

Steps to PRISM

1. Clip Raster
2. Reclassify
3. Regiongroup
4. Generalize Regions
5. Genregionsampleplots
6. Mergesampleplots
7. Isectpolyst
8. Reclassifyrecords
9. Export.csv
10. R.sample

1. CLIP RASTER

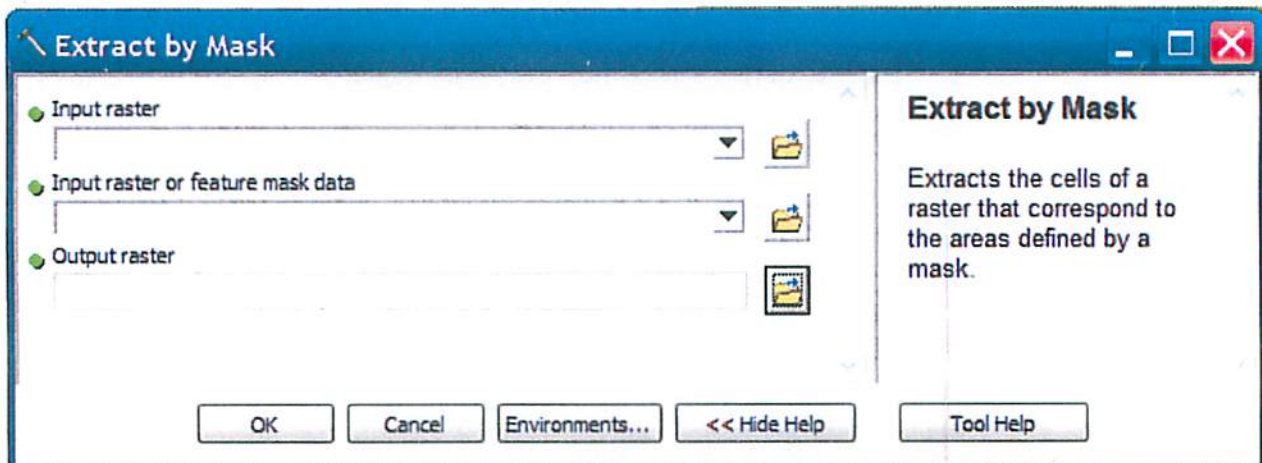
* Ensure that every layer is in the same projection:
"NAD 1983 LAMBERT CONFORMAL CONIC"

Clip landcover raster to the extent of the study area. In order to do this use the Extract By Mask (Spatial Analyst) tool in ArcGIS 10.0 – create a polygon to represent the area to be clipped.

Input raster = the landcover raster

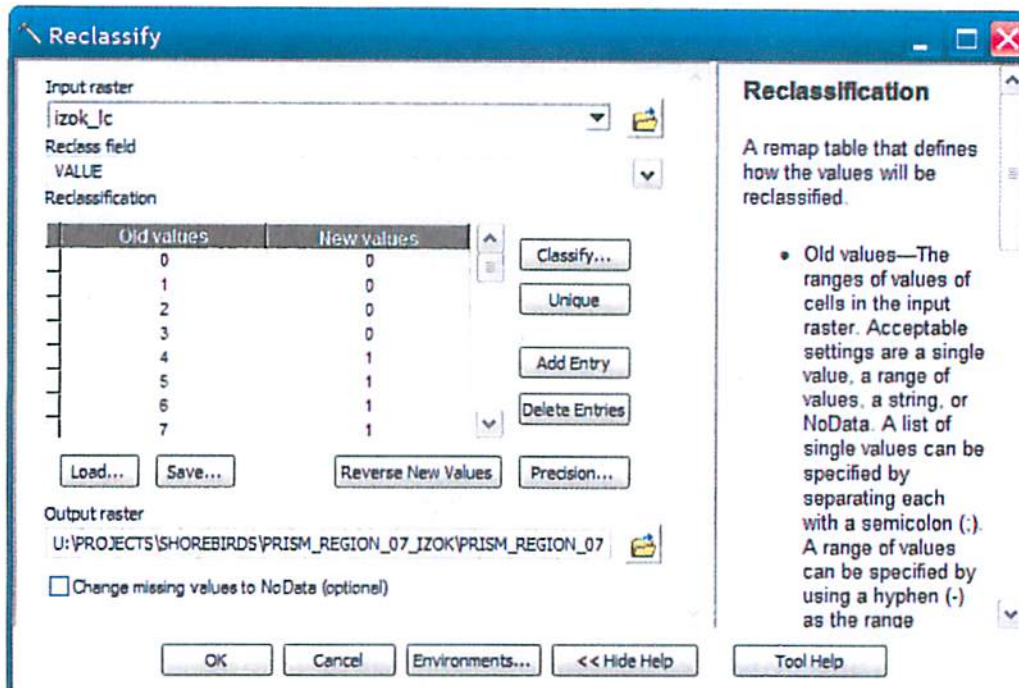
Input raster or feature mask data = polygon representing the extent of area to be clipped

