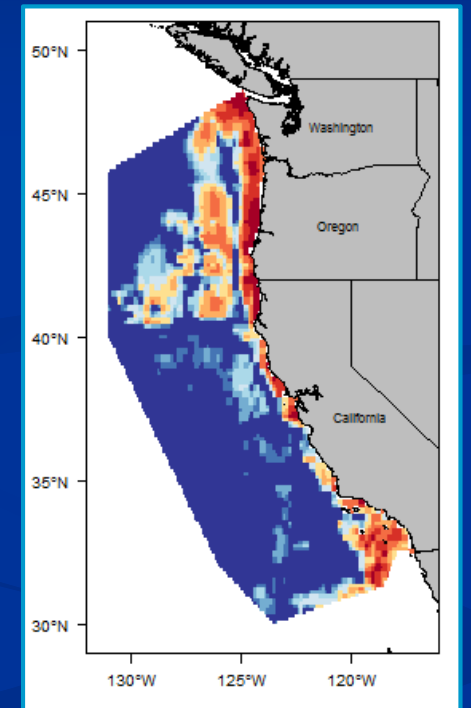
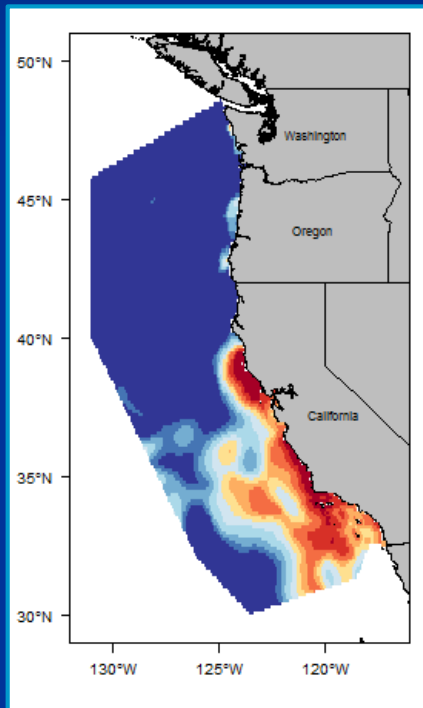


eSDM:

A tool for creating and exploring ensembles of predictions from species distribution models



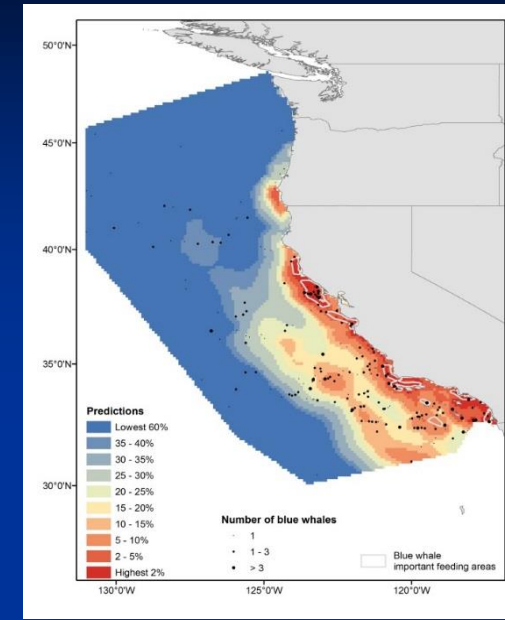
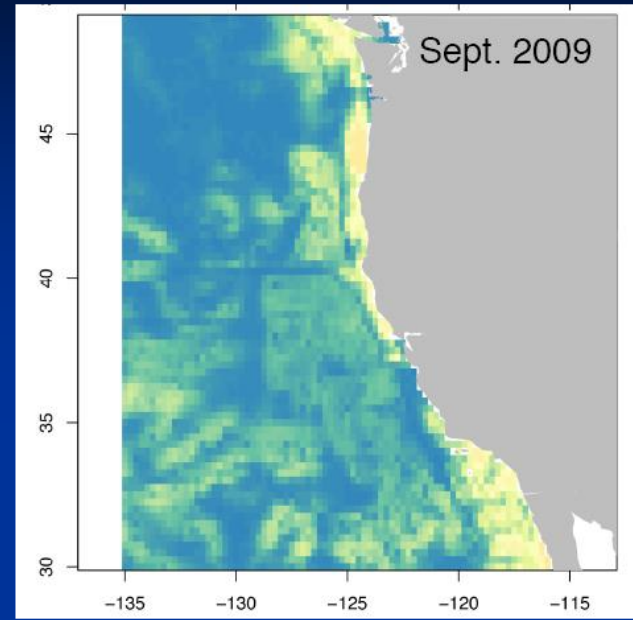
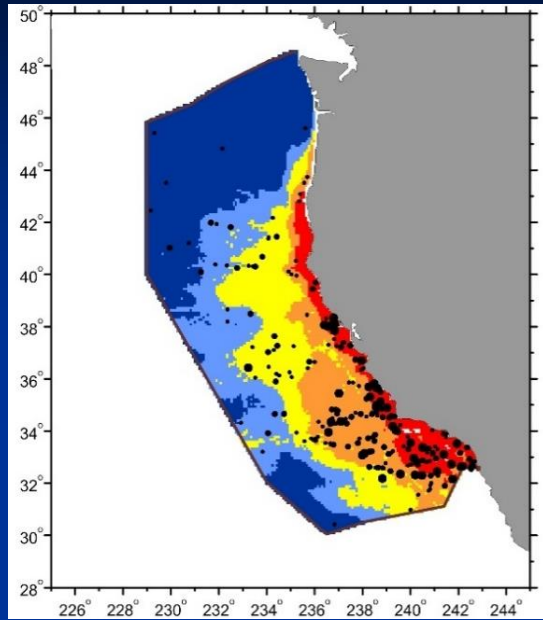
Samuel M. Woodman

