Ensemble Tool for Species Distribution Models

eSDM

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

Samuel Woodman
October 2017

Table of Contents

User Manual
Ensemble Tool for Species Distribution Models (eSDM)
(beta version, Oct 2017)

1. Introduction

1.1. Why create ensemble models?

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 scales. 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). 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 scales, 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 that allows end-users to import the SDM model layers (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

The manual below describes how to use this tool, referencing sample data sets that are provided for download within the tool. The manual is divided into sections demarcated by the individual eSDM tabs, and then by the boxes within each tab. Key steps in the ensemble process include: Loading model prediction (Section 3), Providing study area boundary/coastline files, and selecting a model as the base grid for overlaying all models (Section 4); Selecting methods of creating ensemble predictions and creating those ensembles (Section 5); Calculating performance metrics for models (Section 6); Creating high-quality maps (Section 7), and Exporting model predictions (Section 8). Support and assistance is provided via a PDF manual (Section 9), comments/feedback about the tool (Section 10), and Appendices with details for some of the internal operations and calculations (Section 11).

1.3. Run the eSDM locally

One of the benefits of R Shiny applications such as the eSDM is that they can be hosted online so that users do not have to run them through R themselves. However, running the R Shiny apps locally can be faster than running from them from a server. Thus, you can download the eSDM code from GitHub at this link and run the eSDM locally.

To download all of the necessary files, click the link above and then click the green ‘Clone or download’ button on the right side of the screen and click “Download ZIP”. Unzip the downloaded file, and open the ‘server.R’ file using RStudio. Once you open the ‘server.R’ file, there will be a ‘Run App’ button above the ‘server.R’ code. Click the black dropdown arrow next to the ‘Run App’ text, and select “Run External”. This runs the app through your default browser rather than RStudio, and thus most closely simulates running the app online. Finally, click the green arrow next to the ‘Run App’ text, and the eSDM will open. The eSDM was developed using R version 3.4.2 and RStudio version 1.0.143; it is unlikely, but using other versions could cause errors within the app.

1.4. Supported import file types: SDM prediction data

SDM prediction data requirements for all file types

The prediction data loaded into the app must be either relative or absolute density predictions. Note that for the purpose of the eSDM, probabilities of occurrence and habitat suitability indices are considered a relative density. Abundance predictions are not supported at this time, and thus you must convert them to density predictions to use them into the app. Predictions that are one of 'NA', 'NaN', 'N/A', 'n/a', 'na', ‘Null’, blank, or a negative number will be classified as NA predictions and not used in any of the analyses. For GIS file types, the predictions can be in any coordinate system, but the longitude coordinates must be between the equivalent of -180 and 180 decimal degrees.

Excel .csv

SDM prediction data can be loaded from an Excel .csv file that has headers and data that meets the following requirements. The data must have longitude and latitude coordinates for each prediction, and these coordinates must be in WGS 84 geographic coordinates (decimal degrees). Longitude values may be in the range [-180, 180] or [0, 360]. The coordinates must be on a regular grid and thus equally spaced, but they can be in any order in the file and represent the center or any of the four corners of the grid cells.

GIS raster files

SDM prediction data can be loaded from a raster GeoTIFF (.tif) file.

GIS shapefile files

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 load into the app. There may be as few as four files to load, or more than ten. The extensions of these files will likely be some subset of the following: '.shp', '.shx', '.dbf', '.prj', '.sbn', '.sbx', '.shp.xml', '.fbn', '.fbx', '.ain', '.aih', '.ixs', '.mxs', '.atx', or '.cpg'.

GIS file geodatabase feature class files

To load 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. If possible, use ArcCatalog to get both the file path of the file geodatabase and the name of the feature class object. At this time, data cannot be read from personal geodatabases or from file geodatabase raster datasets.

Note: You can only load SDM predictions from a GIS file geodatabase feature class while running the eSDM locally through RStudio.

1.5. Supported export file types: SDM prediction data

Any set of SDM prediction loaded into or created within the eSDM may be exported in the following formats:

Excel .csv

To export SDM predictions to an Excel.csv file, the app calculates the centroid for each polygon that contains a prediction. The downloaded .csv file consists of columns with the longitude and latitudes of these centroids, in the specified coordinate system specified, the predictions, and the weight values. Note that depending on the coordinate system selected and if the prediction polygons have had any land area clipped from them, you may not be able to reload the saved predictions into the app because the centroids will not be equally spaced.

GIS Shapefile

SDM predictions exported as a GIS shapefile will be exported as polygons with their respective prediction and weight values. These polygons will exactly match the polygons in the app. Predictions cannot be exported to either a GIS file geodatabase or a GIS personal geodatabase.

KML or KMZ files

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.

1.6. Supported export file types: image files

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

2. eSDM Roadmap and Load or Save Environment

2.1. Load and Save App Working Environment

Description: You may save the working environment of your current eSDM session, or load a working environment you saved previously.

Load eSDM working environment

- If you saved the working environment from a previous eSDM session, you can browse to the corresponding .RDATA file and load it to restore the working environment of the previous session.