Output raster = the location where you wish to store your raster output (filename is limited to 9 characters - raster name cannot start with a number otherwise you will receive an error)



2. RECLASSIFY

The raster output from step one is now used to reclassify the habitat types into suitable and non – suitable areas (value of 0 representing unsuitable areas and a value of 1 representing suitable areas).



Value	Code	Description	Reclass	Suitable/Unsuitable
1	WA	Water	4	0
2	CL	Cloud	NoData	0
3	SH	Shadow	NoData	0
4	DA	Mine sites	3	1
5	EG	Exposed gravel/cobble	3	1
6	ES	Exposed silt/sand	3	1
7	RR	Boulder/Bedrock	3	1
8	LV	Lichen veneer	3	1
9	HR	Heath rock	3	1
10	HT	Heath tundra	3	1
11	HC	Heath cryoturbated	2	1
12	HG	Heath graminoid	2	1
13	ST	Tussock sedge	1	1
14	SN	Non-tussock sedge	1	1
15	BF	Riparian shrub	2	1
16	BL	Low shrub	3	1
17	BT	Tall shrub	3	1
18	CO	Conifer open	3	1

Input raster = the raster output from step one

Reclass Field = the field in the attribute table that identifies the habitat types (in the table above this would be the Reclass field).

(ie. 1= Bedrock → suitable; 2= water → unsuitable; 3= wetland → suitable)

Reclassification – enter the new values (enter 0 and 1 for all habitat types to identify which areas are suitable and unsuitable)

Output Raster = the reclassified raster representing unsuitable (0) and suitable (1) habitat types.

****Unsuitable habitat types include water, clouds and no data – assign these areas a value of 0**

****Suitable habitat types include any other types of habitat (ie. Bedrock, wetland, dry upland, low shrub, conifer open, etc.) – assign these areas a value of 1.**

3. REGIONGROUP

The Region Group Tool is designed to identify regions of contiguous cells. In order to run this tool you must have all previous raster and vector data loaded into your ArcMap project and open the Spatial Ecology Geospatial Modeling Environment (GME) tool.

This is a raster tool that identifies contiguous blocks of cells of the same value (i.e. regions) and assigns each region a unique ID number that is written to the output raster dataset. It is designed for use only with thematic (categorical) rasters.

The cell neighbourhood that is evaluated can be all eight cells that surround a cell (this is the default option), or can be limited to only the four cardinal cells using the 'diag=FALSE' option, which prevents diagonal connections being evaluated when defining region membership. For our purposes we will accept the default option. → is this true? YES.

The unique ID numbers in the output raster are arbitrary, but are ordered by when the region is first encountered when moving left to right and top to bottom through the raster dataset. The output also includes a DBASE (dbf) table that summarizes the number of cells that belong to each region.

For large raster datasets this tool can take quite some time to run (a few hours for very large rasters). It is therefore recommended that you test the tool on a small raster dataset before applying it to the full dataset

The formula for this tool is as follows: regiongroup(in, out, [diag]);

- in → the input integer raster data source
- out → the output integer raster data source
- [diag] → (TRUE/FALSE) allow cells to be diagonally connected when defining regions? (default=TRUE)


```
regiongroup(in="C:\izok_lc_rec", out="C:\izok_rec_rg");
```

Processing complete.

4. GENERALIZE REGIONS

This tool generalizes / simplifies a thematic (categorical) raster dataset by recoding the smallest contiguous groups of pixels of the same value with the pixel value of the largest neighbouring group. You must run the regiongroup command before running this command as the output of the regiongroup command serves as the input to this command.