Karin A. Forney, Elizabeth A. Becker, Monica L. DeAngelis, Elliot L. Hazen,
Daniel M. Palacios, Jessica V. Redfern

Outline

- Why ensemble?
- Tools for creating ensembles – old and new
- eSDM overview
- eSDM example analysis using blue whale (*Balaenoptera musculus*) SDM predictions

How to handle multiple SDMs for the same species in the same region?



Study:

Becker *et al.* 2016
Remote Sensing

Hazen *et al.* 2017
Journal of Applied Ecology

Redfern *et al.* 2017
Diversity and Distributions

Data source:

1991-2009 shipboard line-transect survey data

1994-2008 blue whale satellite telemetry data

1991-2009 shipboard line-transect survey data

Spatial res:

0.09 x 0.09 degrees (~10 x 10 km) grid

0.25 x 0.25 degrees (~25 x 25 km) grid

10 x 10 km equal area grid

Unit type:

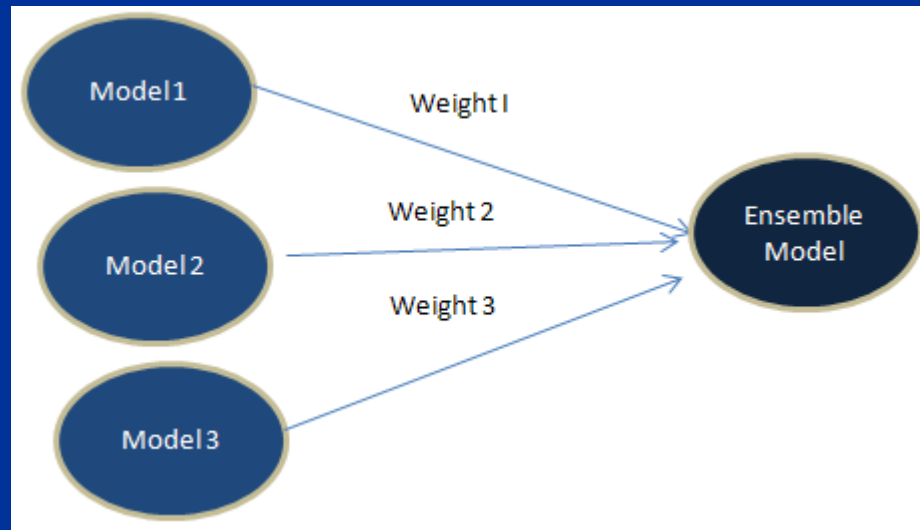
Absolute density

Probability of occurrence

Relative density

May 2015: Joint NMFS-IWC workshop to examine blue whale models and develop a plan for evaluating/combining diverse models

- No single dataset or study is correct at all areas, times (seasons, years, decades), and scales
- Different sampling techniques may provide different results with varying strengths/weaknesses



An ensemble of predictions from a set of models is often more robust than the original predictions

e.g. Araújo and New 2007, Thuiller et al. 2008, Opper et al. 2012 Gritti et al. 2013, Scales et al. 2016, Abrahms et al. (in review)