Save eSDM working environment

- You may save the current eSDM working environment to an .RDATA file so that during a later eSDM session you can load the file and resume your work from the previous session. To save the working environment of the current session, use *Filename with which to save current eSDM working environment* to provide a file name, including the '.RDATA' extension, for the .RDATA file and click *Download eSDM working environment* to download the file to your computer. You can load this file later using *Load saved eSDM session*. Aspects of the eSDM working environment that are saved include but are not limited to: original model predictions, study area polygon, land polygon, overlaid model predictions, ensemble predictions, validation data, and calculated evaluation metrics. Note that if you click this button before the eSDM fully finishes loading on start-up, which takes about 5-10 seconds, the eSDM will download a non-.RDATA file. If this happens, wait a few seconds and try the download again.

2.2. Roadmap

Description: This section describes the order in which you can use the tabs and functionality of the eSDM. You can also click the *Download sample data* button to download a zip file with sample data that you can use within the eSDM.

3. Load Model Predictions Tab

3.1. Load Model Predictions

Description: Load SDM predictions into the eSDM so that you can use them in one or more of the subsequent tabs. The *Data file type* pull down menu allows you to specify the type of input file, and adjusts the page to request relevant additional details, as described below. For each file type, an error message will appear if you select a file that does not match the file type selected in *Data file type*.

3.2.1: "Excel csv file"

Description: Select this option if the SDM predictions you wish to load are in an Excel .csv file with columns for at least longitude, latitude, and model prediction values.

Location of point in grid cell

- Specifies the part of each prediction grid cell represented by the longitude and latitude coordinates in the loaded .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 SDM prediction data you wish to load into the app. The longitude and latitude points must be equally spaced, WGS 84 geographic coordinates (decimal degrees). This file must 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 value type

- Select "Absolute density" if all necessary correction factors have been applied to the model predictions for the applicable study area, and thus they provide a true prediction of density. Select "Relative density" if the model 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 model predictions. Also select "Relative density" if the model predictions are probability of occurrence predictions or habitat suitability indices. If you select "Absolute density", then the abundance will be displayed in the *Loaded Model Predictions table* and you will be able to select "None" as a rescaling option in the ‘Create Ensemble Predictions’ tab. Otherwise absolute and relative density predictions are used identically within the eSDM.

Column with weight data

- 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 model predictions, then select “N/A”. Weight data can be used as pixel-level spatial weights in a weighted ensemble in the ‘Create Ensemble Predictions’ tab.

NA prediction values message

- A message detailing how many of the provided prediction values will be classified as NA. A prediction value will be classified as NA if the provided data is one of the following: 'NA', 'NaN', 'N/A', ‘n/a', 'na', ‘Null’, blank, or a negative number.

Load model predictions

- Click to load model predictions and other specified data from the uploaded Excel .csv file into the eSDM.

3.2.2: “GIS raster (GeoTIFF)”

<u>Description:</u> Select this option if the SDM predictions you wish to load are a band of a raster GeoTIFF file. eSDM does not currently support loading model predictions from other raster formats.

Band number of prediction data

- The band number of the prediction data within the loaded raster .tif file.

Upload raster GeoTIFF file (.tif extension)

- Browse to and load the TIFF file that has the extension '.tif. The raster coordinates can be in any coordinate system, but they must be between the equivalent of -180 and 180 decimal degrees. An error message will appear if the selected raster does not have any data at the provided band number.

Prediction value type

- See *Prediction value type*

NA prediction values message

- See *NA prediction values message*

Load model predictions

- Click to load the model predictions and other specified data from the uploaded .tif file into the eSDM.

3.2.3: “GIS shapefile”

<u>Description:</u> Select this option if the SDM predictions you wish to load are in a GIS

shapefile with at least one associated data column for the predictions values.

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 file name (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

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

Prediction value type

- See *Prediction value type*

Column with weight data

- See *Column with weight data*

NA prediction values message

- See *NA prediction values message*

Load model predictions

- Click to load the model predictions and other specified data from the uploaded shapefile into the eSDM.

3.2.4: "GIS file geodatabase (.gdb) file"

Description: Select this option if the SDM predictions you wish to load are in a file geodatabase feature class with at least one associated data column for the predictions values. eSDM does not currently support loading a raster dataset from a file geodatabase or loading a model from a personal geodatabase.
Note: You can only load SDM predictions from a GIS file geodatabase feature class while running the eSDM locally through RStudio.

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 load. 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 load. You can find this name in ArcCatalog, where the file type will be 'File Geodatabase Feature Class'.

Upload file geodatabase 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

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

Prediction value type

- See *Prediction value type*

Column with weight data

- See *Column with weight data*

NA prediction values message

- See *NA prediction values message*

Load model predictions

- Click to load model predictions and other specified data from the uploaded file geodatabase feature class into the eSDM.

3.2. Loaded Model Predictions

Description: This window contains a table with summary information about each loaded set of model predictions, which can be used to select one or more sets to display a preview plot within the app, download a preview plot, or remove loaded set(s) of model predictions from the eSDM. You may select a set of model predictions by clicking on the corresponding row in the table; this will highlight the model row in gray-blue. You may select multiple sets at one time if desired, but you can only select them when *Display additional information* is unchecked.