The cell neighbourhood that is evaluated can be all eight cells that surround a cell (this is the default option), or can be limited to only the four cardinal cells using the 'diag=FALSE' option, which prevents diagonal connections being evaluated when defining region membership. You will probably want to use the same option here that was used with the regiongroup command.

This is an iterative command whereby the maximum allowed number of iterations is controlled by the limit option. If your raster is simple (relatively homogeneous) then only one or two iterations may be required. However, if portions of the raster are highly heterogeneous, and many small groups of pixels are mixed together, then more iterations will be required to arrive at a solution that no longer changes. As a general rule, 20 iterations is a reasonable starting point, but it is recommended that you inspect the raster to ensure the command has produced consistent results. Watch the number of groups to classify decrease in the progress bar as it processes. If it reaches 0 then it has reached an endpoint that cannot be further improved.

```
generalizeregions(in, table, limit, out, [diag]);
```

in the input raster region data source that was created from the regiongroup raster output.

table the input table data source resulting from the regiongroup command

limit the minimum region size expressed as the minimum number of connected raster pixels

out the output raster data source

[diag] (TRUE/FALSE) allow cells to be diagonally connected when generalizing regions?
(default=TRUE)

****For our purposes the limit = 3 and accept the default to allow cells to be diagonally connected when generalizing regions.**

```
generalizeregions(in="C:\izok_rec_rg", table="C:\izok_rec_rg.dbf", limit=3,  
out="C:\izok_recl_gr");
```

4
if raster doesn't complete as GRID use .img. or .tif export format

17623 cells recoded

5. GENREGIONSAMPLEPLOTS

This tool generates a sampling grid: a series of regularly sized and spaced square or rectangular polygons. Unlike the simpler 'genvecgrid' command, however, this command provides options for generating both plots and zones (plots are the smallest sampling unit, and zones are groups of plots), for characterizing the plots and zones based on a suitability raster, and for taking into account barriers in the landscape. For instance, this tool can be used to generate sampling plots that do not cross rivers.

The extent of the sampling grid can be specified by referencing a feature data source, or by directly specifying the minimum and maximum x and y coordinates. An ideal choice of reference layer would be a polygon dataset containing a study area boundary. Sampling plots are only generated in areas covered by the features in this layer. Alternatively, if you specify the extent of a rectangle then plots are generate throughout that area.

The primary output of this tool is a plot polygon layer. You may optionally also generate zones (see the 'outzones' and 'dimzones' command options). Zones would be useful if you wish to perform stratified random sampling in order to cluster sample plots in zones. For instance, if the extent of the area you are sampling is very large then pure random sampling of plots may create a distribution of sample plots that is impractical to visit in reality. A more detailed discussion of sampling designs and the issues that motivate different designs is included in the sampling chapter.

The dimensions of the plots (and zones) are specified in coordinate system units (e.g. meters for the UTM projection). It is recommended that you do not use this tool with unprojected data (geographic coordinate systems). The dimension of the zones should be exactly divisible by the dimension of the plots. For instance, plots of 100m and zones of 1km is suitable, but plots of 30m and zones of 1km is not.

The snap option forces the spatial position of the grid to be aligned with the underlying coordinate system. For instance, specifying snap=1000 would force the vector grid lines to be aligned with the exact 1km positions in a UTM grid. For most sampling purposes this is not an important consideration.

The 'raster' option refers to a suitability raster that classifies the landscape into areas that can be sampled (1) and those that cannot (0). For instance, if you are sampling forest habitat in a landscape that is a mixture of forest and open areas, then the raster would depict all forested areas as 1, and all open areas as 0. This raster is used to calculate the proportion and total area of suitable habitat in each plot (and zone), stored in two fields in the attribute table.

This tool adjusts the plots and zones to reflect barriers in the landscapes, which can be depicted as polygons and lines. For the polygons, the tool will remove all plots that are completely

contained by barrier polygons, and will clip plots that overlap the barrier polygons. Barrier lines are used to split plots into components that do not cross the barrier lines. You can specify barriers as polygons only, lines only, or both lines and polygons.

After the geometry of the plots has been modified by these barriers, an area threshold (expressed as a proportion of the plot area) is used to identify plot fragments that will be merged with adjacent plots. Any fragments greater than this threshold are considered to be close enough to a full plot size that they can be retained without merging in the plot dataset. Each of the smaller fragments is merged with the neighbouring plot with which it shares the longest boundary, provided this does not involve crossing a barrier.

