Ensemble tool for predictions from Species Distribution Models (eSDM) GUI

User Manual

The web-based interface of a user-friendly spatial tool that allows users to import SDM predictions and create and explore ensemble predictions to inform management and explore spatial uncertainties

Samuel Woodman
September 2018

Table of Contents

Ensemble tool for predictions from Species Distribution Models (eSDM) GUI user manual

1. Introduction

1.1. Why create ensembles?

As the field of species distribution modeling (SDM) for marine species in their dynamic environment has advanced, the resulting spatial models have greatly enhanced the ability of marine resource managers to assess and mitigate potential impacts to protected species at the appropriate spatial resolutions. However, different data sets or different analytical approaches often yield different modeled results, creating uncertainty and challenges for management decision-making. As an example, there are currently multiple spatial and habitat-based models of blue whale (*Balaenoptera musculus*) density, distribution, and biologically important areas off the U.S. West Coast (Forney *et al*. 2012, Redfern *et al*. 2013, Irvine *et al*. 2015, Calambokidis *et al*. 2015, Becker *et al*. 2016, Hazen *et al*. 2016, Redfern *et al*. 2017). These models can directly inform management decision-making, e.g. for reducing ship-strike risk, but different models may suggest different actions because of strengths, biases, and limitations for underlying data set and model.

When original data sources are available to support an integrated analysis, i.e. in a Bayesian hierarchical framework, this can provide a robust, probabilistic assessment of the combined information on spatial distributions of protected species. However, this is generally a time-consuming and (at times) analytically challenging approach, and highly disparate data sources may be too complex to allow an effective joint analysis given computational capabilities. In other cases, the different investigators may only be willing to share the output from their SDMs, and not the original survey, telemetry, or other data sources.

For these reasons, a method of combining model outputs (i.e., predicted species distributions) is needed, allowing end-users and managers to explore and evaluate 'ensemble predictions' that take into account each model's strengths and weaknesses and highlight areas of greater or lesser uncertainty about species occurrence. Predictions derived from an ensemble of models are generally more robust than single-model predictions. Integration and comparisons of predictions across models provides critical information on uncertainty, and provides a foundation for improving modeling methodologies by highlighting key assumptions and limitations of each model and/or data set. Combining multiple models can be complex when models are developed at different spatial resolutions, with different coordinate systems and possibly projections, and using differing prediction units (e.g. animal density, probability of occurrence, presence/absence).

The eSDM has been designed as a user-friendly spatial tool with a web-based interface (the GUI) that allows end-users to import the SDM predictions (i.e. as GIS layers or raster data) and create and explore ensemble predictions to inform management and explore spatial uncertainties.

1.2. Structure of manual

This document describes how to use this tool. The manual is divided into sections demarcated by the individual GUI tabs, and then by the boxes within each tab. Key steps in the ensemble process include: Importing predictions (Section 3); Providing study area and erasing polygon files, and overlaying original predictions (Section 4); Selecting ensemble predictions using various methods (Section 5); Calculating evaluation metrics for predictions (Section 6); Creating high-quality maps (Section 7); and Exporting predictions (Section 8). Support and assistance is provided via a PDF manual (Section 9), and Appendices (Section 10) with details for some of the internal operations and calculations.

1.3. Run the GUI locally

R Shiny applications such as the GUI can be hosted and used online; however, running the R Shiny apps locally can be faster than running from them online. You can install the eSDM from GitHub (https://github.com/smwoodman/eSDM). To install the eSDM package and run the GUI locally, run the following code in the RStudio console:

```
if (!require('devtools')) install.packages('devtools')
# install from GitHub
devtools::install_github('smwoodman/eSDM')

# open the GUI
eSDM::eSDM_GUI()
```

The above code will launch the GUI in your system's default web browser. You can also run launch R Shiny apps in an RStudio window, but it is recommended to run the GUI through a web browser. Most development and testing was done running the GUI in Google Chrome.

1.4. Supported import file types: SDM prediction data

SDM prediction data requirements for all file types
Prediction data imported into the GUI must be absolute density, relative density, or abundance predictions. For the purposes of the GUI, probability of occurrence is considered a relative density. Abundance predictions will be converted to density. Predictions that are one of 'NA', 'NaN', 'N/A', 'n/a', 'na', blank, or a negative number will be classified as NA predictions and will be ignored for most analyses. The major exceptions are that you can display the NA predictions in plots, and if a set of predictions is chosen as the base geometry then prediction polygons from that set with NA values will still be included in the base geometry. The GUI does support predictions that span the antemeridian (180 decimal degrees).

Excel .csv
SDM prediction data can be imported from an Excel .csv file that has headers and

the coordinates of regular points on a grid of prediction polygons, as well as the prediction values themselves. These coordinates must: 1) be in WGS 84 geographic coordinates (decimal degrees); 2) have all longitude values either in the range [-180, 180] or [0, 360]; 3) be on a regular grid and thus equally spaced, although they can be in any order in the file and represent the centroid or any of the four corners of the grid.

GIS raster
SDM prediction data can be imported from an IMG (.img) or TIFF (.tif) file.

GIS shapefile
Shapefiles are stored as multiple files on local file systems, and thus you must browse to and select all of the files from the shapefile you wish to upload into the GUI. There may be as few as four files to upload, or more than ten. The extensions of these files will likely be a subset of the following: '.shp', '.shx', '.dbf', '.prj', '.sbn', '.sbx', '.shp.xml', '.fbn', '.fbx', '.ain', '.aih', '.ixs', '.mxs', '.atx', or '.cpg'. If you have the shapefile open in ArcCatalog or ArcMap, it does not matter if you select the 'LOCK File' when uploading the files.

GIS file geodatabase feature class
To import SDM prediction data from a file geodatabase feature class, you must enter the absolute file path of the file geodatabase, which will end with '.gdb', as well as the name of the feature class object within the file geodatabase. Because you must use the file path to upload the file geodatabase feature class, you can only use this feature when running the GUI locally rather than from a server. If possible, use ArcCatalog to get both the file path of the file geodatabase and the name of the feature class object. Data cannot be imported from personal geodatabases or from file geodatabase raster datasets.

1.5. Supported export file types: SDM prediction data
Original, overlaid, and ensemble predictions may be exported in any coordinate system, and in the longitude range of [-180, 180] or [0, 360] decimal degrees, or the equivalent in a non - lat/long coordinate system.

Excel .csv
To export SDM predictions to an Excel.csv file, the GUI calculates the centroid for each polygon that contains a prediction. The .csv file consists of columns with the longitude and latitudes of these centroids, the prediction values, and (if applicable) the weight values. Note that depending on the selected coordinate system and if the prediction polygons are regular, you may not be able to import the saved predictions into the GUI because the centroids will not be equally spaced.

GIS Shapefile
SDM predictions exported as shapefiles will be exported as polygons with their respective prediction and (if applicable) weight values. These polygons will exactly match the polygons in the GUI. Predictions cannot be exported to any other GIS file format.

KML or KMZ file

Within the KML or KMZ file, predictions will be represented as polygons with a red outline. Currently you cannot color-code the polygons by prediction value. The polygons will have their respective prediction values and (if applicable) weight values as descriptions.

1.6. Supported export file types: image files

You may download plots as JPG, PDF, or PNG file types at low (72 ppi) or high (300 ppi) resolutions.

2. eSDM GUI Roadmap and Load or Save Workspace

2.1. Load or Save the GUI Workspace

<u>Description:</u> You may save the workspace of your current GUI session, or load a previously saved workspace. Note that user selections (e.g. high quality map parameters) and displayed plots are not saved in the workspace and cannot be restored.

Load a saved GUI workspace
- Browse to the desired .RDATA file and upload it to restore the saved workspace.

Save current GUI workspace
- To save the workspace of the current session, use *Filename (must have '.RDATA' extension)* to provide a filename, including the '.RDATA' extension, for the .RDATA file and click *Download GUI workspace* to download the file to your computer. You can load this file later using *Load a saved GUI workspace*. The 'GUI workspace' consists of data that has been imported into the GUI (e.g. predictions, a study area polygon, or validation data) or created using the GUI (e.g. overlaid or ensemble predictions, evaluation metrics, or saved high quality maps) Thus, this data is saved in the downloaded '.RDATA' file and can be loaded back into the GUI. However, user selections (e.g. high quality map parameters) and displayed plots are not saved in the workspace and cannot be restored.

2.2. eSDM GUI Roadmap

<u>Description:</u> This section describes the basic order in which you may use the tabs and functionality of the eSDM GUI. You can click the *Download sample data* button to download a zip file with sample data that you can use within the GUI, or the *Download manual* to download the manual as a PDF. There also is a 'GUI tips' section of notes about using the GUI.