Tools for creating ensembles

- Creating ensembles from a single data source
 - Use different model algorithms (GAM, GLM, Maxent, etc) and the original data to make several SDMs
 - Average corresponding predictions from each SDM to create ensemble
- State of the art: integrated analyses, e.g. a Bayesian hierarchical framework
 - Time intensive, computationally challenging, and not possible if original data sources are not available
- A tool for creating an ensemble from SDM predictions was identified in 2016 as a project for the NOAA Fisheries Spatial Toolbox Initiative

eSDM overview

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

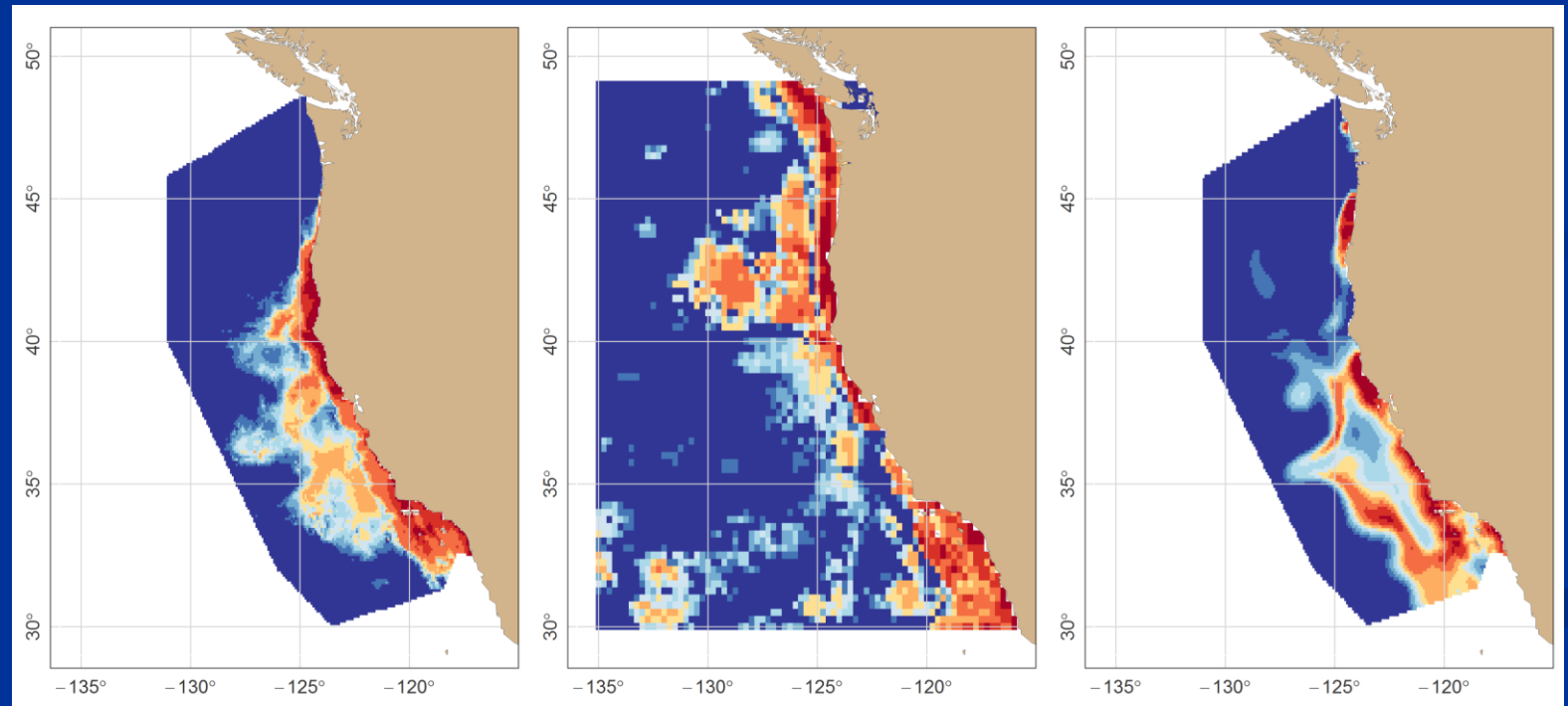
Evaluation metrics

High quality maps

Export predictions

- eSDM is an R package which includes the eSDM GUI, an R Shiny app

Example analysis: Create ensemble from 2008 predictions



Becker *et al.* 2016

Hazen *et al.* 2017

Redfern *et al.* 2017

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

The screenshot shows the 'Ensemble tool for predictions from Species Distribution Models' interface. The left sidebar contains a navigation menu with the following items: 'eSDM GUI Roadmap and Load or Save Workspace', 'Import Predictions' (highlighted), 'Overlay Predictions', 'Create Ensemble Predictions', 'Evaluation Metrics', 'High Quality Maps', 'Export Predictions', and 'Manual'. A 'Close GUI' button is located at the bottom of the sidebar. The main panel displays the 'Import Predictions' workflow. A red rounded rectangle highlights the 'Data file type' dropdown menu, which is currently set to 'Excel .csv'. The dropdown list includes the following options: 'Excel .csv', 'GIS raster', 'GIS shapefile', and 'GIS file geodatabase feature class'. Below the dropdown, there are instructions: 'Column with prediction data: Please ensure that missing predictions are represented by 'N/A', 'n/a', 'na', 'Null', blank, or a negative number.', 'Prediction unit: Select "Relative density" if the predictions are proportional to relative density.', and 'Column with weight data: Select "N/A - No pixel-level spatial weight data.'. An 'Upload Excel .csv file (.csv extension)' section shows a 'Browse...' button, a file named 'Predictions_Becker_et_al_2016.csv', and a blue 'Upload complete' button. At the bottom, there are dropdown menus for 'Column with longitude data' (set to 'lon') and 'Column with latitude data' (set to 'lat').