3.2.1: *Select loaded model predictions with which to perform an action*

- Select or deselect a set of model predictions by clicking on the row of that set of model predictions in the table. You may select multiple sets of model predictions, and you can perform action(s) with these set(s) of model predictions below the table. If selected, a row is highlighted grey-blue. The table has two sets of information that you may toggle using the *Display additional information* check box. The table illustrated above is shown when the box is unchecked, and displays the name of the file that was loaded, the names of columns from which prediction and weight data were loaded, and the specified prediction type. The 'Weight' column is blank if "N/A" was specified for that column. If *Display additional information* is checked, the table displays the resolution of the model predictions (see below for more details), the number of cells in the grid of model predictions, the count of the non-NA predictions, the predicted abundance (if the predictions are absolute densities), and the range of the model predictions.

- 'Resolution' column: When model predictions are loaded into eSDM, the app attempts to determine the resolution at which the predictions were made, meaning the distance between the centroids of adjacent grid cells. eSDM only attempts to calculate the resolution in the native coordinate system of the loaded model predictions; thus if model predictions were generated on an equal area grid but were loaded into eSDM via a shapefile in some form of lat/long coordinates, then eSDM would not be able to calculate the exact resolution. This functionality is still in a developmental stage, so please check the reported resolution if you know the resolution of your predictions.

3.2.2: *Action to perform with selected model predictions*

Description: Select the action you wish to perform with the set(s) of model predictions selected in the table. These actions include, for one or more sets of model 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 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 model 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.

Preview selected model predictions

- Click to generate the preview in the "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 preview"

Units

- See *Units*

Resolution

- The resolution of the downloaded image. It is recommended to use the "High" resolution for multiplot previews.

Image file format

- The file format in which to download the preview of the selected model predictions. The supported 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 model 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 preview of selected set(s) of model predictions with the parameters and filename specified in the *Action option(s)* box.

3.2.2.3: "Remove from app"

Remove selected model predictions

- Click to remove the selected set(s) of model predictions from the app.

4. Overlay Model Predictions Tab

4.1. Load 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. If the study area boundary cuts through a prediction cell, then only the portion of the prediction cell that is within the boundary is kept for the overlay. Use this feature if you have model predictions that cover a broad area, but you only want to create ensemble predictions in a specific study area within the broader area. For instance, if you have model predictions that span the US West Coast but only want ensemble predictions in the Southern California Bight, then you could upload a study area polygon that covers the Bight. Use the checkbox to indicate whether or not you want to use a study area polygon, and use the *File type* selection to specify how you wish to load the study area polygon. Additional options shown depend on the specified *File type*.

Use a study area polygon as the boundary for the base grid in the overlay process

- Check to upload and use a study area polygon, and uncheck to remove a loaded study area polygon. Note that a message is displayed in blue text when a study area polygon is loaded. If this box is checked, then you must load a study area polygon or uncheck the box before previewing the base grid or performing an overlay.

4.1.1: "Excel csv file"

Upload Excel csv file (.csv)

- Browse to and load 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 in the range [-180, 180], and provided in decimal format (e.g. 37.345). Multiple polygons may be demarcated using blank cells or cells with 'NA' entries. If the provided points do not form a closed polygon, then the last point is connected to the first point. Please be aware that this could create an invalid polygon.

4.1.2: "GIS shapefile"

Upload GIS shapefile files

- See *Upload GIS shapefile files*

4.1.3: "GIS file geodatabase (.gdb) file"

Note: You can only load SDM predictions from a GIS file geodatabase feature class while running the eSDM locally through RStudio.

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 file geodatabase feature class
See *Upload file geodatabase feature class*

4.2. Load Land Polygon

Description: If a land polygon is loaded, then during the overlay process all land area specified by the land polygon will be erased from the base grid. Therefore, all overlaid models will have the land erased as well. Use the checkbox to indicate whether or not you want to use a land area polygon, and use the *Land polygon source* to specify if you want to use a provided land polygon or load your own. If you are loading your own, then use the *File type* selection to specify how you wish to load the study area polygon. Additional options shown depend on the *Land polygon source* and *File type* selections.

Use a land polygon in the overlay process

- Check this box to upload a land polygon, and uncheck this box to remove a loaded land polygon. Note that a message is displayed when a land polygon is loaded. If this box is checked, then you must load a land polygon or uncheck the box before previewing the base grid or performing an overlay.

4.2.1: "Use provided"

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 use one of the provided land polygons from the Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) Database. See the GSHHG website for more information about these land polygons.

Resolution of land polygon

- The resolution options are, from highest to lowest, "full", "high", "intermediate", "low", and "crude". Each resolution option is an approximately 80% reduction in size and quality from the previous option.

Load provided land polygon

- Load the GSHHG land polygon with the specified resolution

4.2.2: "Load personal"

Description: Use the *File type* selection to specify how you wish to load the study area polygon. Additional options shown depend on the *File type* selections.

4.2.2.1: “Excel csv file”

Upload Excel csv file (.csv)

- Browse to and load the file with the .csv extension that contains the desired polygon(s). 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 in the range [-180, 180] and in decimal format (e.g. 37.345). Multiple polygons may be demarcated using blank cells or cells with 'NA' entries. If the provided points do not form a closed polygon, then the last point is connected to the first point. Please be aware that this could create an invalid polygon.

4.2.2.2: “GIS shapefile”

Upload GIS shapefile files

- See *Upload GIS shapefile files*

4.2.2.3: “GIS file geodatabase (.gdb) file”

Note: You can only load SDM predictions from a GIS file geodatabase feature class while running the eSDM locally through RStudio.