3. Import Predictions Tab

3.1. Import Predictions

<u>Description:</u> Import SDM predictions into the GUI. The *Data file type* pull down menu allows you to specify the input file type, and adjusts the page to display relevant inputs, as described below. See <u>Supported import file types: SDM prediction data</u> for file format requirements.

Import Predictions   —

Data file type

[Excel .csv ▼]

Browse to and open the desired file with the extension '.csv'. This file must have headers. Note that larger files will take longer to load.

Column with longitude data and *Column with latitude data:* Longitude and latitude points must be in WGS 84 geographic coordinates and equally spaced in decimal degrees.
Column with prediction data: Please ensure that missing prediction values are one of the following: 'NA', 'NaN', 'N/A', 'n/a', 'na', 'Null', blank, or a negative number.
Prediction value type: Select "Relative density" if the predictions are probabilities of occurrence.
Column with weight data: Select "N/A - No pixel-level spatial weight data" if the data does not have weight data.

Upload Excel .csv file (.csv extension)   Location of point in grid cell

[Browse... | Sample_predictions_1.csv]   [Center ▼]
        Upload complete

Column with longitude data   Column with latitude data

[Lon ▼]   [Lat ▼]

Column with prediction data   Prediction value type

[PredModel1 ▼]   [Relative density ▼]

Column with weight data (optional)   21672 prediction values were classified as NA
                                  No weight values were classified as NA

[Weight ▼]

[Import predictions]

3.1.1: "Excel csv file"

<u>Description:</u> Use when SDM predictions are in an Excel .csv file with columns consisting of longitude, latitude, prediction, and (possibly) weight values.

Location of point in grid cell
- Specifies the part of each prediction grid cell represented by the longitude and latitude coordinates in the uploaded Excel .csv file. The options are "Center", "Top right", "Top left", "Bottom left", and "Bottom right".

Upload Excel csv file (.csv extension)
- Browse to and select the .csv file that contains the data you wish to import. The longitude and latitude points must be geographic coordinates that are equally spaced in decimal degrees. This file must be a .csv file and have headers.

Column with longitude data
- Select the name of the column with the longitude data, which must be in WGS 84 geographic coordinates (decimal degrees).

Column with latitude data
- Select the name of the column with the latitude data, which must be in WGS 84 geographic coordinates (decimal degrees).

Column with prediction data
- Select the name of the column with the prediction data.

Prediction unit
- Select "Absolute density" if all necessary correction factors have been applied to the density predictions for the applicable study area, and thus provide a true prediction of density. Select "Relative density" if the predictions have **not** had all necessary correction factors applied for the applicable study area, and thus only accurately predict the density relative to the other predictions. Also select "Relative density" if the predictions are probability of occurrence predictions or habitat suitability indices. Select "Abundance" if the prediction values are abundances. If you select "Absolute density" or "Abundance", then the abundance will be displayed in the *Imported Original Predictions table*.

Column with weight data (optional)
- Select the name of the column with the weight data. All weight values must be between zero and one, inclusive. If you do not have weight data for these predictions, then select "N/A – No pixel-level spatial weight". Weight data can be used as pixel-level spatial weights in a weighted ensemble in the 'Create Ensemble Predictions' tab.

Predictions info message
- A message detailing how many of the provided prediction values and (if applicable) weight values will be classified as NA, as well as if any non-NA predictions have NA weight values. A prediction value will be classified as NA if the provided data is one of the following: 'NA', 'NaN', 'N/A', 'n/a', 'na', blank, or a negative number. In addition, if you select "Abundance" for *Prediction unit* a message detailing how the prediction values will be processed will be displayed.

Import predictions
- Click to import predictions from the uploaded Excel .csv file into the GUI.

3.1.2: "GIS raster"
Description: Use when SDM predictions are a band of a raster file. The GUI supports rasters from IMG (.img) and TIFF (.tif) files; it may not be able to properly process rasters from other file types. Note that raster data cannot have spatial-level pixel weight data since the GUI can only read data from one raster band at a time.

Band number of prediction data
- The band number of the prediction data within the uploaded raster file.

Upload raster file (.img or .tif extension)
- Browse to and upload the raster file(.img or .tif extension). The raster can be in any coordinate system. The GUI will display an error message if the selected raster does not have any data at the provided band number or if the GUI was not able to process the uploaded file.

Prediction unit
- See *Prediction unit*

Predictions info message
- See *Predictions info message*

Import predictions
- Click to import predictions from the uploaded raster file into the GUI.

3.1.3: "GIS shapefile"
Description: Use when SDM predictions are stored in a GIS shapefile.

Upload GIS shapefile files
- Browse to and select all files of the desired GIS shapefile. Although they will have different file extensions, these files will all have the same filename (this will be the name of the shapefile in ArcCatalog). An error message will appear if not all of the files of the desired shapefile are selected, or if extraneous files are selected.

Column with prediction data
- See *Column with prediction data*

Prediction unit
- See *Prediction unit*

Column with weight data (optional)
- See *Column with weight data (optional)*

Predictions info message
- See *Predictions info message*

Import predictions
- Click to import predictions from the uploaded files into the GUI.

3.1.4: "GIS file geodatabase feature class"
Description: Use when SDM predictions are in a file geodatabase feature class with at least one associated data column for the predictions values. You cannot import predictions from a personal geodatabase. You must upload the file geodatabase feature class, and then you can finish importing the predictions.

Full path to file geodatabase
- Enter the full path up to and including the file geodatabase that contains the file geodatabase feature class you wish to upload. In the file path, the file geodatabase will act as a folder and will have a '.gdb' extension. Do not put any additional text, such as a '/', after the .gdb extension. You can copy and paste the file path from the top bar of ArcCatalog after navigating to the desired file geodatabase. On a Windows machine, you can also copy and paste the file path from the top bar of the Windows Explorer.

Name of file geodatabase feature class
- Enter the name of the file geodatabase feature class you wish to upload. You can find this name in ArcCatalog, where the file type will be 'File Geodatabase Feature Class'.

Upload feature class
- Click to upload the specified file from the specified path. An error message will appear if the app does not find a file with the given name in the specified file geodatabase.

Column with prediction data
- See *Column with prediction data*

Prediction unit
- See *Prediction unit*

Column with weight data (optional)
- See *Column with weight data (optional)*

Predictions info message
- See *Predictions info message*

Import predictions
- Click to import predictions from the uploaded feature class into the GUI.

3.2. Imported Original Predictions

<u>Description:</u> This window contains a table with summary information about each set of imported predictions, and is also where you select predictions to preview (plot) or remove from the GUI.

3.2.1: *Select original predictions with which to perform an action*

- Select or deselect a set of original predictions by clicking on the row of that set of predictions in the table. Selected predictions are highlighted blue-grey. Depending on the action selection, you may select multiple sets of predictions, but you can only select predictions when *Display additional information* is unchecked. The table has two sets of information that you can toggle using the *Display additional information* checkbox. When *Display additional information* is unchecked, the name of the file that was imported, the names of columns from which prediction and weight data were imported, and the specified prediction unit are displayed. The 'Weight' column is blank if the predictions do not have weight data. When *Display additional information* is checked, the table displays the resolution of the predictions (see below for more details), the number of prediction polygons, the number of prediction polygons with non-NA predictions, the predicted abundance (if applicable), and the spatial range of the predictions.
  - 'Resolution' column: When predictions are imported into the GUI, the GUI attempts to determine the resolution at which the predictions were made, meaning the distance between the centroids of adjacent grid cells. The GUI only attempts to calculate the resolution in the native coordinate system of the imported predictions; for example if predictions were generated on an equal area grid but were imported into the GUI via a shapefile in in a decimal degrees coordinate system, then the GUI would not be able to calculate the resolution.

3.2.2: *Action to perform with selected original predictions*

<u>Description:</u> Select the action you wish to perform with the selected set(s) of predictions. These actions include, for one or more sets of predictions: showing a preview,

downloading a preview, or removing them from the app. The options shown in the *Action option(s)* box depend on the selected action, and are described below.

3.2.2.1: "Plot interactive preview"

Units

- Select "Percentages" to have the colors of predictions depend on the prediction value relative to the other prediction values in the set of predictions. For instance, the color scheme will delineate, among other ranges, the top 2% of prediction values, the prediction values in the top 2% to 5%, and the prediction values in the top 5% to 10% of all of the prediction values. Select "Values" to have the colors of predictions depend on the numerical values of the predictions.

Plot interactive preview

- Click to generate the interactive preview in the 'Interactive Preview' box. This can only be a preview of a single set of predictions.

3.2.2.2: "Plot static preview"

Units

- See *Units*

Plot static preview

- Click to generate the static preview in the 'Static Preview' box. This can be a preview of a single set of predictions or a preview of multiple sets of predictions.

3.2.2.2: "Download static preview"

Units
- See *Units*

Resolution
- The resolution of the downloaded image. It is recommended to use the "High" resolution for plots of multiple sets of predictions.

