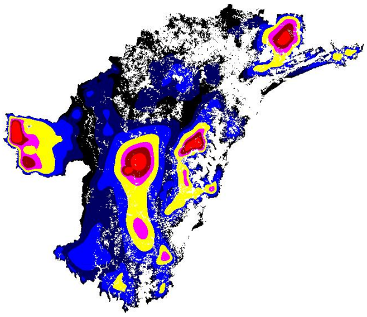
Administrative units

Conservation decisions are often taken at the national or regional level. Different administrative units may, however, have different conservation priorities even though species population dynamics and connectivity effects extend across administrative unit borders. By incorporating administrative units into a Zonation analysis, it is possible to balance the local and global conservation needs of biodiversity features and to account for connectivity across borders. See section 2.12 for more information about using administrative units in an analysis.

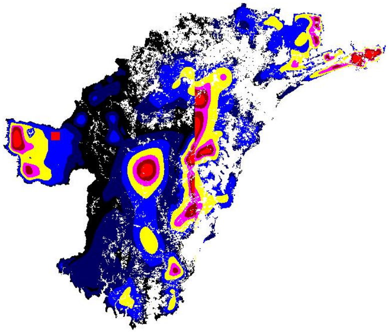
In this exercise, assume that the planning region contains three different administrative units: Eastland, Westland, and Boxland. These areas are defined in the ADMU layer file **ADMU_map.asc** and can be seen by opening this map layer in the Zonation GUI. To begin this exercise, assume that all species are considered equal both globally and within each region (all weights equal in **splist_ADMU.spp** and **ADMU_weights.txt**). Assume also that all regions are equally valuable (G_A is 0.333 for all regions, **ADMU_descriptions.txt**). Be sure to include the appropriate settings in the run settings file (see **set_ADMU.dat** in the tutorial files). See section 3.3.3.16 for more information about the files and settings needed to conduct an analysis with administrative units. Use the batch file **do_ADMU.bat** to run this analysis from the command line or the Zonation GUI.

With the weak variant (ADMU mode 1, image below, left) the solution does not differ considerably from an analysis that does not distinguish between administrative units. The strong variant result (ADMU mode 2, image below, right), however, clearly shows each region's own priorities, especially when almost all value has been given to local considerations (p=0.001). Increasing the global weights (*p*) value will give increasing emphasis to the global component, and the result will again start to resemble a global solution. Please note that a run settings file for ADMU mode 2 is not provided with this tutorial. To conduct this analysis, adjust the run settings file so that ADMU mode 2 is selected and provide a value between 0 and 1 for the global weight (*p*) value. Note also that the *p* value is not actually a weight, but rather it balances how much weight is give to local vs. global priorities. Remember to change the name of the output file to avoid overwriting the ADMU mode 1 solution.

ADMU mode 1:
All weights equal

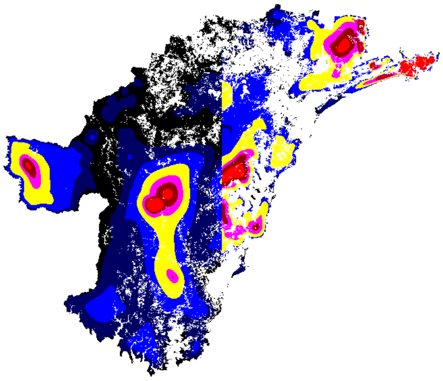


ADMU mode 2:
All weights equal, p=0.001



Next, assume that Eastland is a region with a small human population. Therefore, the habitats in Eastland are in better condition and the threats to biodiversity are lower. Since the species have better habitat in Eastland and assuming that there are fewer pressures on land use due to low human population, it is possible to give Eastland a higher weight to ensure higher protection there. To do this in ADMU mode 1, adjust the **ADMU_descriptions.txt** file to include the following weights: $G_{Eastland}$=0.5, $G_{Westland}$ and $G_{Boxland}$=0.25. It is also possible to prioritize regions under the most pressure and ignore the areas where the species will be safe anyway by giving higher weights to those regions under the most pressure.