eSDM example analysis

eSDM GUI workflow

Import predictions

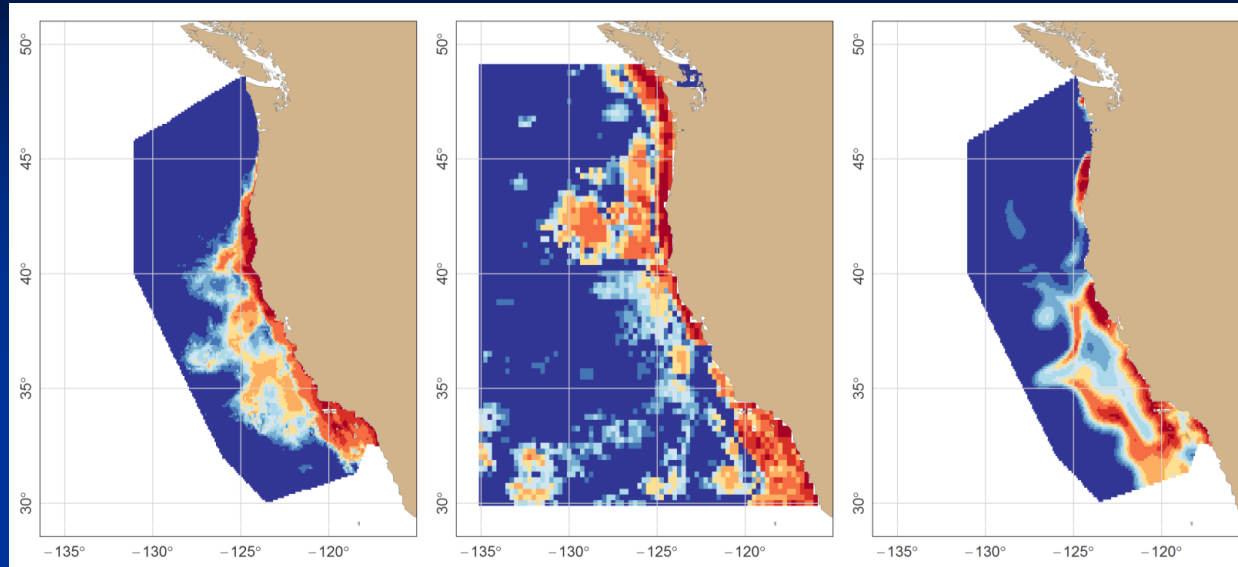
Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