File format
- The file format in which to download the preview of the selected predictions. The current options are JPEG, PDF, and PNG file types.

File dimensions
- The file format in which to download the preview of the selected predictions. The current options are JPEG, PDF, and PNG file types.

Filename
- The filename of the preview when it is downloaded. The filename is reset to the default if different predictions are chosen to be plotted, or if any of the other inputs in the *Action option(s)* box are changed.

Download
- Click to download the static preview of selected set(s) of predictions with the parameters and filename specified in the *Action option(s)* box.

3.2.2.3: "Remove from GUI"
Remove selected original predictions
- Click to remove the selected set(s) of predictions from the GUI.

4. Overlay Predictions Tab

4.1. Import Study Area Polygon

Description: A study area polygon restricts the region in which the overlay will be performed to the area covered by the study area polygon by clipping the base geometry to the study area polygon. The study area polygon performs the same function as the 'clip feature' in the clip tool in ArcGIS. Use this feature if you have predictions that cover a broad area, but you only want to create ensemble predictions in a specific study area, e.g. if you have marine predictions that span the U.S. West Coast but only want ensemble predictions in the Southern California Bight. Check the checkbox to use a study area polygon, and specify your study area polygon format using the *File type* selection. A message is displayed in blue text when a study area polygon is imported.

Clip the base geometry to a study area polygon in the overlay process
- Check to upload a study area polygon, and uncheck to remove an imported study area polygon.

4.1.1: "Excel .csv"
Upload Excel .csv file (.csv extension)
- Browse to and upload the file with the .csv extension that contains the desired polygon. The file must have headers, the first column must contain the longitude values, and the second column must contain the latitude values. The longitudes and latitudes must be in geographic coordinates (decimal degrees), and the longitudes must be either in the range [-180, 180] or [0, 360]. The study area polygon must be a single polygon, and thus cells with blank or 'NA' values are not permitted. The provided points must form a closed polygon.

4.1.2: "GIS shapefile"
Upload GIS shapefile files
- See *Upload GIS shapefile files*

4.1.3: "GIS file geodatabase feature class"
Full path to file geodatabase
- See *Full path to file geodatabase*

Name of file geodatabase feature class
- See *Name of file geodatabase feature class*

Upload feature class
- See [Upload feature class](#)

4.2. Import Erasing Polygon

Description: An erasing polygon can be used to erase area from the base geometry before the overlay, e.g. if you want to remove land area from predictions in a marine environment. The erasing polygon performs the same function as the '[erase feature](#)' in the erase tool in ArcGIS. Check the checkbox to use an erasing polygon. You can either import a personal erasing polygon, or use the provided erasing polygon. A message is displayed in blue text when an erasing polygon is imported.

Erase area from the base geometry in the overlay process
- Check to import an erasing polygon, and uncheck to remove an imported erasing polygon.

4.2.1: "Use provided erasing polygon"
Description: If you do not have a land polygon of your own but would like to use one to remove land from the base grid, you may import the provided erasing polygon: the low resolution L1 (continents other than Antarctica) and L6 (Antarctica) polygons from the Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) Database. See the [GSHHG website](#) for more information.

Import provided land polygon
- Import the provided, low resolution GSHHG land polygon

4.2.2: "Upload personal erasing polygon"
Description: Specify your erasing polygon format using the *File type* selection. Additional options shown depend on the *File type* selection.

4.2.2.1: "Excel csv file"
Upload Excel csv file (.csv)
- Browse to and upload the file with the .csv extension that contains the desired polygon. The file must have headers, the first column must contain the longitude values, and the second column must contain the latitude values. The longitudes and latitudes must be in geographic coordinates (decimal degrees), and the longitudes must be either in the range [-180, 180] or [0, 360]. Multiple polygons may be demarcated using blank cells or cells

with 'NA' entries. The provided points must form a closed polygon.

4.2.2.2: "GIS shapefile"
Upload GIS shapefile files
- See *Upload GIS shapefile files*

4.2.2.3: "GIS file geodatabase feature class"
Full path to file geodatabase
- See *Full path to file geodatabase*

Name of file geodatabase feature class
- See *Name of file geodatabase feature class*

Upload feature class
- See *Upload feature class*

4.3. Imported Original Predictions

Description: See Description.

	SDM filename	Predictions	Weights	Prediction type
Original 1	Becker_data.csv	Density_preds_2008		Absolute density
Original 2	WW_data_fit.csv	Density_preds_2008		Relative density
Original 3	JVR_preds	Dens2008		Relative density

Imported Original Predictions —

Select original predictions to use as the base geometry: Click on a row in the table below to select or deselect predictions.

☐ Display additional information - NOTE that you can only select or deselect a row when this box is unchecked

Select original predictions to use as the base geometry
- See *Select original predictions with which to perform an action* for how to make selections using a table. The only difference is that you can only select one row in this table because the base geometry must come from a single set of predictions.

4.4. Overlay Predictions

Description: Specify desired overlay options, and then overlay all original predictions onto the base geometry to create overlaid predictions. See Appendix 1: The overlay process for a detailed description of the overlay process, including how the *Overlay option* inputs are used in the overlay process.

Overlay options: study area and erasing polygons
- Import study area and erasing polygons if desired

Overlay options: base geometry
- Choose the set of predictions whose geometry will be used as the base geometry

Overlay Predictions　　　　　　　　　　　　　　　　　　　　　　　　　　　　−

1) Overlay options: study area and erasing polygons

Import these polygons in their respecitve boxes: 'Import Study Area Polygon' and 'Import Erasing Polygon'.

Note that the study area polygon performs the same function as the 'clip feature' in the clip tool in ArcGIS, while the erasing polygon performs the same function as the 'erase feature' in the erase tool .

2) Overlay options: base geometry

Choose the base geometry in the 'Imported Original Predictions' box.

3) Overlay options: coordinate system

The overlay process involves calculating the intersection and area of polygons, and thus the coordinate system in which the overlay is performed will have an effect on the results.
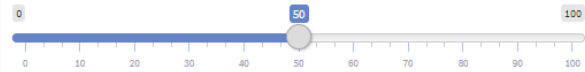
☑ Perform the overlay in the native coordinate system of the specified base geometry

4) Overlay options: percent overlap threshold

Specify the percent overlap threshold

The percent overlap threshold is the minimum percentage of a base geometry polygon that must overlap with overlaid prediction polygons. All base geometry polygons with an overlap percentage less than this threshold will be assigned an overlaid prediction value of 'NA'.
A threshold of '0' means that base geometry polygons will not be assigned a prediction value of 'NA' if they overlap with any original predictions.

0　　　　　　　　　　　　　　　　50　　　　　　　　　　　　　　　　100

0　　10　　20　　30　　40　　50　　60　　70　　80　　90　　100

5) Perform overlay

Reminder: imported study area and land polygons will be used during the overlay process. This process may take several minutes.

[Overlay all predictions onto the specified base geometry]

Overlaid models have been created using the following overlay options:
1) Both a study area polygon and an erasing polygon were used
2) Using the geometry of the 'JVR_preds' SDM as the base geometry
3) In the native coordinate system of the base geometry
4) With a percent overlap threshold of 50%

Overlay options: coordinate system
Perform the overlay in the native coordinate system of the specified base geometry
Description: If this box is checked, all objects involved in the overlay will be projected to the native coordinate system of the base geometry.

"Perform overlay in WGS 84 geographic coordinates"
- The base geometry and all original predictions will be transformed to WGS 84 geographic coordinates for the overlay process. Thus, the following assumptions will be made for all area and overlap calculations: 1) 'Equatorial axis of ellipsoid' = 6378137 and 2) 'Inverse flattening of ellipsoid' = 1/298.257223563. See this article for more details about assumptions made when using WGS 84 geographic coordinates.

"Select predictions with desired coordinate system"
Perform the overlay in the coordinate system of the selected original predictions
- Specify the original predictions whose coordinate system in which you want to perform the overlay.

"Enter numeric EPSG code"
EPSG code
- Enter the numeric EPSG code of the coordinate system in which you want to perform the overlay. You can find valid codes at http://epsg.io/ or http://www.epsg.org/

Overlay options: percent overlap threshold
Specify the percent overlap threshold
- The slider bar allows the user to specify the percent overlap threshold, i.e. the minimum percentage of a base geometry polygon that must overlap with overlaid prediction polygons. All base geometry polygons with an overlap percentage less than this threshold will be assigned a prediction value of 'NA'. A threshold of '0' means that base geometry polygons will not be assigned a prediction value of 'NA' if they overlap with any original predictions.

Perform overlay
Overlay all predictions onto the specified base grid
- Click this button to overlay all sets of original predictions onto the specified base geometry using the eSDM::overlay_sdm() function.