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 file geodatabase feature class

See *Upload file geodatabase feature class*

4.3. Loaded Model Predictions

Description: See *Description*. The only difference is that in this table you can only select one row at a time because the base grid must come from a single set of model predictions.

Select loaded model predictions to use as the base grid

- See *Select loaded model predictions with which to perform an action*. The only difference is that in this table you can only select one row at a time because the base grid must come from a single set of model predictions.

4.4. Overlay Model Predictions

Description: Specify options used in the overlay process, and then overlay all original model predictions onto the base grid to create overlaid model predictions that you may use to create ensemble predictions. It is strongly recommended to save the app environment before overlaying in case you are disconnected from the server during the process. If all of your model predictions were made using the same base grid, then use *Overlay version to perform*

to select “Perform same-grid overlay”; otherwise select “Perform standard overlay. This is a simplified version of the standard overlay process designed to overlay the model predictions quicker if their grids are the same. See *Appendix 1: The standard overlay process*
for a more detailed description of the standard overlay process, including how the *Overlay option* inputs are used in the standard overlay process. See *Appendix 2: The same-grid overlay process*
The same-grid for more information about the same-grid overlay process.

4.4.1: “Perform standard overlay”

Description: Overlay models with different resolutions and/or coordinate systems onto a single base grid.

Overlay options: coordinate system

Perform overlay in WGS 84 geographic coordinates

- If you check this box, then the base grid and all loaded model predictions will be in WGS 84 geographic coordinates during 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 that must be made when calculating the area using WGS 84 geographic coordinates.

Use the coordinate system of the selected model predictions during the overlay process

- If *Perform overlay in WGS 84 geographic coordinates* is unchecked, then you must specify the coordinate system in which you wish the overlay process to be performed. You specify this coordinate system by selecting the loaded model predictions that were loaded into the eSDM with the desired coordinate system

Overlay options: percent overlap

- The slider bar specifies the percent that the original model predictions must overlap a base grid cell for that base grid cell to have a non-NA overlaid prediction value. A slider bar value of "0" means that cell will have a non-NA overlaid prediction value if there is any overlap with any original model prediction.

Overlay all predictions onto the specified base grid

- Click this button to overlay all sets of model predictions onto the specified base grid using the standard overlay process. It is strongly recommended to save the app environment before overlaying in case you are disconnected from the server during the process.

4.4.2: “Perform same-grid overlay”

Description: Original model predictions that were made on the same grid and have same resolution and coordinate system do not have to go through the standard overlay process because they are already on the same grid and the predictions can simply be averaged in some fashion to create an ensemble. If the model predictions do not meet these requirements, the app will display error messages and will not allow you to perform a same-grid overlay. The app will display warning messages if the loaded model predictions could be overlaid using the same-grid overlay but certain flags are raised.

Perform same-grid overlay

- Overlay all loaded model predictions using the same-grid overlay process.

4.5. Preview of Base Grid

Description: Click *Preview* to generate a preview of the selected base grid, including loaded study area and land polygons. The preview will not be generated if *Use a study area polygon in the overlay process* is checked and no study area polygon is loaded, or if *Use a land polygon in the overlay process* is checked and no land polygon is loaded. Note that if the base grid has a high resolution, then the base grid in the preview may appear to be completely black because of the small grid cell size.

4.6. Preview of Overlaid Model Predictions

Description: Click *Preview* to generate a preview of the set(s) of overlaid model predictions selected in *Overlaid model prediction to preview*. This can be a preview of a single set of overlaid model predictions or a multiplot of multiple sets of overlaid model predictions. This feature was designed to allow you to be able to quickly preview the overlaid models, possibly at a coarser resolution than the actual overlaid predictions, and see how they differ from the original model predictions. If you wish to generate high quality images of the overlaid model predictions, please use the 'High Quality Maps' tab.

Overlaid model predictions to preview

- Select the set(s) of overlaid model predictions to preview. The sets of overlaid model 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 *Loaded Model Predictions* section, meaning that the "Overlaid 1" set of model predictions is the "Original 1" model predictions overlaid onto the base grid, and so on. You may select and thus preview multiple sets of model predictions at one time.

5. Create Ensemble Predictions Tab

5.1. Overlaid Model Predictions

Description: Specify the overlaid model predictions that you wish to use when creating ensemble model predictions.

Select overlaid model predictions to ensemble

- If the *Create ensemble using a subset of the overlaid model predictions* check box is unchecked, then all overlaid models will be used in the ensemble. If the box is checked, then you must click on rows of the table to select or deselect the overlaid model predictions you wish to use when creating ensemble model predictions.

5.2. Create Ensemble Predictions – Ensembling method

Description: For all ensembling methods, the ensemble is created by, for each grid cell, performing an average of all the predictions in that grid cell. This average may be an unweighted or weighted average, depending on the *Ensembling method* selection. If any of the overlaid models predictions are NA for a particular cell, then that NA is ignored for that cell and the other predictions are averaged together. The options displayed in this section depend on the *Ensembling method* selection, and these options are described below.

5.2.1: "Unweighted"

Description: The predictions will be averaged together via an unweighted average (simple mean).