The final barrier-related parameter is a tolerance that defines the distance from the edge of the plot that barrier polygons or lines can be ignored. This parameter is needed because, with certain data formats, there can be very small movements of vertices following the splitting procedure (as a result of coordinate precision). Usually this is only a few centimetres and is negligible, but does create a problem when resolving merges of adjacent plot polygons. The default tolerance is 1, and the units for this tolerance are the coordinate system units (e.g. meters for UTM). If you use a barrier line dataset that contains overshoots, then you may need to increase this tolerance to allow the merging algorithm to resolve acceptable merges. It is particularly important when using a barrier line dataset that 1) lines that are supposed to cross do actually cross (e.g. roads at a T junction), and 2) that overshoots at such crossings are not too long. The former issue results in plots that fail to be split by a barrier line, while the latter issue can result in a failure of the algorithm to deal with plot fragments.

genregionsampleplots(extent, out, dim, [outzones], [dimzones], [snap], [ráster], [barrierpoly], [barrierline], [minarea], [tol]);

extent	the reference layer that defines the extent of the vector grid, or a set of four values that define the extent (min x, max x, min y, max y) This is the polygon that was created in step 1.
out	the output polygon data source for the plots
dim	the dimensions of the plots in coordinate system units, e.g. 100 or c(100,200); specifying one value results in square plots
[outzones]	the output polygon data source for the zones (if unspecified no zone output is produced; if specified, dimzones must also be specified)
[dimzones]	the dimensions of the zones in coordinate system units (this value should be an exact multiple of the plot dimensions), e.g. 1000 or c(1000,2000); specifying one value results in square zones
[snap]	a value > 0; controls whether the vector grid is aligned with a major coordinate system interval (supplying a value of 1000 will result in the grid being aligned to the 1000-mark intervals of the coordinate system); default=0 (no snap)
[ráster]	a suitability (1/0) raster that is used to characterize plots (see help documentation for further details)
[barrierpoly]	a polygon data source representing barriers that cannot be sampled (see help

.. 416 739 4216

Paul -

- documentation for further details)
- [barrierline] a line data source representing barriers that cannot be crossed (see help documentation for further details)
- [minarea] the threshold, expressed as a proportion of the plot area, that is used to determine whether a plot fragment will be retained as a fragment, or merged (default=0.8, see help documentation for further details)
- [tol] the tolerance that is used in the merging of plot fragments resulting from barriers (see help documentation for further details)

****For our purposes the dimensions of the plots is 300metres across and 400metres in height. In order to show this the proper syntax is dim=c(300,400) – see formula below. The dimensions for the zones is always 1200 (dimzones=1200). Ignore [snap], [barrierpoly], [barrierline], [minarea], and [tol]. The raster is the suitability raster that was created in step 2 once it had been reclassified.**

→ generates zones, then plots, then linking.

`genregionsampleplots(extent="C:\Extent.shp", out="C:\plots.shp", dim=c(300,400), outzones="C:\zones.shp", dimzones=1200, raster="C:\izok_recl_gr");`

45484 zones created

544611 plots created

Plots updated with zone ID's

ensure that all layers are in the same projection

Warning: One or more plots were detected with partial overlap of the suitability raster. The proportion of suitable habitat calculation may, therefore, be incorrect for these polygons. It is recommended that you supply a suitability raster that covers all of the zones and plots.

Warning: One or more zones were detected with partial overlap of the suitability raster. The proportion of suitable habitat calculation may, therefore, be incorrect for these polygons. It is recommended that you supply a suitability raster that covers all of the zones and plots.

Obtained suitability data

300243 plots filtered based on suitability

23505 zones filtered because they contained no plots

244368 plots created.

→ additional fields include PLOTAREA, PcsuitHab, AcsuitHab.

A warning for this step will appear – this is simply indicating the some of the zones/plots fall partially outside of the raster in which case the suitable habitat calculation may be incorrect. Such plots and zones cannot be used as plots/zones to be surveyed – if randomly selected perform another random selection to avoid any plots/zones that have partial overlap outside of the raster.