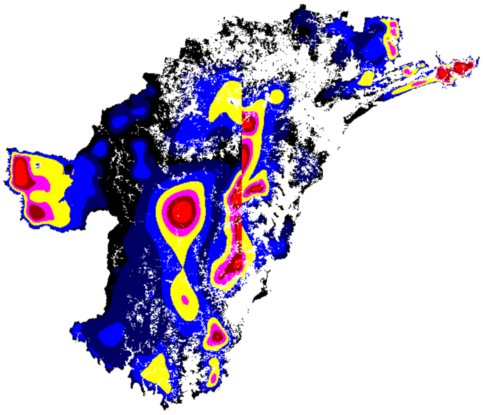
ADMU mode 1:
$w^G_{Eastland}$ = 0.5, $G_{Westland}$ and $G_{Boxland}$=0.25



With strong priorities (ADMU mode 2), representation is required for all species within all regions regardless of their size. This may produce politically unacceptable outcomes, such as prioritizing the entire Boxland in this example. Area-standardized regional weights are one option to alleviate this problem. Try assigning area-standardized weights (G_A in the **ADMU_descriptions.txt** file) of 0.739, 0.259 and 0.002 to Westland, Eastland, and Boxland, respectively. To see the effect best, run it with a low p value, such as 0.001.

ADMU mode 2:

$w^G_{Westland}$=0.739, $w^G_{Eastland}$=0.259, $w^G_{Boxland}$=0.002



Modify the global (from the file **splist_ADMU.spp**) and regional (from the file **ADMU_descriptions.txt**) species weights to see their influence on prioritization. When species weights are not equal, the $ß_A$ parameter comes into use. Remember that the $ß_A$ parameter tunes how much weight administrative unit A places on global considerations. So far its value has not had an influence because all global species weights have been one. Try adjusting it as well to see the influence. Modify one thing at a time to keep the changes in results tractable.

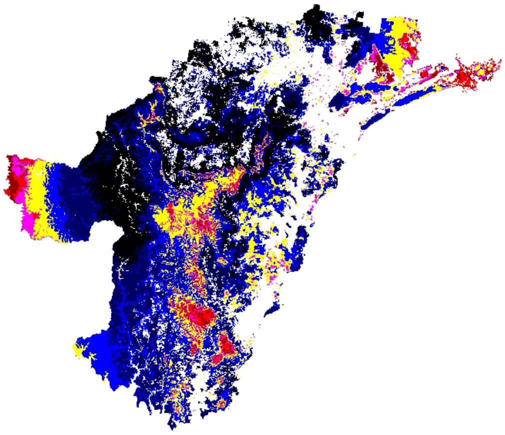
6.10 Exercise 10

Community level analysis with similarity expansion

This exercise illustrates the community-based approach to reserve selection. Three layers representing different community types are considered in this analysis:
community1.asc
community2.asc
community3.asc

In this exercise, output files are directed to a folder called **out_community**. This folder is included in the tutorial download for Zonation v.3.1 and higher. If this folder is not present, please create it or adjust the name of the output file in all batch files to save the output files in the desired location.

First, run the analysis with the three community layers only. Use the batch file **do_comm_nosimilarity.bat** to run the analysis in the Zonation GUI or from the command prompt. This batch file calls the community type list file **bdlist_community.spp** and the run settings file **set.dat**. The priority rank map for this analysis can be seen below.



Next, run an analysis with a community similarity matrix. Use the batch file **do_comm_similarity.bat** included with the tutorial. The same community type list file is used as before, but now a community similarity matrix file called **community_similarity.txt** is included. The community similarity matrix (see figure below) is the basis for a similarity expansion and its use is briefly explained below the figure. The run settings for this analysis are included in the run settings file **set_community_similarity.dat**. See section 2.8 for more information about the theory of community level analysis and the community similarity matrix. See sections 3.3.3.4 and 5.3.1 for more information about input file and analysis set-ups.

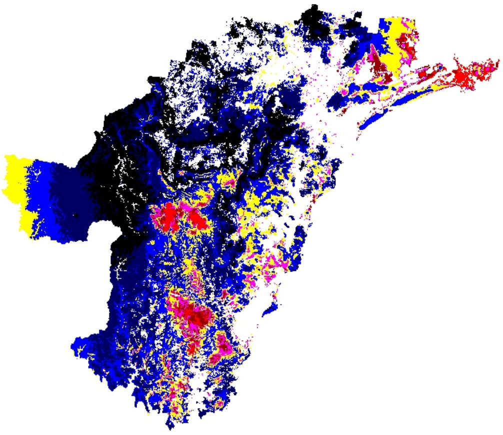
		Community		
		1	2	3
Community	1	1	0.6	0.1
	2	0.6	1	0.1
	3	0.1	0.1	1

Example of a community similarity matrix.