5.2.2: "Weighted"

Description: The 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.

5.2.2.1: "Manual entry"

Description: Weight each set of model predictions by a single value.

Ensemble weights

- The values entered into the text box will be used as weights in the weighted average of the predictions. The first value will be applied to the first set of overlaid model predictions specified in *Select overlaid model predictions to ensemble*, the second value will be applied to the second set of specified overlaid model predictions, and so on. These values must be numbers greater than or equal to zero that are separated by a single comma and a single space.

5.2.2.2: "Evaluation metric"

Description: Weight each set of model predictions by an evaluation metric value. To use this weighting method, go to the 'Evaluation Metrics' tab and, for all of the overlaid models to be used in the ensemble, calculate the metric you wish to use as a weight.

Metric to use for weights

- This section will display the metric(s) that you have calculated for all of the overlaid models you have selected to use in the ensemble. Select the metric that you wish to use as weights for the sets of model predictions. The table displayed shows the metric calculated for each of the overlaid models selected to be in the ensemble, as well as their weights relative to the largest metric.

5.2.2.3: "Pixel-level spatial weights"

Description: Weight each model prediction individually. 'Pixel-level spatial weights' refer to the weight data specified via the *Column with weight* data entry when the prediction data was loaded into the app in the 'Load Model Predictions' tab. These weights are applied to the prediction in their same pixel. At least one set of model predictions must have weight data to use this weighted ensembling method. The table displayed shows which set(s) of model prediction have pixel-level spatial weights. The set(s) of model predictions that do not have pixel-level spatial weights will be included in the ensemble with a 'weight' value of 1 for all predictions.

5.2.2.4: "Polygon(s) with spatial weights"

Description: Weight model predictions within the area defined by a loaded polygon. You may load polygons, hereafter referred to as 'weight polygons', and specify the weight that will be applied to model predictions that have a specified percentage of their area within that weight polygon. Each weight polygon can be applied to one or more set(s) of overlaid model predictions. Some of the options shown depend on the *File type* selection (sections 4.2.2.4.1 – 4.2.2.4.4). The assigned weight polygons are summarized in a table below the load and assign options. You may also preview the weight polygons for a single set of overlaid model predictions by specifying the overlaid predictions to preview and clicking *Preview* in the 'Polygon(s) with weights method (cont)' box in the right part of the window.

Overlaid predictions to which to apply polygon weights

- Select the set(s) of overlaid predictions to which to apply the loaded weight polygon when you click *Assign loaded weight polygon to selected predictions*.

Weight for ___ polygon(s)

- The weight that will be applied to the model predictions within the weight polygon.

Percentage of overlap…

- Specify the percentage of a model prediction grid cell that must be overlapped by the weight polygon for that weight polygon to be applied to that model prediction.

Assign loaded weight polygon to selected predictions

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

Remove selected weight polygons

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

5.2.2.4.1: "Excel .csv"

Upload Excel .csv file (.csv extension)

- See *Upload Excel csv file (.csv extension)*

5.2.2.4.1: "GIS raster GeoTIFF"

Upload raster GeoTIFF file (.tif extension)

See *Upload raster GeoTIFF file (.tif extension)*

5.2.2.4.3: "GIS shapefile"

Upload GIS shapefile files

- See *Upload GIS shapefile files*

5.2.2.4.4: "GIS file geodatabase (.gdb) file"

Note: You can only load SDM predictions from a GIS file geodatabase feature class while running the eSDM locally through RStudio.

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 file geodatabase feature class

- See *Upload file geodatabase feature class*

5.3. Creating Ensemble Predictions – Rescaling method

Description: If any of the loaded model predictions have a prediction type of 'Relative abundance' rather than 'Absolute abundance', then you will have to select a rescaling method. This means that the model predictions will be rescaled in using one of the provided methods before they are ensembled together via a weighted or unweighted average.

5.3.1: "None"

Description: Overlaid predictions will not be rescaled before being ensembled together. This option is only available if all prediction types are 'Absolute density'.

5.3.2: "Abundance"

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

5.3.3: "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 overlaid predictions: $X_{new} = (X - X_{min}) / (X_{max} - X_{min})$

5.3.4: "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 overlaid predictions: $X_{new} = (X - \mu) / \sigma$

5.3.5: "Sum to 1"

Description: Overlaid predictions will be rescaled so that they sum to one for each set of model predictions.

5.4. Creating Ensemble Predictions – Create

Description: Click *Create ensemble* to create an ensemble using your specified ensembling and rescaling method inputs.

5.5. Created Ensemble Predictions

Description: This window contains a table with summary information about each set of created ensemble predictions, and can be used to select one or more ensembles to create or download a plot preview, remove from the app, 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 model row in gray-blue. You may select multiple model predictions at one time if desired.

This section has the same actions to perform and action option(s) as those described under *Select loaded model predictions with which to perform an action*, except you also have the option to calculate the predicted abundance for the selected ensembles. However, you only can calculate the predicted abundance of an ensemble if the rescaling method was "None" or "Abundance" when creating that ensemble.

6. Evaluation Metrics Tab

6.1. Select Predictions to Evaluate

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

6.2. Load Validation Data

Description: Load validation data into the eSDM that you wish to use to evaluate original, overlaid, and/or ensemble model predictions. 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 app adjusts the page to request relevant additional details depending on the file type and data type selected. A blue 'Validation data loaded' message will be displayed if validation data is loaded in the app.