4.5. Base Geometry and Overlaid Predictions Previews

4.5.1: "Base geometry preview"
<u>Description:</u> Click *Preview base geometry* to generate an interactive preview of the base geometry, including imported study area and erasing polygons. You can only generate the preview after you have selected original predictions to use as the base geometry. In addition, the *Clip the base geometry to a study area polygon in the overlay process* cannot be checked if no study area polygon has been imported, and the *Erase area from the base geometry in the overlay process* cannot be checked if no erasing polygon has been imported. If the base geometry was made at a high resolution, then it may appear to be completely black in the preview when zoomed out. In the preview, the base geometry polygons will be outlined in black, the study area polygon will be outlined in red, and the erasing polygon will be filled in tan. In addition, if the any base geometry polygons that span the antemeridian (i.e. 180 decimal degrees or their equivalent in the specified coordinate system) will appear to be split at the antimeridian, even though they are still treated as a single polygon.

4.5.2: "Overlaid predictions preview"
<u>Description:</u> Click *Preview selected overlaid predictions* to generate a preview of the overlaid predictions selected in *Select overlaid predictions to preview*. This can be a preview of a single set of overlaid predictions or a multiplot of multiple sets of overlaid predictions. This feature was designed to allow you to be able to quickly preview the overlaid predictions; if you wish to generate high quality images of the overlaid predictions, use the 'High Quality Maps' tab.

Select overlaid predictions to preview; 'Overlaid' numbers correspond to 'Original' numbers in the table above
- Select the set(s) of overlaid predictions to preview. The sets of overlaid predictions are listed as "Overlaid 1", "Overlaid 2", etc. These numbers correspond to the "Original 1", "Original 2", etc., entries in the far left column

of the table in the *Imported Original Predictions* section, meaning for example that the "Overlaid 1" set of predictions is the "Original 1" predictions overlaid onto the base geometry. You may select and preview multiple sets of overlaid predictions at one time.

Units
- See *Units*

Preview selected overlaid predictions
- Click to generate a preview of the selected overlaid predictions used the specified units.

5. Create Ensemble Predictions Tab

5.1. Overlaid Predictions
<u>Description:</u> Specify the overlaid predictions to use when creating ensemble predictions.

Overlaid Predictions   −

Overlaid predictions were created using the geometry of the 'JVR_preds' SDM as the base geometry. See the 'Overlay Predictions' tab for other overlay information.

Select overlaid predictions to ensemble: All sets of overlaid predictions will be used when creating ensemble predictions. You can select specific sets of overlaid predictions to ensemble when *Create ensemble using a subset...* is checked.

	SDM filename	Predictions	Weights	Prediction type	Resolution	Cell count	Non-NA prediction count	Abundance	Long, lat range
Overlaid 1	Becker_data.csv	Density_preds_2008	NA	Absolute density	10 km	11419	11392	1481	-131, -117, 30, 49
Overlaid 2	WW_data_fit.csv	Density_preds_2008	NA	Relative density	10 km	11419	11373	N/A	-131, -117, 30, 49
Overlaid 3	JVR_preds	Dens2008	NA	Relative density	10 km	11419	11378	N/A	-131, -117, 30, 49

☐ Create ensemble using a subset of the sets of overlaid predictions

Select overlaid predictions to ensemble
- If the *Create ensemble using a subset of the sets of overlaid predictions* checkbox is unchecked, then all overlaid predictions will be used in the ensemble. If the box is checked, then you can click on rows of the table to select or deselect the overlaid predictions you wish to use when creating ensemble predictions. You must select at least two sets of predictions to create an ensemble.

5.2. Create Ensemble Predictions

Create Ensemble Predictions   −

1) Ensemble options: rescaling method
- ○ None
- ● Abundance
- ○ Normalization
- ○ Standardization
- ○ Sum to 1

Abundance to which to rescale predictions

[0]

<u>Description:</u> For each set of overlaid predictions, rescale them so that the predicted abundance is the value entered above

2) Ensemble options: regional weighting
- ☐ Regionally weight overlaid predictions before creating ensemble

3) Ensemble options: ensembling method
- ● Unweighted
- ○ Weighted

Unweighted ensembling method: Calculate the mean of all corresponding predictions

4) Create ensemble

[Create ensemble]

5.2.1: Ensemble options: rescaling method
<u>Description:</u> If one set of predictions has a different prediction unit or was made using a much larger numerical scale than the other predictions, then it would contribute disproportionately to a basic, unweighted ensemble. This ensemble option allows you to rescale predictions before ensembling them, therefore allowing you to create realistic ensembles of predictions with different units or numerical scales. Unless the prediction unit of all of the selected overlaid predictions are either 'Absolute density' or 'Abundance', then the predictions likely need to be rescaled. An exception would be if you know that all of the predictions were created using the same numerical scale. If you select a rescaling option other than "None", then all of the selected overlaid predictions will be rescaled

before they are regionally weighted or the ensemble is created.

5.2.1.1: "None"
Description: Overlaid predictions will not be rescaled before the ensemble is created. This option is only recommended if all prediction units are one of 'Absolute density' or 'Abundance'.

5.2.1.2: "Abundance"
Description: Overlaid predictions will be rescaled so that the predicted abundance for each set of overlaid predictions is the value entered in *Abundance to which to rescale predictions*. This value must be greater than zero.

5.2.1.3: "Normalization"
Description: Overlaid predictions will be rescaled so that each set of overlaid predictions has a range of [0, 1]. Normalization uses the following formula, where X is the current set of overlaid predictions:
$X_{new} = (X - X_{min}) / (X_{max} - X_{min})$

5.2.1.4: "Standardization"
Description: Overlaid predictions will be rescaled so that each set of overlaid predictions has a mean (μ) of 0 and a standard deviation (σ) of 1. Standardization uses the following formula, where X is the current set of overlaid predictions:
$X_{new} = (X - \mu) / \sigma$

5.2.1.5: "Sum to 1"
Description: Overlaid predictions will be rescaled so that they sum to one for each set of predictions. Sum to 1 uses the following formula, where X is the current set of overlaid predictions:
$X_{new} = X / sum(X)$

5.2.2: Ensemble options: regional weighting
Description: Regionally weight overlaid predictions by importing polygons, hereafter referred to as 'weight polygons', and specifying the weight that will be applied to selected overlaid predictions that have a specified percentage of their area within that weight polygon. You can apply each weight polygon to one or more set(s) of overlaid predictions. The assigned weight polygons are summarized in a table at the bottom of the regional weighting window. You may also preview the assigned weight polygons for each set of overlaid predictions.

Overlaid predictions to which to assign polygon weights
- Select the set(s) of overlaid predictions to which to assign the imported weight polygon.

Use 'NA' as weight and *Weight*
- The value that the overlaid predictions within the weight polygon will be multiplied by. If the *Use 'NA' as weight* checkbox is checked, then the overlaid

predictions within the weight polygon will be set as 'NA'.

Percentage of prediction polygon…
- Specify the percentage of an overlaid prediction polygon that must be overlapped by the weight polygon for the value of that prediction polygon to be multiplied by the weight.

Assign weight polygon to selected predictions
- Click this button to assign the imported weight polygon to the selected set(s) of overlaid predictions. These changes will be reflected in the summary table.

Remove selected weight polygons
- Click to remove the weight polygons selected in *Select assigned weight polygon(s) to remove* from the GUI. These changes will be reflected in the summary table.

5.2.2.1: "Excel .csv"
Upload Excel .csv file (.csv extension)
- See *Upload Excel csv file (.csv extension)*

5.2.2.2: "GIS shapefile"
Upload GIS shapefile files
- See *Upload GIS shapefile files*

5.2.2.3: "GIS file geodatabase feature class"
Full path to file geodatabase
- See *Full path to file geodatabase*

Name of file geodatabase feature class
- See *Name of file geodatabase feature class*

Upload feature class
- See *Upload feature class*

5.2.3: Ensemble options: ensembling method
Description: Ensembles can be created using a weighted or unweighted average of the (possibly) rescaled predictions. Weights can be the evaluation metrics of the overlaid predictions or assigned by the user to the overlaid predictions at two resolutions: the entire study area or individual polygons within the overlaid predictions. Overlaid predictions that are 'NA' are ignored when creating the ensemble.

5.2.3.1: "Unweighted"
Description: The predictions will be averaged together via an unweighted average, i.e. simple mean.

5.2.3.2: "Weighted"

Description: The selected overlaid predictions will be averaged together via a weighted average. You can specify the weighting method via your *Weighted ensembling method* selection. The options displayed and the process of getting the different weights are described below.

5.2.3.2.1: "Manual entry"
Description: Weight each set of overlaid predictions by a user-provided value.

Ensemble weights
- The values entered into the text box will be used as the weights in the weighted average of the predictions. The first value will be applied to the first set of overlaid predictions selected in *Select overlaid predictions to ensemble*, the second value will be applied to the second set of selected overlaid predictions, and so on. These values must be numbers greater than or equal to zero and separated by a single comma. Spaces after the comma(s) are optional.