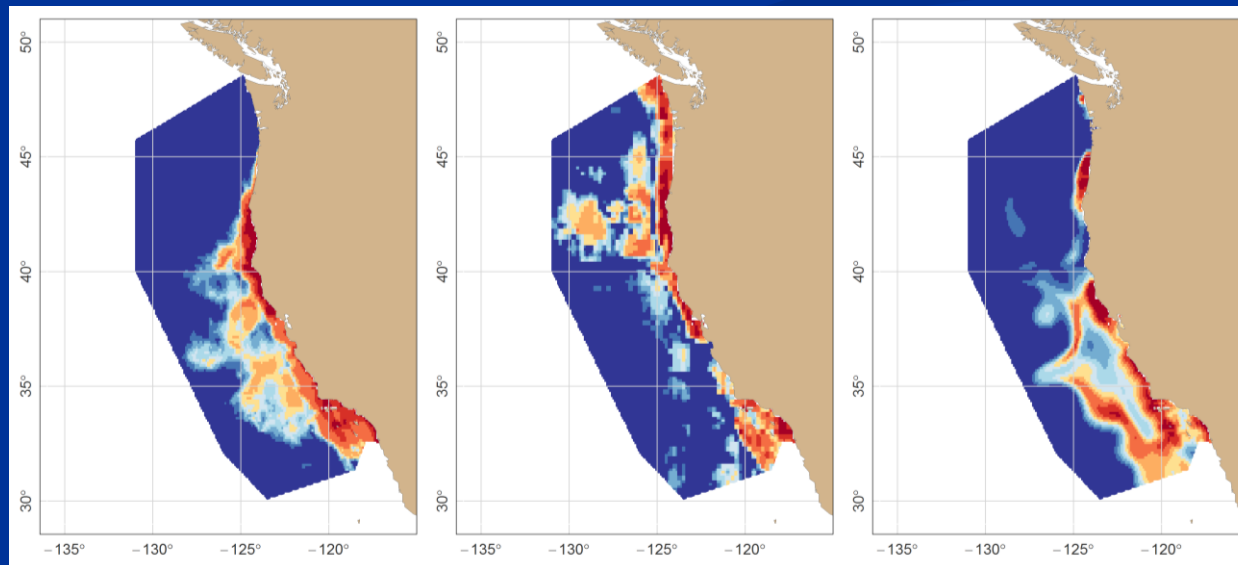


Becker *et al.* 2016

Hazen *et al.* 2017

Redfern *et al.* 2017

Original predictions



Overlaid predictions

eSDM example analysis

eSDM GUI workflow

Import predictions

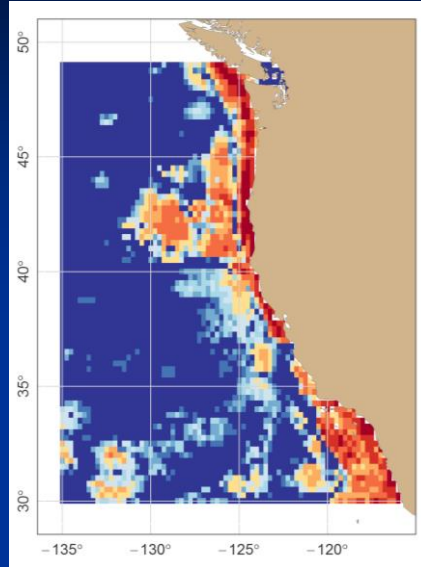
Overlay predictions

Create ensemble predictions

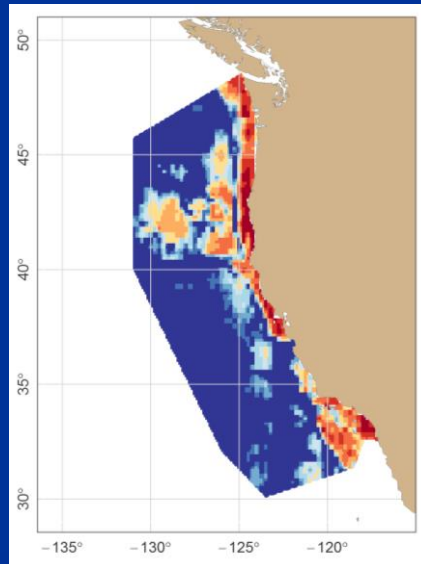
Evaluation metrics

High quality maps

Export predictions



Hazen *et al.* 2017



eSDM options

- Study area polygon
 - California Current study area
- Erasing polygon
 - High res land polygon
- Coordinate system
 - Cylindrical equal area