6.2.1: “Excel .csv file”

Upload Excel .csv file

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 app 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*

Select the validation data column for the uploaded GIS file

- 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.

6.2.3: “GIS file geodatabase (.gdb) file”

Note: You can only load SDM predictions from a GIS file geodatabase feature class while running the eSDM locally through RStudio.

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 file geodatabase feature class

- See *Upload file geodatabase feature class*

Select the validation data column for the uploaded GIS file

- 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 “Presence or absence” is selected. 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 "Presence or absence" is selected. 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 presence code(s)*, and you cannot classify a code as both a presence and an absence code.

Load specified validation data into app

- Load the validation data into the eSDM with the specified inputs.

6.3. Calculate Metrics

Description: The *Validation data info* table displays information about the loaded validation data. Once you have selected the set(s) of model predictions for which you wish to calculate metrics, then specify and calculate the desired metrics.

Metric(s) to calculate

- Check the appropriate boxes to calculate area under the curve (AUC), True skill statistic (TSS), and/or root mean squared error (RMSE). You may only calculate RMSE if the loaded 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 to calculate the selected metrics for the specified set(s) of model predictions using the loaded validation data.

6.4. Metrics

Description: This section displays the calculated metrics for the specified models after you click *Calculate metrics*. You may also download the metric values.

Download metrics

- Click to download an Excel .csv file that contains the metric value(s) and model information for the models 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

Description: This section provides additional information about the supported evaluation metrics.

7. High Quality Maps Tab

7.1. Select Predictions to Map

Description: In *Select model predictions to map*, to click on the row of the original, overlaid, and/or ensemble model predictions to select the one you wish to map. Currently, you can only create a map of one set of model predictions at a time.

7.2. Map Control

Description: Generate a high quality map in-app, and if desired download the map to your local drive.

7.2.1. Generate map in-app

Generate map

- Click this button to generate a map in the box above with the parameters specified in the ‘Map Parameters’ sections. Plotting a large set of model predictions may take several minutes.

7.2.2. Download map

Description: Unlike downloading the previews in previous tabs, the image that will be downloaded is the image displayed in the box above. Thus, you must click 'Generate map' to generate a map before then downloading that map. The option to download the map will not be displayed until you have generated a map.

Resolution

- The resolution of the downloaded image.

Image file format

- The file format in which to download the preview of the selected model 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 model predictions are chosen to be plotted, or if *Resolution* or *Image file format* are changed.

Download map

- Click this button to generate a map in the box above with the parameters specified in the ‘Map Parameters’ sections. Downloading a large set of model predictions may take several minutes.

7.3. Map Parameters – Section 1

Description: Map parameters for high quality maps.

7.3.1. Map coordinate system and range

Description: Specify the coordinate system and the range of the map. 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.

Generate map in WGS 84 geographic coordinates (decimal degrees)

- If checked, the map will be generated in WGS geographic coordinates (decimal degrees). If unchecked, you may specify the coordinate system in which to generate the map with *Filename of original model predictions with desired coordinate system for map*

Filename of original model predictions with desired coordinate system for map

- The map will be generated in the native coordinate system of the specified original model prediction.

Longitude minimum

- The left-most limit of the map. This value must have the same units as the specified coordinate system and be in the range equivalent to [-180, 180] decimal degrees.

Longitude maximum

- The right-most limit of the map. This value must have the same units as the specified coordinate system and be in the range equivalent to [-180, 180] decimal degrees.

Latitude minimum

- The bottom-most limit of the map. This value must have the same units as the specified coordinate system and be in the range equivalent to [-90, 90] decimal degrees.

Latitude maximum

- The top-most limit of the map. This value must have the same units as the specified coordinate system and be in the range equivalent to [-90, 90] decimal degrees.

7.3.2. Title and axis labels

Map title

- The title of the map, which will be displayed at the top of the image. Currently this title may be only one line. Leave this box blank if you do not wish to have a title.

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 x-axis label. Currently this label can only be displayed as horizontal text.

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.

7.3.3. Tick marks and tick labels

Description: Control whether tick marks will be displayed on the map, and if so where they will be. There are no minor ticks, and tick marks and tick labels will only be displayed on the bottom and left sides of the map. If tick marks are displayed, then you also control whether or not tick labels are generated at the specified ticks. If ticks are not displayed, then tick labels cannot be displayed either.

Include tick marks in the map

- Check the box to include tick marks in the map. The other options for tick marks and tick labels are only displayed if this box is checked.

Tick location options

- Specify whether you want the tick locations to be the default tick locations determined by R ("Use default tick locations") or if you want to specify the tick locations yourself ("Enter tick locations manually").

Longitude tick locations

- Only displayed if "Enter tick locations manually" is selected. Enter the longitude values at which you want tick marks. These values must have the same unit as the coordinate system specified in the 'Map coordinate system and range' box, and must be numbers separated by ", " (e.g. "-135, -130, -115"). If the unit of the specified coordinate system is degrees, then the values must be decimal degrees. Values must be in the range equivalent to [-180, 180] decimal degrees.