With the community similarity matrix, Zonation calculates representation for community type 1 for one grid cell as follows:

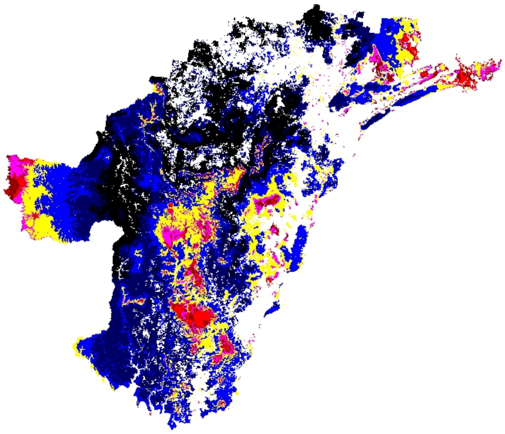
representation(community1) = occurrence(community1) + occurrrence(community2)*0.6 + occurrence(community3)*0.1

The other community types are treated in a similar way. Below is an example of the priority rank map. Compare this map to the previous analysis where the community similarity matrix was not used to see how the matrix effects the ranking of cells.



Please note that the representation levels in the **.curves.txt** output file are calculated for the expanded occurrences of community types (the values calculated with the similarity matrix). Use solution loading to see how the original community types are represented in the solution (see section 3.5.2). Call the program with **do_comm_original_rep.bat** where Zonation will remove cells in exactly the same order as it did in the previous analysis, but while calculating representation levels for the original community types. This analysis produces a file called **output_commsim_original.curves.txt**, which shows how the representation of the community types themselves change during cell removal.

Zonation can also identify conservation priorities based on both community types and individual species. In the biodiversity feature list file (**sp_community_list.spp**), list the community types first followed by the individual species that will be included in the analysis. It is very important that the communities are listed first in the biodiversity feature list file as the community similarity matrix will read the features from the beginning of the file. See section 5.3.2 for more information about this. It is useful to create a grouped output so that the representation curves for community types and species can be examined separately. The groups file **groups_community_sp.txt** is included in this analysis where community types are assigned to output group 1 and species are assigned to output group 2. The analysis settings are defined in the run settings file **set_community_sp.dat**. See sections 3.3.3.4 and 5.3.2 for more information about the specific settings needed for this analysis. Call Zonation with the batch file **do_comm_sp.bat**. The priority rank map from this analysis is shown below.



As community types and species were assigned separate output groups, a groups file (**output_comm_sp.grp_crvs.txt**) is created as part of the output. This file summarizes minimum, mean, and maximum representation of both output groups separately. The representation curves of individual features are described in the **.curves.txt** file.

6.11 Exercise 11

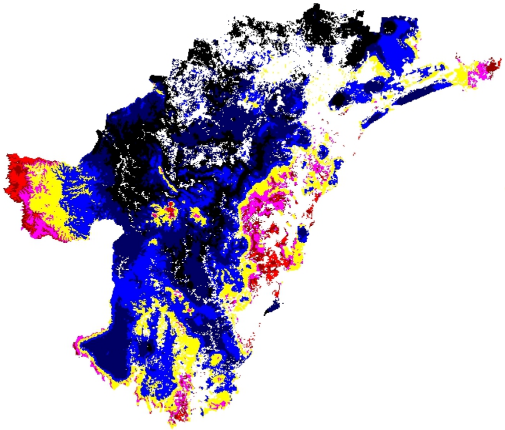
Landscape condition and retention

Use of condition layers is one of the simpler and more useful techniques available in Zonation. It is both ecologically highly relevant and technically simple to include condition information for habitats. In contrast, using the retention feature is much harder.