5.2.3.2.2: "Evaluation metric"
Description: Weight each set of selected overlaid predictions by an evaluation metric value. To use this weighting method, you must first use the 'Evaluation Metrics' tab to calculate the metric you wish to use as a weight for all of the overlaid predictions you wish to use in the ensemble.

Metric to use for weights
- This section will display the metric(s) that you have calculated for all of the overlaid predictions you have selected to use in the ensemble. Select the metric that you wish to use for the weights. The table shows the metric calculated for each of the overlaid predictions selected to be in the ensemble ('AUC', 'TSS", or 'RMSE'), as well as their weights relative to the largest metric ('Relative weights'). The number in the 'Relative weights' column is what is actually used as the weight.

5.2.3.2.3: "Pixel-level spatial weights"
Description: Weight each individual overlaid prediction value. 'Pixel-level spatial weights' refer to the weight data specified via the *Column with weight data* selection when initially importing the predictions into the GUI. The individual predictions are multiplied by their corresponding weight. At least one set of overlaid predictions must have weight data to use this method, and the table shows which set(s) of overlaid predictions have pixel-level spatial weights. Predictions with weights of 'NA' will be ignored when creating the ensemble, while the set(s) of predictions that do not have pixel-level spatial weights will still be included in the ensemble and will have a 'weight' of 1.

5.2.4: Ensemble options: create ensemble
Description: Click *Create ensemble* to create an ensemble of the selected overlaid

predictions using the specified ensemble options.

5.3. Created Ensemble Predictions

Description: This window contains a table with summary information about each set of created ensemble predictions (see summary information abbreviations below), and can be used to select one or more ensembles to create or download a plot preview, remove from the GUI, or calculate the predicted abundance of the ensemble. You may select ensemble predictions by clicking on the corresponding row in the table; this will highlight the predictions row in blue-grey. Depending on the action, you may select multiple ensemble predictions at one time if desired.

This section has the same actions to perform and action option(s) as those described in Imported Original Predictions. In addition, you can calculate the predicted abundance for the selected ensemble(s) if they were all created with an appropriate rescaling method.

5.3.1: Ensemble summary table
- Overlaid predictions used: Either 'All', indicating that all overlaid predictions were used in the ensemble, or numbers indicating that a subset of the overlaid predictions were used in the ensemble. The numbers correspond to the 'Overlaid #' number.
- Rescaling method: "Abundance: ###" indicates that the abundance rescaling method and a value of ### was used before creating this ensemble.
- Ensembling method: Either "Weighted", "Unweighted", or "Regionally weighted + …" where … is either "weighted" or "unweighted"; indicates which top-level ensembling method was used and whether regional weighting was performed before ensembling.
- Weights: Blank if the ensembling method was "Unweighted"; If the ensembling method was one of the weighted methods then the actual weight applied to each set of predictions is displayed.

6. Evaluation Metrics Tab

6.1. Select Predictions to Evaluate

<u>Description:</u> Click on the rows of the original, overlaid, and/or ensemble predictions to select those for which you wish to calculate evaluation metric(s).

Select Predictions to Evaluate

Select predictions to evaluate: Click on row(s) in the table(s) below to select or deselect predictions.

	SDM filename	Predictions	Weights
Original 1	Becker_data.csv	Density_preds_2008	
Original 2	WW_data_fit.csv	Density_preds_2008	
Original 3	JVR_preds	Dens2008	

	SDM filename	Predictions	Weights
Overlaid 1	Becker_data.csv	Density_preds_2008	
Overlaid 2	WW_data_fit.csv	Density_preds_2008	
Overlaid 3	JVR_preds	Dens2008	

	Overlaid predictions used	Rescaling method	Ensembling method	Weights
Ensemble 1	All	Abundance: 858	Regionally weighted + weighted	0.992, 0.814, 1

6.2. Import Validation Data

Description: Import validation data into the GUI. You can specify the type of input file via the *Validation data file type* menu, and you can specify the type of validation data via the *Validation data type* menu. The GUI adjusts the page to display relevant inputs depending on the file type and data type selected. A blue 'Validation data successfully imported' message will be displayed if validation data has been imported.

Import Validation Data   −

Note that the GUI can only store one set of validation data at a time; importing validation data will overwrite any previously imported validation data

Validation data file type   Validation data type

◉ Excel .csv   ◉ Count (numerical)
○ GIS shapefile   ○ Presence/absence
○ GIS file geodatabase feature class

Browse to and open the desired file with the extension '.csv'. This file must have headers. Note that larger files will take longer to load.

The Excel .csv file must have columns with the longitude and latitude coordinates of the data, in addition to the column with validation data. The longitude and latitude coordinates are assumed to be WGS 84 geographic coordinates.

Upload Excel .csv file (.csv extension)

| Browse... | No file selected |

Validation data successfully imported

6.2.1: "Excel .csv"
Upload Excel .csv file (.csv extension)
- See *Upload Excel csv file (.csv extension)*

Select, in this order, the longitude, latitude, and validation data column for the uploaded .csv file
- When you click on the input box, the GUI will display the headers of all of the columns in the uploaded file. Select, in order, the headers that correspond to the longitude column, the latitude column, and the column that contains the validation data. You can delete selected column names from the input box by selecting them and clicking your computer's delete button. If *Validation data type* is "Counts (numerical)", then the data column must consist of numbers.

6.2.2: "GIS shapefile"
Upload GIS shapefile files
- See *Upload GIS shapefile files*

6.2.3: "GIS file geodatabase feature class"
Full path to file geodatabase
- See *Full path to file geodatabase*

Name of file geodatabase feature class
- See *Name of file geodatabase feature class*

Upload feature class
- See *Upload feature class*

Validation data type
- Specify the type of validation data. If "Count (numerical)", all values in the specified column must be numbers.

Select the validation data column for the uploaded GIS file
- Only displayed when *Validation data file type* is "GIS shapefile" or "GIS file geodatabase feature class".
- Select the name of the data column in the shapefile that contains the validation data. If *Validation data type* is "Counts (numerical)" then the data column must consist of numbers.

Select presence code(s)
- Only displayed when *Validation data type* is "Presence/absence".
- You must specify the code(s) that specify presence observations. You must classify all of the codes as either presence or absence codes between this input and *Select absence code(s)*, and you cannot classify a code as both a presence and an absence code.

Select absence code(s)
- Only displayed when *Validation data type* is "Presence/absence".
- You must specify the code(s) that specify absence observations. You must classify all of the codes as either presence or absence codes between this input and *Select presence code(s)*, and you cannot classify a code as both a presence and an absence code.

Import specified validation data
- Import the validation data into the GUI.

6.3. Calculate Metrics

Description: Get information about the imported validation data, select the metrics you wish to calculate, and then calculate the metrics. See Appendix 2: Calculating evaluation metrics for more details about how evaluation metrics are calculated.

Calculate Metrics   —

Validation data info

Filename	PA_LT5km_HR90.csv
Data type	Presence/absence
Number of presence points	599
Number of absence points	25396

See 'Metrics Descriptions and References' section below for metric information

Metric(s) to calculate

☑ AUC

☑ TSS

Calculate metrics

Metrics calculated

Metric(s) to calculate
- Check the respective boxes to calculate area under the receiver operating curve (AUC), true skill statistic (TSS), and/or root mean squared error (RMSE). You may only calculate RMSE if the *Validation data type* of the imported validation data is "Count (numerical)". See Fielding and Bell 1997 for more information about AUC, Allouche *et al.* 2006 for more information about TSS, and this page for more information about RMSE.

Calculate metrics
- Click this button use the imported validation data to calculate the selected metrics for the selected set(s) of predictions.

6.4. Metric Results

<u>Description:</u> Displays the calculated metrics of the specified predictions and information about predictions and validation data overlap. You may also download the metric values.

Metric Results   —

Predictions	AUC	TSS
Overlaid 1	0.865	0.583
Overlaid 2	0.710	0.337
Overlaid 3	0.872	0.639

⬇ Download metrics

Note: The downloaded Excel csv file will have both metric values and information about each set of predictions. Ensemble predictions have different information than original and overlaid predictions, and thus if evaluation metrics have been calculated for both ensemble predictions and original and/or overlaid predictions then some column headers will be formatted as 'Original+Overlaid info name/Ensemble info name'.

The predictions being evaluated each had 0 validation points, respectively, that fell on the boundary between two or more prediction polygons; the predictions from these polygons were averaged for the evaluation

Download metrics

- Click to download an Excel .csv file that contains the metric value(s) and information about the predictions for which the metrics were calculated. Because ensemble predictions have different information than original and overlaid predictions, if evaluation metrics have been calculated for both ensemble predictions and at least one of original and overlaid predictions, then some column headers will be formatted as 'Original+Overlaid info name/Ensemble info name'.