6. MERGE SAMPLE PLOTS

The purpose of this tool is to merge sampling plots (and/or zones) that contain less than a specified area of suitable habitat. It is designed for use with plot and zones layers that have been created with the `genregionsampleplots` command. This tool can also be applied to a plot layer that has no corresponding zone layer. The difference is that when a zone is also specified, merging of plots occurs preferentially within the same zone.

It is essential that you specify the zone and plot layer that were created simultaneously using the `genregionsampleplots` tool. If you have run that tool more than once, you must not mix a zone and plot layer from different sessions (the unique ID values are unlikely to match and this will result in this merging tool producing nonsensical data). It is also not advisable to use a previously merged zone/plot layer as the input to this tool.

Define the minimum suitable habitat area in coordinate system units e.g. square meters for UTM. This value corresponds to the values in the `ArSuitHab` field that were produced by the `genregionsampleplots` command. Any zones with a value of `ArSuitHab` less than the value specified will be merged using the merging rule you specify. When a plot is merged, the attribute fields are updated and the polygon it has been merged with is deleted. Note that the new, merged plot is not evaluated to determine if the minimum suitable area criteria has been met. However, this merged polygon is available for other plots to be merged with. Note that after merging the plot IDs will still be unique, but may not be a consecutive series as some of the plots have been deleted. After the plots are merged, the zones are evaluated for merging. If a zone is merged then the plots that are contained within this zone are updated to reflect the new zone ID value.

Merging Rules. For all merging rules the term neighbour refers to the four bordering plots in the four cardinal directions. A plot on the edge of a zone, however, will have fewer than four neighbours because merging of plots across zones is not permitted.

1. **Dominant Neighbour Rule.** The plot (or zone) is merged with the neighbour that contains the greatest total area of suitable habitat (the maximum `ArSuitHab` value).

2. **Longest Border Rule.** The plot (or zone) is merged with the neighbour that shares the border passing through the greatest amount of suitable habitat. The line representing the shared border is intersected with the suitability raster and the proportion of suitable habitat cells intersected by that line is calculated. The border with the greatest value defines which polygons are merged.

3. **Simple Contiguity Rule.** The plot (or zone) is split into halves: north and south halves, and east and west halves (i.e. two pairs of halves, not four quarters). The half containing the greatest proportion of suitable habitat defines the neighbour that is selected for merging.

In landscapes in which the suitable area is highly fragmented, the Dominant Neighbour Rule is probably the best choice.

`mergesampleplots(in, out, area, rule, [inzones], [outzones], [areazones], [rulezones], [raster]);`

`in` the input plot data source that was created using the `genregionsampleplots`

→ 1:3000

command (it must contain a `ArSuitHab` field)
out the new, merged, output plot data source
area the minimum area of suitable habitat per plot (identifies the plots that will be merged)
rule the number representing the plot merge rule: 1=Dominant Neighbour, 2=Longest Border, 3=Simple Contiguity (see full help documentation for details)
[inzones] the input zone data source that was created using the `genregionsampleplots` command
[outzones] the new, merged, output zone data source
[areazones] the minimum area of suitable habitat per plot (identifies the plots that will be merged)
[rulezones] the number representing the zone merge rule: 1=Dominant Neighbour, 2=Longest Border, 3=Simple Contiguity (see full help documentation for details)
[raster] a suitability (1/0) raster that is used to characterize plots, required for rules 2 and 3 (see help documentation for further details)

```
mergesampleplots(in="C:\plots.shp", out="C:\plots_merge.shp", area=36000, rule=3,  
inzones="C:\zones.shp", outzones="C:\zones_merge.shp", areazones=432000, rulezones=3,  
raster="C:\izok_recl_gr");
```

12443 plots identified for merging

test → 127

991 zones identified for merging

test → 4

Merging complete

**** Minimum plot area for suitable habitat is 30% of the plot which is always 36000 metres squared.**

**** Minimum zone area for suitable habitat is 30% of a zone which is always 432000 metres squared.**

**** The rule for both plots and zones is rule #3 (simple contiguity)**

7. ISECTPOLYRST

This tool summarizes the raster cell values that are contained by a polygon. The output consists of summary fields that are added to the polygon attribute table. It processes thematic (i.e. categorical) and continuous rasters in different ways. For thematic rasters, a new field is added

for each unique value in the raster and is populated with either a count of the number of cells of each raster value within that polygon, or the proportion of cells of each raster value. For continuous rasters a statistical summary is produced (the mean, median, minimum, maximum, standard deviation, count, and the 95% quantiles).