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

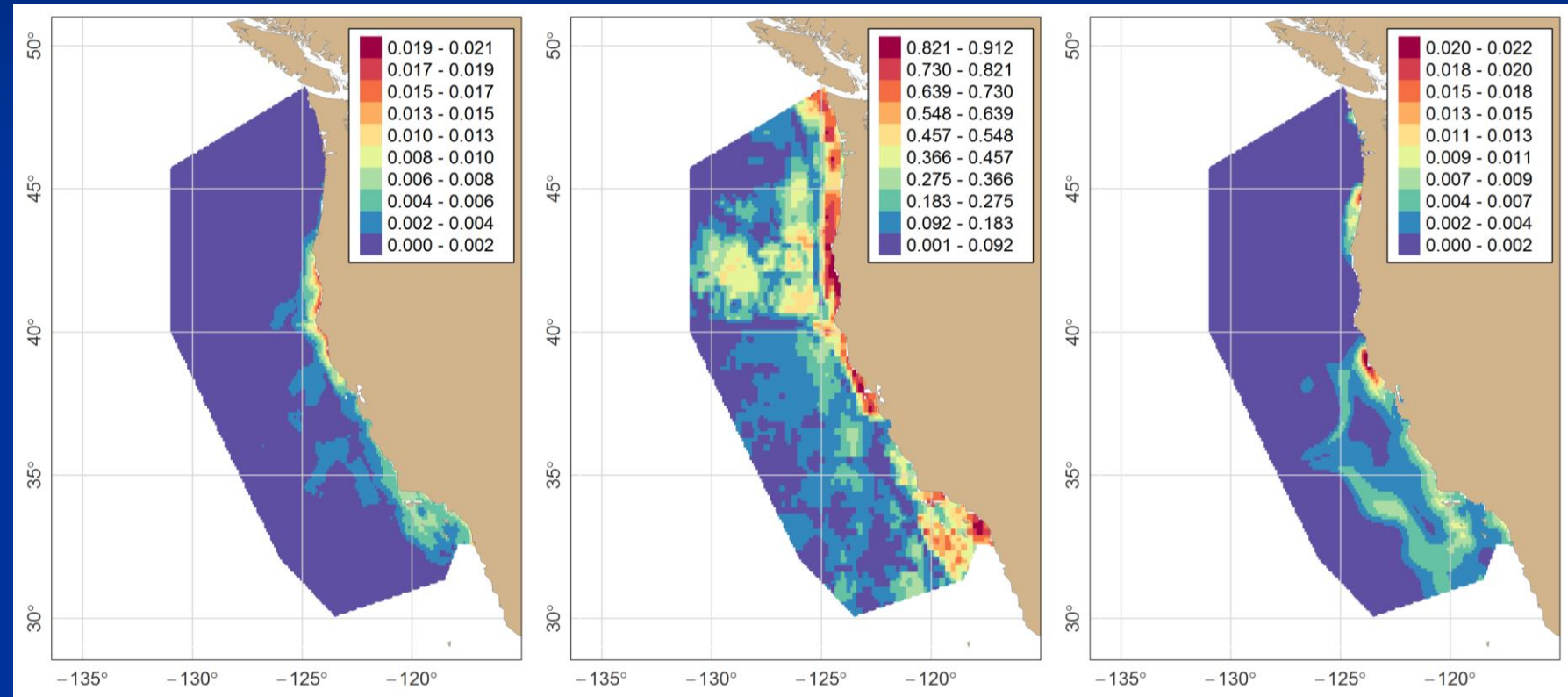
High quality maps

Export predictions

Becker *et al.* 2016

Hazen *et al.* 2017

Redfern *et al.* 2017



Absolute density

Probability of occurrence

Relative density

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

eSDM options

➤ Rescaling method

- Abundance from Barlow 2016 (NOAA TM LJ-16-01)

➤ Regional weighting

- Down-weighted Hazen *et al.* 2017 predictions off OR and WA

➤ Ensembling method

- Weighted by an evaluation metric

➤ Coming soon: calculate and map among-model variance

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

Calculate Metrics

Validation data info

Filename: Validation_data.csv

Data type	Count
Number of points with non-zero counts	8
Number of points with counts of 0	92
Range of non-zero counts	1 to 7

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

Metric(s) to calculate

AUC

TSS

RMSE

Calculate metrics

Metrics calculated

Metric Descriptions and References

Area under the curve (AUC)

True skill statistic (TSS)

Root mean squared error (RMSE)

AUC description: 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.

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

Predictions	AUC	TSS
Becker <i>et al.</i> (2016) original	0.862	0.583
Hazen <i>et al.</i> (2017) original	0.707	0.331
Redfern <i>et al.</i> (2017) original	0.872	0.639
<hr/>		
Becker <i>et al.</i> (2016) overlaid	0.865	0.583
Hazen <i>et al.</i> (2017) overlaid	0.710	0.337
Redfern <i>et al.</i> (2017) overlaid	0.872	0.639
<hr/>		
Ensemble	0.877	0.628

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

Predictions	AUC	TSS
Becker <i>et al.</i> (2016) original	0.862	0.583
Hazen <i>et al.</i> (2017) original	0.707	0.331
Redfern <i>et al.</i> (2017) original	0.872	0.639
Becker <i>et al.</i> (2016) overlaid	0.865	0.583
Hazen <i>et al.</i> (2017) overlaid	0.710	0.337
Redfern <i>et al.</i> (2017) overlaid	0.872	0.639
Ensemble	0.877	0.628

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

Predictions	AUC	TSS
Becker <i>et al.</i> (2016) original	0.862	0.583
Hazen <i>et al.</i> (2017) original	0.707	0.331
Redfern <i>et al.</i> (2017) original	0.872	0.639
Becker <i>et al.</i> (2016) overlaid	0.865	0.583
Hazen <i>et al.</i> (2017) overlaid	0.710	0.337
Redfern <i>et al.</i> (2017) overlaid	0.872	0.639
Ensemble	0.877	0.628

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

Predictions	AUC	TSS
Becker <i>et al.</i> (2016) original	0.862	0.583
Hazen <i>et al.</i> (2017) original	0.707	0.331
Redfern <i>et al.</i> (2017) original	0.872	0.639
Becker <i>et al.</i> (2016) overlaid	0.865	0.583
Hazen <i>et al.</i> (2017) overlaid	0.710	0.337
Redfern <i>et al.</i> (2017) overlaid	0.872	0.639
Ensemble	0.877	0.628

eSDM example analysis

eSDM GUI workflow

Import predictions

Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions

Predictions	AUC	TSS
Becker <i>et al.</i> (2016) original	0.862	0.583
Hazen <i>et al.</i> (2017) original	0.707	0.331
Redfern <i>et al.</i> (2017) original	0.872	0.639
Becker <i>et al.</i> (2016) overlaid	0.865	0.583
Hazen <i>et al.</i> (2017) overlaid	0.710	0.337
Redfern <i>et al.</i> (2017) overlaid	0.872	0.639
Ensemble	0.877	0.628