6.5. Metric Descriptions and References

<u>Description:</u> Provides additional information about the metrics (AUC, TSS, and RMSE).

Metric Descriptions and References   —

- ⦿ Area under the curve (AUC)
- ○ True skill statistic (TSS)
- ○ Root mean squared error (RMSE)

AUC decription: The area under the receiver operating curve (AUC) of the ROCR plot is a threshold independent metric that evaluates the percentage of the time a random selection from the positive group will have a score greater than a random selection from the negative class (Deleo 1993).

See Fielding and Bell 1997 for more information.
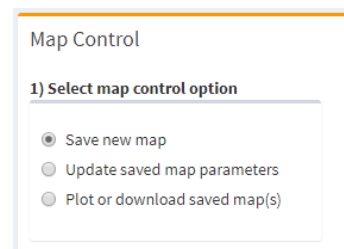
7. High Quality Maps Tab

7.1. High Quality Maps

Description: Display window for plotting high quality maps. The window resizes itself based on the plot dimensions specified by the user.

High Quality Maps
High quality map(s) will be displayed here
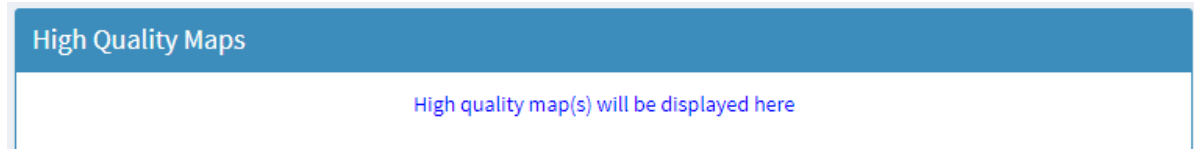
7.2. Map Control

Description: Each of the three map control sections has three displays: 1) the map control selection, 2) a table from which to select the predictions you want to save or a saved map you wish to use, and 3) a section with final inputs and buttons to perform. You always need to make your selection in the table before performing the action

Map Control

1) Select map control option

- Save new map
- Update saved map parameters
- Plot or download saved map(s)

7.2.1. Map control option: "Save new map"

2) Select predictions and specify parameters below

Select predictions to map: Click on a row in the table(s) below to select or deselect predictions. Map parameters will appear after you select a set of predictions.

	SDM filename	Predictions	Weights
Original 1	Becker_data.csv	Density_preds_2008	
Original 2	WW_data_fit.csv	Density_preds_2008	
Original 3	JVR_preds	Dens2008	

	SDM filename	Predictions	Weights
Overlaid 1	Becker_data.csv	Density_preds_2008	
Overlaid 2	WW_data_fit.csv	Density_preds_2008	
Overlaid 3	JVR_preds	Dens2008	

	Overlaid predictions used	Rescaling method	Ensembling method	Weights
Ensemble 1	All	Abundance: 858	Regionally weighted + weighted	0.992, 0.814, 1

3) Specify map ID and save map

Saved map ID

Map ID

Note that most plot parameters below (including loaded additional objects) will stay the same unless changed by user, even when a different set of predictions is selected. Thus, be sure to check the parameters before saving a new map

Save map

Select predictions to map

- Select the set of predictions to map by clicking on a row in the table(s). You can only save one map at a time, and thus you can only select one row at a time. The selected set of predictions will be highlighted blue-grey.

Saved map ID

- The identifier text for the map being saved. This ID must be unique, i.e. it cannot be the same as the ID of any already-saved maps. The ID can be made up of letters, numbers, symbols, and/or spaces

Save map

- Click this button to save the map of the selected set of predictions with the specified parameters and ID.

7.2.2. Map control option: "Update saved map parameters"

Description: This section allows you to either remove a saved map or update the parameters of a saved map. If you select a map and click *Update saved map parameters*, then a window will appear in which you can navigate to the parameter you wish to update and update it. Note that you cannot update some parameters, such as the coordinate system of the map, or add additional objects to the map in this window. In the update window, be sure to click *Save parameter* after updating each parameter. When you are finished, click *Done – save the updated parameters* to finish updating the map parameters and close the window. Click *Cancel and discard the updated parameters* if you wish to close the window without saving the updated parameters.

2) Select saved map to update	3) Update saved map parameters
Select saved map to update Click on a row in the table below to select or deselect a saved map. Then click the 'Update saved map parameters' button in the next box.	Update saved map parameters Remove saved map

Predictions ⇕	ID ⇕
Ensemble 1	Ensemble 1a
Ensemble 1	Ensemble 1b-valid data

Select saved map to update
- Select the saved map for which to update the parameters, or that you wish to remove for the GUI.

Update saved map parameters
- Select a saved map and click this button to open a window where you can update the parameters of the selected map.

Remove saved map
- Select a saved map and click this button to remove that map from the GUI.

7.2.3. Map control option: "Plot or download saved map(s)"

<u>Description:</u> Select one or more saved maps to plot or download, specify the desired dimensions, and then plot or download.

Select saved map(s) to plot or download
- Select the saved map(s) you wish to plot or download. Maps will be plotted left to right, top to bottom, in order you select them.

Number of rows and *Number of columns*
- Specify how you want the selected map(s) displayed, meaning the number of rows and columns that they will be in. The product of these two numbers must be at least the number of selected saved maps.

Plot width (inches) and *Plot height (inches)*
- The size, in inches, of either the plot displayed in the 'High Quality Maps' window or in the downloaded file.

Plot map(s)
- Click to plot the selected saved map(s). Downloading a plot may take several minutes.

Resolution
- The resolution of the downloaded plot. It is recommended to use the "High" resolution for most plots.

File format
- The file format in which to download plot. The current options are JPEG, PDF, and PNG file types.

Filename
- The filename of the file when it is downloaded. The filename is reset to the default if different maps are selected to be plotted, or if *Resolution* or *File format* are changed.

Download map(s)
- Click this button to download a plot of the selected map(s) with the specified dimensions. Downloading a plot may take several minutes.

7.3. Map Parameters – Section 1

Description: Map parameters for high quality maps. This section is only displayed when "Save new map" is selected and a selection has been made in *Select prediction to map*. Parameters that depend on the selected predictions update each time you make a selection.

7.3.1. Map coordinate system and range

Description: Select the map's coordinate system and specify the map range. If the specified coordinate system is degree-based, then map range values must be entered as decimal degrees with the range [-180, 180] for the longitude values and [-90, 90] for the latitude values. If the specified coordinate system is not degree-based, then the map range values must be entered as the unit of the specified coordinate system. The exception is if the selected set of predictions spans the antemeridian (180 decimal degrees); then the longitude values must be in the equivalent of [0, 360] decimal degrees. If the selected set of predictions spans the antemeridian, then a red message stating "The selected predictions span the antemeridian…" is displayed.

Map coordinate system
Description: The same user choices and inputs as [Overlay options: coordinate systems](#).

Longitude minimum and *Longitude maximum*
- The left- and right-most limits of the map, respectively. These values have the same units as the specified coordinate system. They must be in the equivalent of the range [-180, 180] decimal degrees, or the equivalent of [0, 360] if the selected set of predictions spans the antemeridian (if this is the case, a red message stating "The selected predictions span the antemeridian…" is

displayed).

Latitude minimum and *Latitude maximum*
- The bottom- and top-most limits of the map. These values have the same units as the specified coordinate system, and they must be in the equivalent of the range [-90, 90] decimal degrees.

7.3.2. Background color and prediction color scheme

Description: Specify the background color, color of predictions with NA values, and the prediction color scheme. The prediction color scheme consists of the units with which to color-code the predictions by, the color palette and, if applicable, the number of colors 'bins' into which to divide the predictions. The *Color scheme preview* is a preview of the specified color palette with the specified number of colors on the right side of the box. The highest number in the preview corresponds to the first color bin (highest 2% or largest value), and so on.

Click to select background color
- Click to select the background color of the map. This color will be displayed everywhere not covered by a prediction, legend, or additional object, including in the inner margin area.

Make NA predictions transparent
- Uncheck this box to color the predictions within the selected set of predictions. If you uncheck this box and none of the selected set of predictions have NA values, then a message is displayed rather than the color selection input. Otherwise, you can specify the NA prediction color using *Click to select color of NA predictions*.

Color-code predictions by:
- "Relative percentage": the predictions will be split into 10 bins: highest 2%, 2% - 5%, 5% - 10%, 10% - 15%, 15% - 20%, 20% - 25%, 25% - 30%, 30% - 35%, 35% - 40%, and lowest 60%.
- "Numerical value": the predictions will be split into *Number of colors* bins, where the bins are equal size (and thus likely not the equal number of objects within each bin).