**** For our purposes we are using thematic rasters so we will not be accepting the default option – the expression must contain thematic = TRUE in order for the process to run properly.**

Polygons are processed sequentially, so are not affected in any way by overlapping polygons (in contrast to related tools in ArcGIS). This means that it can take a considerable time to process many polygons.

This tool cannot distinguish between thematic and categorical rasters automatically. By default it assumes the raster represents continuous data, **so if the raster is a thematic raster you must explicitly use the 'thematic=TRUE' option.**

NoData cells are ignored. Therefore, if the polygon contains NoData cells this may influence the accuracy of the statistical summary. It is highly recommended you ensure that raster data exists for the entire area covered by the polygons. By default, any polygons that are entirely or partially outside of the extent of the raster will not be processed – they will be coded with NoData values (-2147483648) so should be easily identifiable. This behaviour is a precaution against potentially biased estimates, but can be overridden using the allowpartialoverlap parameter. Setting this parameter to TRUE will force all polygons to be processed, even if they only partially overlap the extent of the raster dataset.

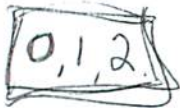
Note that this single command provides the same functionality as both the 'Zonal Statistics ++' and 'Thematic Raster Summary' tools in Hawth'sTools.

Recent changes to the interface with R mean that the median and quantiles options (calculated of medquant=TRUE) will tend to run quite slowly. I am working on resolving this speed issue for a future update.

`Isectpolyst(in, raster, prefix, [thematic], [proportion], [metrics], [allowpartialoverlap], [medquant], [where]);`

in	the input polygon data source
raster	the input raster data source
prefix	a short prefix to use in the summary statistic fields that are added, which end with MN, MIN, MAX, STD, CNT, MED for continuous raster data, and V# for thematic raster data where # is each of the unique cell values. The prefix should be no longer than 6 characters, and should be related in some way to the raster dataset.

- [thematic] (TRUE/FALSE) controls whether the raster should be treated as a thematic (categorical) raster (default=FALSE)
- [proportion] (TRUE/FALSE) only applies to thematic rasters: records the proportion of the polygon in each cell value rather than the area of the polygon (default=FALSE)
- [metrics] the statistical metric(s) to calculate, expressed as a single value (e.g. "MIN") or as a vector of metrics (e.g. c("MN","STD","CNT")); options are MN/MEAN, MIN, MAX, STD, SUM, CNT, MED, QUPP, QLOW (see help for details)
- [allowpartialoverlap] (TRUE/FALSE) if TRUE, will process polygons that only partially overlap the extent of the raster dataset (default=FALSE)
- [medquant] (TRUE/FALSE) controls whether the R derived median and quantiles are included in the statistical summary statistics (default=FALSE)
- [where] the selection statement that will be applied to the feature data source to identify a subset of features to process (see full Help documentation for further details)



**** The expression must contain proportion=TRUE so that the expression returns the proportion of the polygon in each cell instead of the area of the polygon.**

**** The raster to be used is the raster containing the habitat types (0=no data/clouds, 1=good, 2=medium, 3=poor, 4=water) DO NOT use the suitability raster (0, 1) for this step.**

**isectpolrst(in="C:\PRISM_REGION_07_IZOK\DATA\plots1merge.shp",
raster="C:\PRISM_REGION_07_IZOK\RASTER_ANALYSIS\CLASSIFIED_GRIDS\izok_lc_recl",
prefix="PRHAB", thematic=TRUE, proportion=TRUE);**

189858 polygons processed.

4124 polygons processed

8. RECLASSIFY RECORDS

A new field within the attribute table must be created for Plots_merged output. The new field should be called 'Reclass' as this is where the plots are reclassified based on a specific set of conditions (see below). Use the select by attributes to go through the set of conditions.

** This set of conditions needs to be changed in order to eliminate selecting plots that are greater than a certain size and ~~so~~ that an equal # of low, medium, poor plots are selected*

Select plots with plot areas ≤ 36000

0 = unsuitable.
1 = suitable.
2 = snow + ice.

only looks at plots!