**Readme file for the marine protected area (MPA) tourism model**

This readme document describes the different files in this GitHub repository to be able to run the MPA tourism model.

scripts/**tourism-mpa.R** – The main code of the dive tourism model. This is where you should start.

There are helping functions and data used to be able to run the scripts/**tourism-mpa.R.**

The **functions** used in the main code are described below:

1. scripts/func\_evaluateMPA\_explicit.R – Function for evaluating the biomass change from a network of MPAs.
2. scripts/tourism-mpa-figure-plot.R – Code for plotting the main figures in the paper. This uses the saved summarized data generated from the main code, i.e., scripts/03-figures/Figure\_data.RData.
3. scripts/functions/calculate\_relative\_bio\_benefit.R – Code for calculating changes in biodiversity scores, given a network of MPAs.
4. scripts/functions/func\_evaluate\_divefee.R – Code used for the sensitivity analysis. This code calculates the changes in dive fee revenue under different assumptions like different numbers of dive estimates, different price estimates, etc.

**Raw data**

data/dive/tourism\_model\_input.RData – Key scuba dive tourism input data. This file contains the following:

"ocean\_df\_with\_eezs" – dataframe containing all the ocean pixels (with pixel ids)

"dives\_input" – Number of scuba dives per pixel

"suitability\_input" – Pixels that have existing scuba diving are identified here

"price\_constant\_input" – Price per dive per pixel, assuming that a constant dive price is used for all dive sites in the word (i.e., global median dive price)

"price\_country\_region\_input" – Price per dive per pixel, assuming median dive price per region.

"price\_interpolated\_input" - Price per dive per pixel generated by spatially interpolating actual dive price data

data/MegaData\_Sala.rds – Fisheries data from Sala et al. (2021).

data/transformed\_stockdistrib.rds – Stock distribution data used in Sala et al. (2021).

data/homerange\_pld\_predictions/pld\_rf\_predictions\_final.csv – This is the pelagic larval duration (PLD) data used in our model. This PLD is based on a combination of a comprehensive literature search and a machine learning model to fill out data gaps. See Bradley et al. (2023).

data/homerange\_pld\_predictions/homerange\_rf\_predictions\_10112022.csv – This is the species home range data used in our model. This PLD is based on a combination of a comprehensive literature search and a machine learning model to fill out data gaps. See Bradley et al. (2023).

load(file = file.path(this\_project\_dir, "data", "02-processed", "model-inputs", "bio\_model\_input.RData")) – Data files necessary to run the biodiversity model. This data file is from Sala et al. (2021).

The google drive project directory is:

#-- Path to the Pristine Seas tourism directory on the emLab Google Drive

this\_project\_dir <- "/Volumes/GoogleDrive/Shared drives/emlab/projects/current-projects/ps-tourism"

data/UN\_territory\_sovereign\_classification.csv – country classification (developed or developing country) used in our analysis.

**Saved data**

Figure\_data.RData – This file contains the saved results generated in the main code to be used to generate the figures reported in the manuscript.

data/collate\_biomass\_equi\_merged.rds – Calculated equilibrium biomass per pixel under the business-as-usual scenario

data/collate\_biomass\_equi\_merged\_dive.rds – Calculated equilibrium biomass per pixel when all known unprotected scuba dive sites in the world are placed in highly or fully protected MPAs.

data/per\_pixel\_delta\_biodiv\_benefit.rds – Biodiversity scores when each unprotected scuba dive pixel of the world is placed in highly or fully protected MPAs.

data/distance-library/merged\_dist\_matrix/merged\_dist\_matrix.rds – Distance matrix library. This matrix file contains the distance of all pixels from other pixels. Unit is in meters.

**Data preparation codes:**

scripts/01-model-prep/prep\_stock\_geogrange.R – converts stock distribution to Mollweide

scripts/01-model-prep/prep\_distance\_matrix.R – reformat global distance matrices. Distance of each pixel is needed because adult movement and larval dispersal are distance-dependent.

scripts/01-model-prep/prep\_eez\_coords.R – code for generating the EEZ shapefile, i.e., this allowed us to assign each dive pixel to EEZ.

scripts/01-model-prep/prep\_biodiversity\_inputs.R – code for preparing the inputs for the biodiversity model

RUN\_MODEL\_example.R – sample code for running the biodiversity model.

**References**

D. Bradley, A. M. Caughman, S. A. Fogg, R. B. Cabral, J. Mayorga, W. Goodell, K. D. Millage, T. D. White, Marine Fish Movement: home range sizes for commercially relevant species. *Sci. Data*. **In review** (2023).

E. Sala, J. Mayorga, D. Bradley, R. B. Cabral, T. B. Atwood, A. Auber, W. Cheung, C. Costello, F. Ferretti, A. M. Friedlander, S. D. Gaines, C. Garilao, W. Goodell, B. S. Halpern, A. Hinson, K. Kaschner, K. Kesner-Reyes, F. Leprieur, J. McGowan, L. E. Morgan, D. Mouillot, J. Palacios-Abrantes, H. P. Possingham, K. D. Rechberger, B. Worm, J. Lubchenco, Protecting the global ocean for biodiversity, food and climate. *Nature*. **592**, 397–402 (2021).