Color palette
- Select the color palette that you wish to use to color the predictions. The palette names are formatted as "'R package the palette came from': 'Name of color palette'". The 'Default: blue to white to red' color palette did not come from an R package. Some of the palettes have requirements for the number of colors.
- If "Relative predictions" is selected, then the palette must have ten colors.

Number of colors
- Enter the number of colors that you wish to use, and thus the number of bins

into which the predictions will be split. Some of the palettes have a min and max number of colors; if those palettes are selected, then there is a '(Min: #; Max: #)' text display after the *Number of colors* label.
- If "Relative predictions" is selected, this input is not displayed.

7.3.3. Legend

Description: Specify the legend details. Note that tmap, the R package that the GUI leverage to produce the high quality maps, will automatically resize the legend text size if the text size is too big for the given legend width, legend location, and plot width. If running the GUI locally, a message will be displayed in the RStudio console if this happens.

Include legend with the map
- Check this box to include a legend in the map

Legend location
- Whether legend is inside or outside the map frame. The entire image will still take up the same amount of size, so putting the legend outside of the map frame and not changing the plot dimensions may cause the map to be smaller.

Legend position
- Check this box to include a legend in the map

Legend width
- Only displayed when *Legend location* is "Outside map frame"
- The width is the horizontal proportion of the plot window taken up by the legend. For example, if this value is 0.2, it means the legend will take up 20% of the width of the plot window. Increase this value to increase the size of the

Legend text size
- Relative size of the text within the legend. See legend description for when text size might be automatically changed.

Include black frame around legend
- Check this box to include a black frame around the legend

Legend labels: number of decimals
- Only displayed when *Color-code predictions by:* is "Numerical value"
- The number of decimal places displayed for values in the legend

7.4. Map Parameters – Section 2

Description: Map parameters for high quality maps, continued.

7.4.1. Title, axis labels, and margins

Map title
- The title of the map, which will be displayed at the top of the center of the image. This title may be only one line. Leave this box blank if you do not wish to have a title.
- Note that if the legend is placed outside the map, the title will be centered over the combined map + legend image.

X-axis label
- The label of the x-axis, which will be displayed at the bottom of the map. Leave this box blank if you do not wish to have an x-axis label.

Y-axis label
- The label of the y-axis, which will be displayed on the left of the map. Leave this box blank if you do not wish to have an y-axis label. This label will be displayed vertically.

Title size
- The size of the title, relative to one. For instance, a size entry of 1.1 makes the title slightly bigger than the default size, while a size entry of 0.4 makes the title much smaller than the default size.

Axis label size
- The size of the axis labels, relative to one. For instance, a size entry of 1.1 makes the labels slightly bigger than the default size, while a size entry of 0.4 makes the labels much smaller than the default size.

Inner margin - (bottom, left, top, right)
- The size of the inner margins for the four sides of the map, i.e. the distance between the map outline and the map itself. Larger numbers will make the margins bigger, and smaller numbers will make them smaller. Values of 0 mean there will be no space between the map outline and the map image.

Outer margin
- The size of the outer margin for the four sides of the map, i.e. the distance between the map outline and outside of the plot window. Larger numbers will make the margins bigger, and smaller numbers will make them smaller. Values of 0 mean there will be no space between the map outline and the plot window.
- This parameter is particularly useful if you put the grid line coordinate labels outside the frame; increasing this parameter will allow you to increase the label size without running out of room for the label in the plot window.

7.4.2. Coordinate grid lines and labels

Description: Use grid lines and labels to show the region your map is depicting. tmap, the package that the GUI leverages for the high quality maps section only supports coordinate grid lines rather than tick marks.

Include coordinate grid lines
- Check the box to include coordinate grid lines on the map. This box must be checked to include coordinate labels

Longitude grid line start and *Longitude grid line interval*
- The left-most vertical grid line, i.e. the grid line with the minimum longitude value, you want displayed on the map and the interval between the longitude grid lines, respectively. These values will have the same units and appropriate range specifications as the *map range* longitude values.

Latitude grid line start and *Latitude grid line interval*
- The bottom-most horizontal grid line, i.e. the grid line with the minimum latitude value, you want displayed on the map and the interval between the latitude grid lines, respectively. These values will have the same units and appropriate range specifications as the *map range* latitude values.

Grid line width
- The relative width of the grid lines; larger numbers mean thicker grid lines and vice versa.

Grid line transparency
- The transparency, i.e. the alpha value, of the grid lines. A value of 1 means the grid lines will be solid, while a value of 0 means they will be completely transparent.
- If you only wish to display the coordinate labels and not the grid lines, you can set this parameter to 0.

Click to select color for coordinate grid lines
- Click to select the color of the grid lines.

Include coordinate labels
- Check this box to include coordinate labels on the map. Note that this will only be an option is *Include coordinate grid lines* is also checked. See *Grid line transparency* for how to display coordinate labels without the grid lines

Coordinate label location
- Whether the coordinate grid labels will be inside or outside the frame. If outside the frame, you may need to increase the outer margins to display the entire numbers.

Coordinate label size
- The relative size of the coordinate labels; larger numbers mean larger labels and vice versa.

7.4.3. Additional objects (points or polygons)
Description: Include other objects in your high quality map(s). Common examples include previously imported objects (study area polygon, erasing polygon, validation data) or other points or polygons, e.g. presence-only points or possible protected areas.

Additional objects (points or polygons)

☑ Include additional objects in map

Uncheck the above checkbox to remove all loaded additional objects

Include additional objects in map
- Check to include the study area and/or land polygon in the map. Uncheck this box to remove/delete all loaded additional objects

7.4.3.1: Loaded additional object(s)
Description: This table summarizes the currently loaded additional objects, and thus the additional objects that would be included with the saved map if you clicked *Save map*. Click on the table to select a row to either update the parameters of an additional object or to remove said object.

Loaded additional object(s)

Select loaded object to update its parameters or remove it from the GUI

Object	Object type	Draw order
Study area polygon	Polygon	In front of SDM
Validation data points	Point	In front of SDM
NorCal_poly.csv	Polygon	Behind SDM

Update parameters of selected object

Remove selected object

Update parameters of selected object
- Click to open an update window where you can update most of the additional object parameters. Be sure to click *Save parameter* after updating each individual parameter, and *Done – save the updated parameters* when you are all finished. Or click *Cancel and discard the updated parameters* to cancel any updates made in the window.

Remove selected object
- Click to remove/delete the selected additional object.

7.4.3.2: New additional object
Description: Either load objects previously imported into the GUI (study area polygon, land polygon, or validation data) or import new additional objects.

Additional object source
- Option 1: Select a previously imported object (study area polygon, land polygon, or validation data).
- Options 2: Select "Upload new object" and upload an Excel .csv file, a shapefile, or a file geodatabase feature class. These files have the same file format requirements as files for *importing an erasing polygon*, with the exception that if a .csv file contains points then it can have no rows with 'NA' coordinates.

Object type
- If *Additional object source* is a previously imported object, this input is ignored because those objects already have a type.
- If *Additional object source* is "Upload new object", then specify whether the data is point data or polygon data. If point data and you are uploading a .csv file, the file cannot have any 'NA' coordinates. Objects from GIS files already have defined geometries; the GUI will throw an error if the wrong object type is selected.

Color 1 (left-most color input) and *Color 2* (right-most color input)

- Check the *Make ___ transparent* box to remove the color input and make the specified feature transparent.
- If *Object type:* is "Point(s)": Only *Color 1* will be displayed and will be for the color of the points. The only exception is when *Additional object source* is "Validation data…", at which point *Color 1* will be for presence points and *Color 2* will be for absence points.
- If *Object type:* is "Polygon(s)": Both inputs will be displayed. *Color 1* will be for the polygon fill color and *Color 2* will be for the polygon border color.

Point type or *Line type of polygon borders*

- If *Object type:* is "Point(s)": Select the symbol to be used for all of the points for this object. Symbols can be previewed here.
- If *Object type:* is "Polygon(s)": Select the line type of all of the polygon borders for this object. Line types can be previewed here.

Point size or *Line width of polygon borders*

- If *Object type:* is "Point(s)": The relative size of the points for this object.
- If *Object type:* is "Polygon(s)": The relative width of the polygon borders for this object.

Load additional object

- Click to load (or import if you uploaded a new object) the specified additional object and its parameters. No 'Object successfully loaded' message will appear below the button, but the object will appear in the 'Loaded additional object' table when it is loaded.

8. Export Predictions Tab

8.1. Select Predictions to Export

Description: Click on the row of the original, overlaid, or ensemble predictions to select the set of predictions you wish to export. You can only export one set of predictions at a time, and thus may only select one row at a time. The selected set of predictions will be highlighted blue-grey.