eSDM example analysis

eSDM GUI workflow

Import predictions

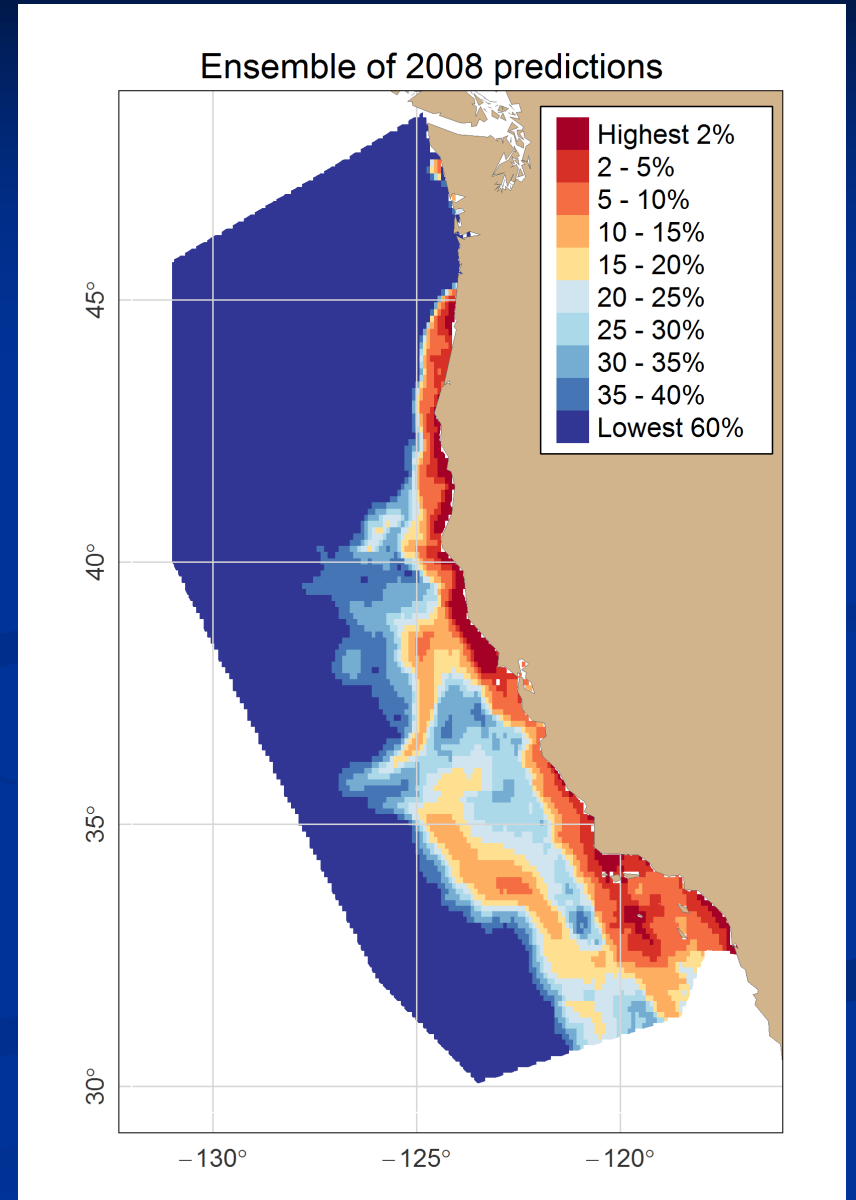
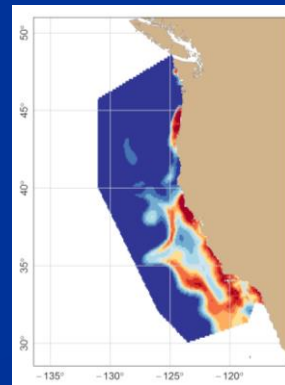
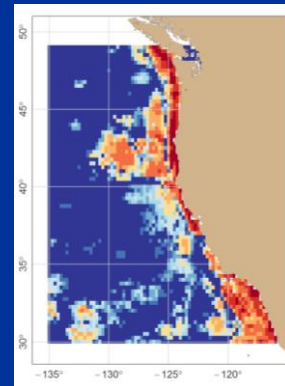
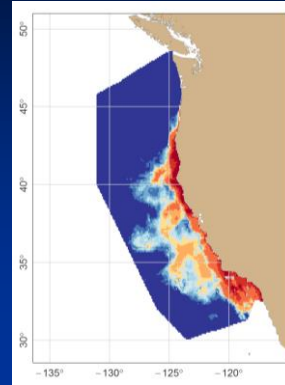
Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions



eSDM example analysis

eSDM GUI workflow

Import predictions

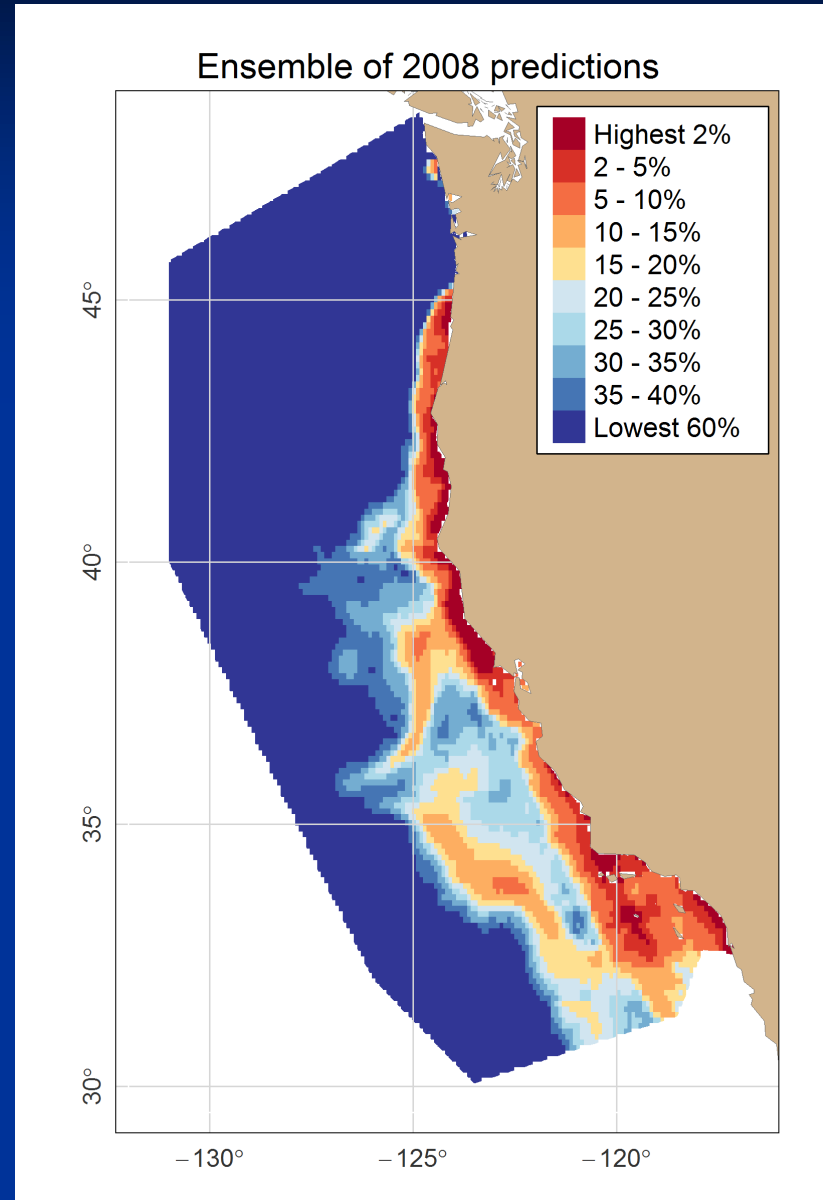
Overlay predictions

Create ensemble predictions

Evaluation metrics

High quality maps

Export predictions



Export file types

- Excel .csv
- GIS shapefile
- KML/KMZ

Takeaways from the eSDM example analysis

- Using eSDM, we successfully created an ensemble despite the different spatial resolutions, data sources, and prediction units of the original models
 - The ensemble identifies known blue whale habitat and has improved performance relative to the original predictions
- Creating an ensemble requires caution
 - Bad original models or uninformed user inputs = bad ensemble
 - If all models have the same biases, an ensemble will also have the same biases
- Researchers are frequently updating and improving SDMs, so the big takeaway here is not this individual ensemble but the utility of the tool

How to access eSDM

- R users: install R package from GitHub
 - <https://github.com/smwoodman/eSDM>
 - Option 1: Run function to open the eSDM GUI
 - Option 2: Incorporate eSDM functions into R scripts
- Non-R users: access GUI online
 - <https://swoodman.shinyapps.io/eSDM/>
- Woodman *et al.* (in review) eSDM: A tool for creating and exploring ensembles of predictions from species distribution models

Acknowledgements

- Funding: NOAA Fisheries Office of Science and Technology as part of the National Protected Species Toolbox initiative
- 2015 NMFS-IWC workshop sponsors and attendees
- Spatial toolbox steering group
- Support during development: Eric Archer, Josh London

Email: sam.woodman@noaa.gov