Longitude tick locations

- Only displayed if "Enter tick locations manually" is selected. Enter the latitude values at which you want tick marks. These values must have the same unit as the coordinate system specified in the 'Map coordinate system and range' box, and must be numbers separated by ", " (e.g. "25, 30, 40"). If the specified coordinate system's unit is degrees, then the values must be decimal degrees. Values must be in the range equivalent to [-90, 90] decimal degrees.

Tick length

- The length of the ticks, relative to one. For instance, a size entry of 1.1 makes the ticks slightly longer than the default size, while a size entry of 0.4 makes the ticks much smaller than the default size.

Include tick labels in the map

- Check the box to include tick labels in the map. These tick labels will be displayed wherever there are tick marks.

Tick label size

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

7.4. Map Parameters – Section 2

Description: Map parameters for high quality maps.

7.4.1. Color scheme of predictions

Description: Specify the units of the color scheme of the predictions in the map, as well as 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 lowest number in the preview corresponds to the first color bin (highest 2% or largest value), second-lowest number in the preview corresponds to the second color bin, and so on.

Prediction display option

- Specify how you want to color-code the predictions. 1) If you select "Color-code predictions by relative percentage", then 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%. 2) If you select "Color-code predictions by numerical value", then the predictions will be split into *Number of colors* bins of equal size.

Color palette

- Select the color palette that you wish to use to color the model 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 was made by hand and did not come from an R package. Some of the palettes have requirements for the number of colors used with them; thus, if "Color-code predictions by relative percentage" is selected, then you cannot select some of the palettes because the “Color-code predictions by relative percentage” selection require that there be ten colors.

Number of colors

- Enter the number of colors that you wish to use, which is also the number of bins into which the predictions will be split. This input option is only available when “Color-code predictions by numerical value” and specific palettes are selected. This is because there must be ten bins for color-coding by relative percentage and some of the palettes require a specific number of bins. Also, some of the palettes have a minimum and maximum number of colors; if those palettes are selected, then there is a ‘(Min: #; Max: #)’ text display after the *Number of colors* label.

7.4.2. Background color and legend

Click to select background color

- When you click the input rectangle, a pop-up window will appear. Click in this window to select the color you want the background of the map to be. Background area is non-prediction and non-land area.

Reset background color to white

- Click this button to set the color specified in *Click to select background color* to white.

Include legend with the map

- Check this box to plot a legend to one side of the map.

Legend position

- Specify the side of the map on which the legend is displayed. The options are “Right”, “Bottom”, “Left”, and “Top”.

Legend labels: number of decimals

- Only displayed if color code predictions by numerical value” is selected. Specify the number of decimals to display in the legend labels.

7.4.3. Additional polygons

<u>Description:</u> This box allows you to include the study area polygon and/or land area polygon from the ‘Overlay Model Predictions’ tab in your high quality map. You can specify the color, line width, and drawing order of these polygons.

Include additional polygons in the map

- Check to include the study area and/or land polygon in the map.

Include selected polygons in the map

- Select the polygons that you want to include in the map. Currently you can only select “Study area polygon” and/or “Land polygon”. If you select multiple polygons, then they will be drawn in the order you selected them, meaning that the second polygon will be drawn over the top of the first. The only time this is not the case is if the first polygon is being drawn after the model predictions and the second polygon is being drawn before the model predictions. In that case, the drawing order is the second polygon, the model predictions, and then the first polygon.

7.4.3.1: “Study area polygon”

Color of study area border

- Specify the color of the line that represents the study area border.

Line width of study area border

- The width of the line representing the study area border, relative to one. For instance, an entry of 1.1 makes the line slightly wider than the default size, while an entry of 0.4 makes the line much narrower than the default width.

Draw study area polygon before model predictions

- Check if you want the study area polygon drawn before the model predictions.

7.4.3.2: “Land polygon”

Color of land

- Specify the color of the polygon(s) that represent the land.

Line width of land outline

- The width of the line representing the land outline, relative to one. For instance, an entry of 1.1 makes the line slightly wider than the default size, while an entry of 0.4 makes the line much narrower than the default width.

Draw land polygon before model predictions

- Check if you want the land polygon drawn before the model predictions.

8. Export Predictions Tab

8.1. Select Predictions to Export

Description: Click on the row of the original, overlaid, or ensemble model predictions to select the one you wish to export. You can only export one set of model predictions at a time.

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: "Excel .csv file"

Description: For predictions to be exported as an Excel .csv file, the centroid is determined for each prediction polygon. The exported .csv file consists of columns with the longitude and latitudes of these centroids, as well as the prediction and weight values for each of those points. Note depending on if any land was clipped from any of the polygons, the centroids might not be equally spaced and thus you might not be able to load this data from the Excel .csv file back into the eSDM.

Filename

- The filename must end in '.csv'.

8.2.2: "GIS shapefile"

Description: Predictions will be exported as polygons with the prediction and weight value for each polygon. The eSDM will produce a zip file with the name 'eSDM_shp_Export.zip' that will contain the various shapefile files.

Filename

- The filename must end in '.shp'.

8.2.3: "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.

File type

- Specify whether you want to download a .kml or .kmz file.

Filename

- The filename must end in '.kml' or .kmz', whichever corresponds to the *File type* specification.

Export predictions in WGS 84 geographic coordinates

- If you check this box, then the predictions will be exported in WGS 84 geographic coordinates.