8.2. Export Predictions

Description: Specify the format in which to export the predictions, the coordinate system in which to export the predictions, and the filename of the downloaded predictions. The *Format in which to export predictions* pull down menu allows you to specify how you wish to export the predictions and adjusts the page to request relevant additional details, as described below.

8.2.1: Export options: file format

    8.2.1.1: "Excel .csv file"

    Description: The centroid of each prediction polygon is calculated, and the exported .csv file consists of columns with the longitude and latitudes of these centroids, as well as the prediction and (if applicable) weight values for each of those points. If the predictions were clipped or had any area erased from them, then the centroids might not be equally spaced and you might not be able to re-import this data from the Excel .csv file back into the GUI.

    8.2.1.2: "GIS shapefile"

    Description: Predictions will be exported as polygons with the prediction and (if applicable) weight value for each polygon. The GUI will produce a zip file with the name 'eSDM_shpExport.zip' that will contain the various shapefile files.

    8.2.1.3: "KML or KMZ file"

    Description: Within the .kml or .kmz file, predictions will be represented as polygons with a red outline. Currently you cannot color-code the polygons by density value. The polygons will have their respective prediction and weight values as descriptions.

8.2.2: Export options: coordinate system
Description: The same user choices and inputs as Overlay options: coordinate systems.

Export predictions with longitude coordinates in a range equivalent to [0, 360] decimal degrees rather than [-180, 180] decimal degrees

- Check this checkbox to export the selected predictions in a range equivalent to [0, 360] decimal degrees. This feature is meant for predictions that span the antemeridian (180 decimal degrees), although it does make the exporting process take longer.

8.2.3: Export options: filename

Filename

- The name of the downloaded file with the exported predictions. The extension must match the extension required by the export format ('.csv' for "Excel .csv file", '.shp' for "GIS shapefile", '.kml' for "KML or KMZ file" and "Export as KML", and ".kmz' for "KML or KMZ file" and "Export as KMZ").

8.2.4: Export predictions

Export predictions

- Click this button to download the selected predictions in the specified coordinate system and specified export format.

9. Manual Tab

<u>Description:</u> Depending on your browser settings, the manual will either be displayed within the GUI or displayed in a separate window that appears. If the manual is displayed within the GUI, then you can download the manual from this tab by clicking the PDF viewer download button. You likely will have to resize the manual display; one way to do so is to click the 'Fit to page' button in the PDF viewer, which is above the '+' and '-' buttons.

If the manual appears in a separate window and you close it, you will not be able to access the manual again during that session. If this happens, you can click the *Download manual* button in the 'eSDM GUI Roadmap and Load or Save Workspace' tab.

10.   Appendix

10.1.   Appendix 1: The overlay process

The GUI transforms all of the original predictions, which are displayed in the table in the 'Imported Original Predictions' box in the 'Overlay Predictions' tab, to the coordinate system specified in *Overlay options: coordinate system*. The geometry of the predictions selected in *Select original predictions to use as the base geometry* is the base geometry. If applicable, the GUI clips the base geometry with the study area polygon and erases area from it with the erasing polygon, e.g. to specify a study area or erase land from marine predictions. The GUI then uses the overlay function, overlay_sdm, to overlay SDM predictions onto a single base geometry so all predictions have the same spatial resolution and coordinate system, as illustrated in Fig. 1.

The overlay function intersects the prediction polygons from an original SDM with the prediction polygons from the user-selected base geometry (i.e. base geometry polygons). It then calculates the percentage of each base geometry polygon that overlaps with these intersected polygons, ignoring intersected polygons that have missing values when calculating this percentage. If this percentage meets or exceeds the user-specified percent overlap threshold, the function calculates the overlaid prediction as an area-weighted average of the predictions in the intersected polygons. Otherwise, the function assigns that base geometry polygon an overlaid prediction of 'NA'. If the original predictions are densities, the overlay function calculates the abundance within each intersected polygon as the product of the original prediction and the area of the intersected polygon. It then uses the sum of the abundances and areas of the intersected polygons within each base geometry polygon to calculate the overlaid density. The function also uses the same steps for calculating the overlaid spatial pixel-level weight values.

If any of the original SDM predictions being overlaid have the same geometry as the base geometry, the GUI can simply match up the base geometry and original predictions by prediction polygon. These 'overlays' are much faster than using the overlay_sdm function.
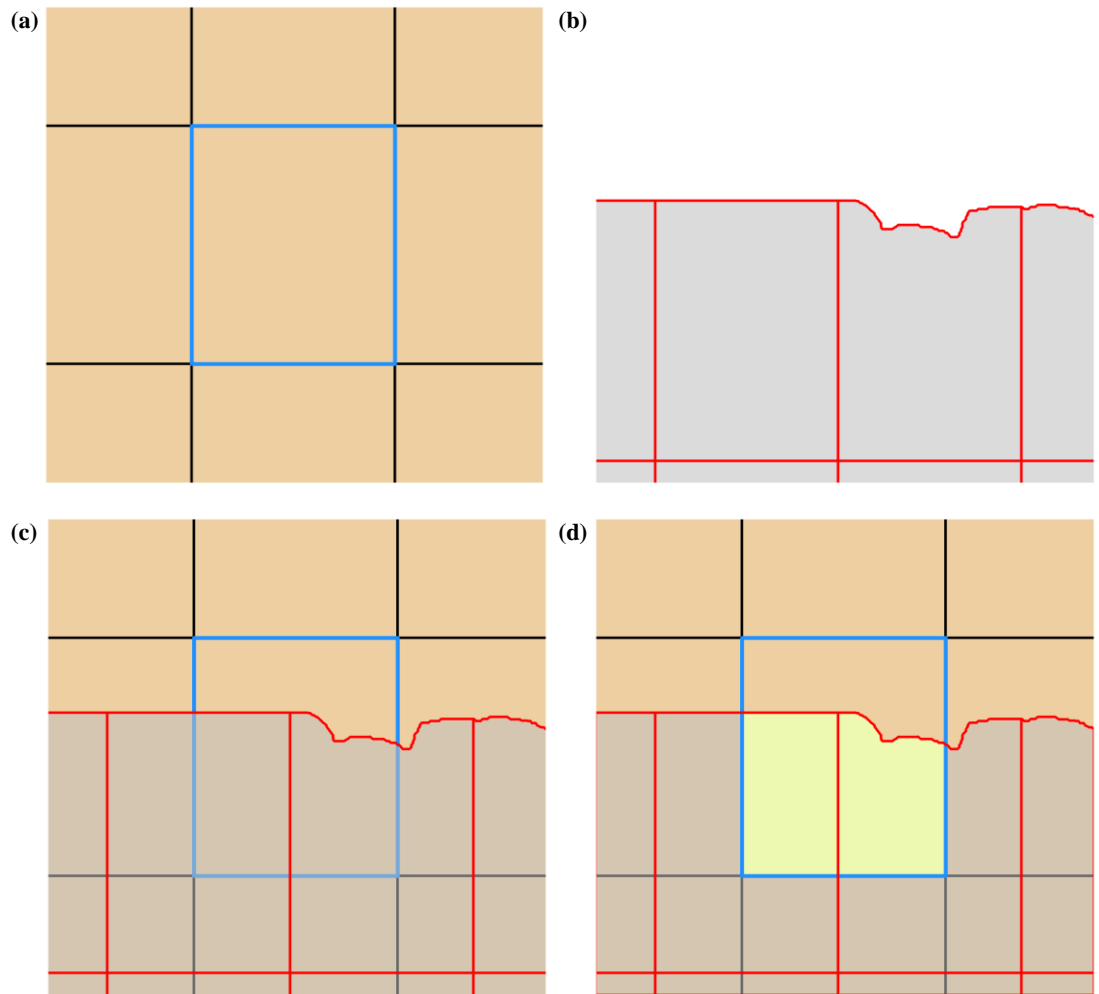
Fig. 1. (a) The base geometry, with the blue outline indicating the current base geometry polygon. (b) The geometry of the SDM predictions being overlaid. (c) The SDM predictions overlaid onto the base geometry. (d) The SDM predictions overlaid onto the base geometry the intersection between the overlaid polygons and the current base geometry polygon colored yellow.

10.2. Appendix 2: Calculating evaluation metrics

For the purpose of this section, assume the GUI is calculating evaluation for a single set of predictions, hereafter referred to as X.

For AUC and TSS calculations, the GUI determines the overlap between the validation data and X, and then can create spatial pairings of the predictions (density) value and either a presence or an absence. Once the GUI has these prediction value – validation data pairings, it can calculate AUC and TSS. For RMSE calculations the GUI creates the same pairings, but after converting the prediction values of X from densities to abundances. If a validation data point lies on the intersection of two or more prediction polygons, then the GUI averages the prediction values of those polygons for that prediction value – validation data pairing. The GUI displays a message below the table of metric results detailing, for each set of predictions, how many validation data points landed on the boundary between two or more prediction polygons.