Pelagic Larval Duration (PLD) predictive model README

We need pelagic larval duration (PLD) information for **811 unique fish and invertebrate species.**

This is the model for home range:

*Home range ~ beta0 + beta1(r) + beta2(K) + beta3(categorical movement parameter) + beta4(IUCN red list status)  + beta5(body length) + beta6(trophic level) + beta7(pelagic or demersal)*

What is our model for the PLD?

Data:

* Our species list can be found here:
* Ramesh et al. (2019)’s life history data table contains 1798 rows but there are duplicates (i.e., data compiled from several studies per species). There are 265 unique entries with larval duration data. **65 matches our modeled species list**.
  + The database can be found here:
    - <https://zenodo.org/record/3249468#.YZbsgJBBzLB>
      * See the “spawn” folder.
    - Or here: <https://github.com/rencabral/tourism-mpa/blob/master/data/pld_ramesh.csv>
* Marshall et al. reported life history data for 806 invertebrate species. 254 of which have “planktonic time” data. **2 matches our modeled species list.**
  + The database can be found here:
    - <https://www.annualreviews.org/doi/suppl/10.1146/annurev-ecolsys-102710-145004>
      * Supplementary Table 1
    - Or here: <https://github.com/rencabral/tourism-mpa/blob/master/data/pld_marshall.csv>
* Check how many species per family.

* Home range data collected via lit review can be found in mpa-fish-flows/data/01\_raw/homerange/all\_homerange.csv; column = ‘hr’
  + <https://drive.google.com/file/d/1nQM44mEgcsfe47QaJK2EwU9lBvs9O919/view?usp=sharing>
* For species list and model parameters, data can be found in the following locations:
  + species list, r, K, categorical movement parameter: MegaData\_Ray.rdf (available: <https://drive.google.com/file/d/1vrrs4SC-pJfWwNbWJGPyXERC7uZUMkQZ/view>)
    - species = SciName
    - K = Kfin
    - r = r\_fin
    - categorical movement parameter = m\_fin
* IUCN red list status: included\_aquamaps\_spp\_list.csv (available: <https://drive.google.com/file/d/1IOiK8s2jCiJ1jKv0ksYlRQti1fstNz3r/view>)
  + IUCN red list status reported in ‘iucn\_category’
  + note this has a different larger associated species list and so will need to be merged via a left join (merging by SciName in the prior file and valid\_sci\_name
  + the script with the original species matching for this exercise is here (if useful): https://github.com/emlab-ucsb/ocean-conservation-priorities/blob/master/data\_prep/01\_get\_species\_list.Rmd
* body length, trophic level, pelagic/demersal: available via Fishbase (see Prior Analyses section below for more details)

Relevant Github repos:

* tourism-mpa: <https://github.com/rencabral/tourism-mpa>
  + The PLD data will be used to parameterize the biological model of the tourism-MPA model. All analyses should go here, noting that there are not yet any scripts associated with any homerange analyses

General approach:

1. Create data frame with model parameters: merge relevant columns from MegaData\_Ray.RDS with included\_aquamaps\_spp\_list.csv with all\_homerange.csv (retaining full species list from MegaData\_Ray.RDS)
2. Get Fishbase parameters from the rfishbase package; use Tim’s method reported in the read me to impute missing values
3. Sample training data based on family group (or some higher taxonomic classification or trophic level if training data are sparse)
4. Run RF model

Model

Pelagic larval duration (in days) will be a function of the following variables

1. Family – same family tend to have the same PLD
2. Home range? Nope. Many sedentary species do have broad PLD. Home range and PLD may be negatively correlated. Any of the pelagics have low PLD while sessile/low mobility species tend to have long PLD (e.g. lobsters).
3. Mobility? Maybe? We can check how mobility categories matches with PLD for matched data.
4. Pelagic vs demersal – yes I think. Good to include.

**References**

Ramesh, N., Rising, J.A. and Oremus, K.L., 2019. The small world of global marine fisheries: The cross-boundary consequences of larval dispersal. *Science*, *364*(6446), pp.1192-1196.

![Chart, histogram

Description automatically generated]()

![Chart, histogram

Description automatically generated]()

![Chart

Description automatically generated]()

![Chart, scatter chart

Description automatically generated]()

A picture containing diagram

Description automatically generated