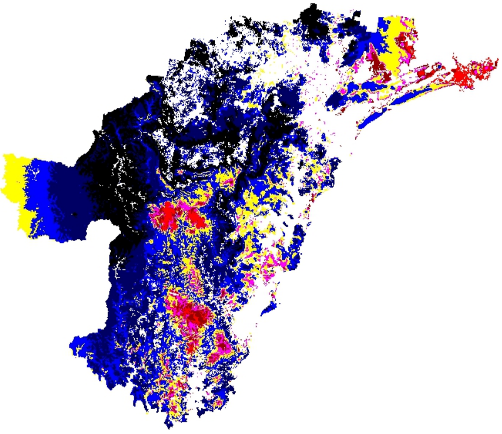
This example illustrates how landscape condition and retention affect the conservation priorities. Community types are often the most practical units for this type of analysis, and the same community types from exercise 10 are used in this example. Community similarity is also considered. The output folder for all analyses in this exercise has been defined as **out_cond_ret**. This folder is included as part of the tutorial download. If it is not present, please create a folder with that name or adjust the batch files so that the output files will be directed to the appropriate location.

First, consider an analysis using only landscape condition. Landscape condition describes the fraction of suitable habitat remaining for each grid cell relative to some historical baseline. Remember from the previous exercise that community types 1 and 2 are rather similar while community type 3 is very different from both of the other two. Because of this, community types 1 and 2 have been assigned to condition group 1 and community type 3 has been assigned to condition group 2 in the second column of the groups file **groups_condition.txt**. This links the community types to the correct landscape condition raster in the condition layer list file. Community types 1 and 2 are linked to the landscape condition file **condition1.asc** and community type 3 to **condition2.asc**. See sections 3.3.3.12 and 3.3.3.14 for more information about the condition groups and the condition layers. The necessary settings for this analysis are defined in the run settings file **set_condition.dat**. Call Zonation with the batch file **do_condition.bat**.

Condition Layer+Community Similarity

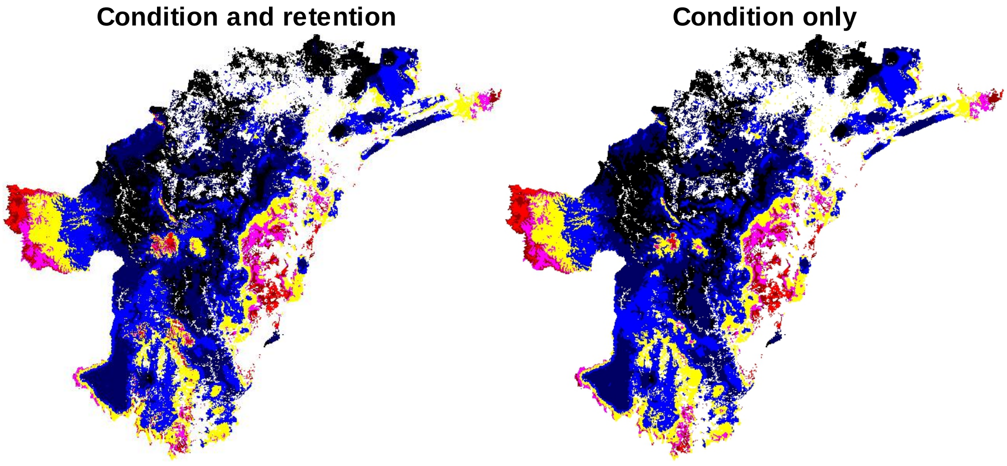


Community Similarity Only



The figures above show the ranking of an analysis conducted with a condition layer and community similarity (left) and an analysis with only community similarity (right). As community types 1 and 2 have good habitat condition (little past habitat loss) in the capes and in the south, their conservation value is increased in those areas. For community type 3, habitat condition is highest at the western edge, which increases the conservation value for that region.