Export predictions in the coordinate system of the selected model predictions

- If *Export predictions in WGS 84 geographic coordinates* is unchecked, then you must specify the coordinate system in which you wish the exported prediction to be. You specify the coordinate system by selecting the loaded model predictions that were loaded into the eSDM with the desired coordinate system.

Filename

- The name of the downloaded file with the exported predictions. The extension must match the extension required by the export format.

Export predictions

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

9. Manual Tab

Description: This tab contains this manual.

10. Submit Feedback Tab

Description: Use this form to: 1) describe functionality you would like to see in future releases, 2) report errors in the tool and 3) comment on any other facet of the tool. In all cases, please provide as much detail as possible. Submissions will be emailed to a eSDM email account that will be checked periodically over the coming months, so if you have an urgent issue please contact Karin Forney at karin.forney@noaa.gov. Thank you!

11. Appendix

11.1. Appendix 1: The standard overlay process

The overlay process is the backbone of this tool, as it allows you to create ensembles of species distribution model (SDM) predictions when the SDMs were made using grids with different coordinate systems and/or spatial resolutions.

The first step of the overlay process is converting all of the loaded sets of model predictions, which are displayed in the table in the 'Loaded Model Predictions' section of the 'Overlay Model Predictions' tab, to the coordinate system specified in *Overlay options: coordinate system*. Next, the eSDM processes the base grid, which you specify via *Select loaded model predictions to use as the base grid*. If you loaded a study area polygon, then the eSDM clips the base grid to the specified study area. If you loaded a land polygon, then the eSDM erases the land area specified by this polygon from the base grid.

Figure 1

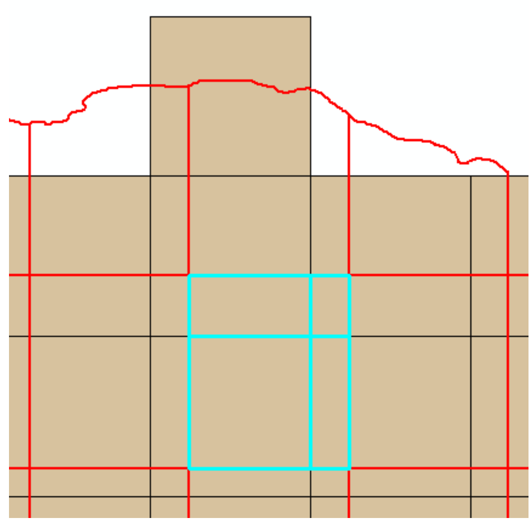

Figure 2

After creating the base grid, the other original predictions are overlaid. To do this, the eSDM performs the following process for each base grid polygon. First, the eSDM determines which polygons from the original model intersect with the base grid polygon, and if the percentage of the base grid polygon that intersects with the original prediction polygons is at least the value specified in *Overlay options: percent overlap*. If the percentage is not high enough, then that base grid cell gets a value of 'NA'. If the percentage is high enough, then the eSDM also calculates the percentage of each of these polygons that overlaps with the current base grid polygon and continues with the overlay process. Figure 1 and Figure 2 illustrate this step. In both figures, the red lines are the base grid, while the brown polygons with black outline are the original model predictions. If the eSDM is currently performing the overlay with the center, red, base grid cell in Figure 1, then you can see that that base grid cell overlaps with four original prediction polygons. The blue polygons in Figure 2 show the intersection polygons, or the intersection between that base grid cell and the four original prediction polygons (overlapping original polygons). The eSDM would then calculate the percentage of the area of the overlapping prediction polygons captured by the intersection polygons (blue polygons).

Next, the eSDM calculates the predicted abundance for the base grid cell by using the prediction value to calculate the predicted abundance for each of the overlapping prediction polygons. Then it multiplies those abundances by the percentage of the area of each overlapping prediction polygon that is in its respective intersection polygon (the blue polygons in Figure 2), and sums these new abundance values to get the predicted abundance of that base grid cell. Finally, the eSDM divides that abundance by the area of the base grid cell in order to get the density value.

The eSDM uses a very similar process for the weight values. If there are any weight values for the model predictions that the eSDM is overlaying, then for each base grid cell the weights of the overlapping original polygons are averaged together using a weighted average. The weights for this weighted average are the percentage of the area of each overlapping prediction polygon that is in its respective intersection polygon (the blue polygons in Figure 2).

The eSDM does this for all of the base grid cells, and thus at the end of the overlay process the overlaid model predictions have a very similar predicted abundance as the original model predictions that went into the overlay process. The predicted abundance might not be exactly the same between the two sets of model predictions if the original model predictions overlapped with some of the base grid cells, but not at the percentage specified in *Overlay options: percent overlap* and thus the cells got 'NA' values.

11.2. Appendix 2: The same-grid overlay process

The same-grid overlap process is intended for you to use when all of the loaded model predictions were made with and thus have the same grid. This normally will happen when you have say an Excel .csv file with multiple prediction columns, such as 'Sample_predictions_3.csv', and you load each separate column as its own set of model predictions. Since all of the predictions are on the same grid, the eSDM does not have to perform the majority of the standard overlay process. However, if you load a study area polygon or a land polygon, all of the loaded model predictions will be clipped to the study area and have the land erased.