Next, consider an analysis where retention is considered along with condition. Retention layers indicate the fraction of local occurrence or habitat suitability that would be retained for biodiversity features in the absence of conservation action. The idea in retention analysis is to balance the representation of features in reserves as well as their retention across the full landscape. To do this, all community types are listed twice in the biodiversity feature list file **bdlist_retention.spp**. The first set of community type files is used to determine representation in reserves while the second set is used to determine loss in occurrence if the cell is not selected for the reserve network. This analysis assumes that all community types have a similar pattern of retention, and therefore just one retention layer (**retention1.asc**) is included in the analysis. Retention layers are listed in the retention layer list file **ret_list.txt** which links the retention layers to the retention groups defined in the groups file **groups_retention_m1.txt**. The groups file also indicates that 1) the analysis will produce separately grouped outputs for the layers used to calculate representation (the first set of layers) and those used to calculate retention (the second set), 2) condition is applied to both sets of layers, 3) retention is applied only to the second set of layers and all layers belong to retention group 1, and 4) retention mode 1 ("stop loss") is used for all features. See section 3.3.3.12 for more information about the groups file. The analysis settings are defined in the run settings file **set_retention_m1.dat**. See section 3.3.3.15 for more information about the necessary settings for this analysis. Call the program with the batch file **do_retention_m1.bat**.



Condition and retention

Condition only

The figures above show the prioritization ranking of an analysis conducted with condition and retention (left) and with condition only (right). Retention is predicted to be high in the central and eastern regions (as seen in the retention layer file). Therefore, these areas gain priority when retention is included. This can most clearly be seen in the central regions of the images above.

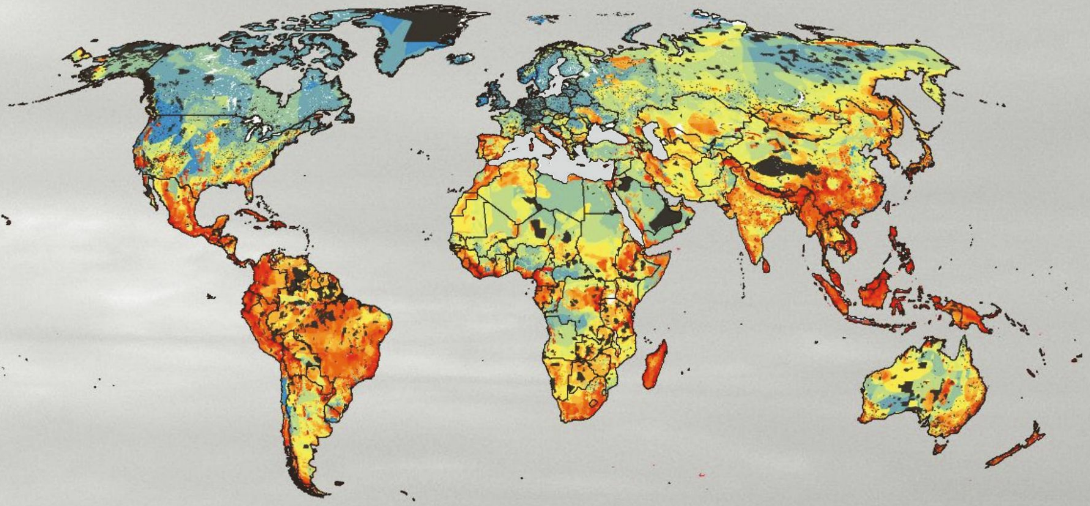
See section 2.10 for more information about the theory and algorithms behind landscape condition and retention.

Index

For the latest version of Zonation, please visit our website: http://cbig.it.helsinki.fi

Zonation implements methods and software for spatial conservation prioritization; it is a decision support tool for conservation resource allocation. Zonation identifies areas important for retaining habitat quality and connectivity for multiple species, habitats, ecosystems, ecosystem services, etc., indirectly aiming at long-term persistence of biodiversity. When including costs, Zonation produces cost-efficient solutions. Typical uses of Zonation include design of reserve networks and their expansions, identification of areas for ecological impact avoidance, and evaluation of proposed conservation solutions developed elsewhere. Zonation also implements traditional target-based planning commonly used within systematic conservation planning.

The operation of Zv4 is explained in this user manual. See also scientific publications and our web site for further information, including a quick introduction guide.



Contact information

C-BIG Conservation Biology Informatics Group
Finnish Centre of Excellence in Metapopulation Biology
Department of Biosciences
P.O. Box 65 (Viikinkaari 1)
FI-00014 University of Helsinki, FINLAND
atte.moilanen@helsinki.fi

ISBN 978-952-10-9919-9 (paperback)
ISBN 978-952-10-9920-5 (PDF)

The following partners have contributed significantly to the development of Zonation