MADS

Multi-Analysis Distance Sampling

CREEM
Centre for Research into Ecological
and Environmental Modelling

University of
St Andrews

Contents

- Analysis Issues
  - Unidentified Sightings
  - Covariate Uncertainty
  - Model Uncertainty

- mads R library
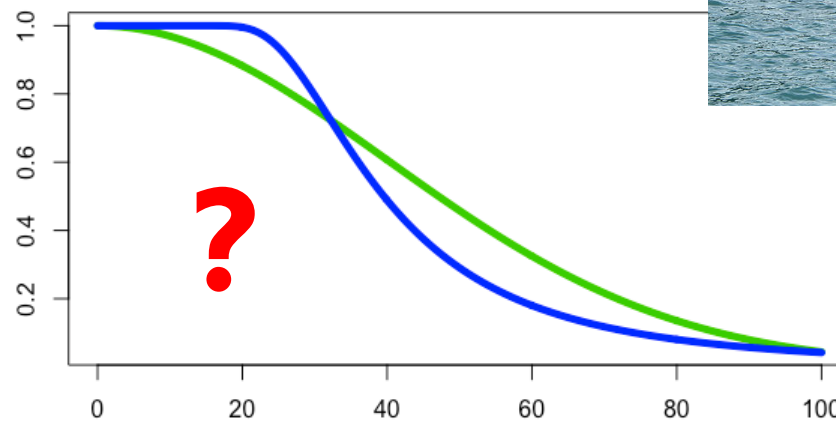
- Distance for Windows

Distance 7.0
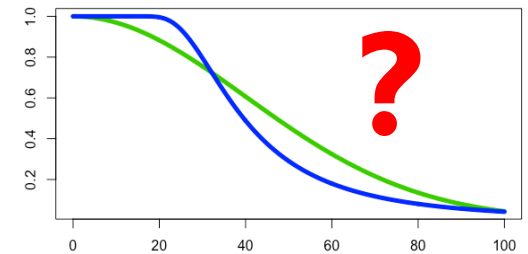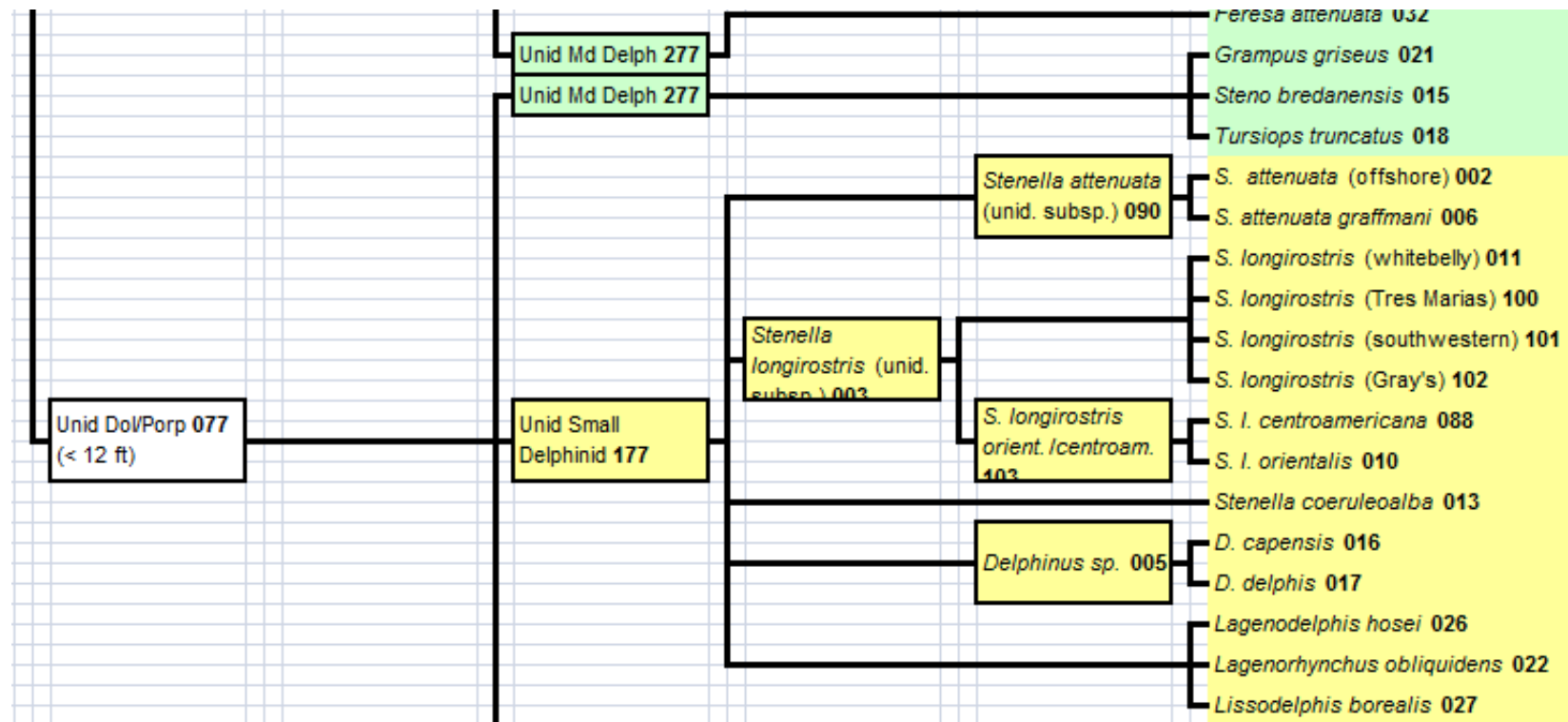
How many?
60, 80, 100?

Analysis Challenges

- Unidentified sightings
  - Sightings cannot be identified to species
  - Causes bias when there are unidentified sightings on the transect

- Covariate uncertainty
  - E.g. uncertainty in cluster size
  - If not included you may under-estimate variability

- Model uncertainty
  - Multiple detection functions with small differences in selection criteria (e.g. AIC) but large differences in estimated density
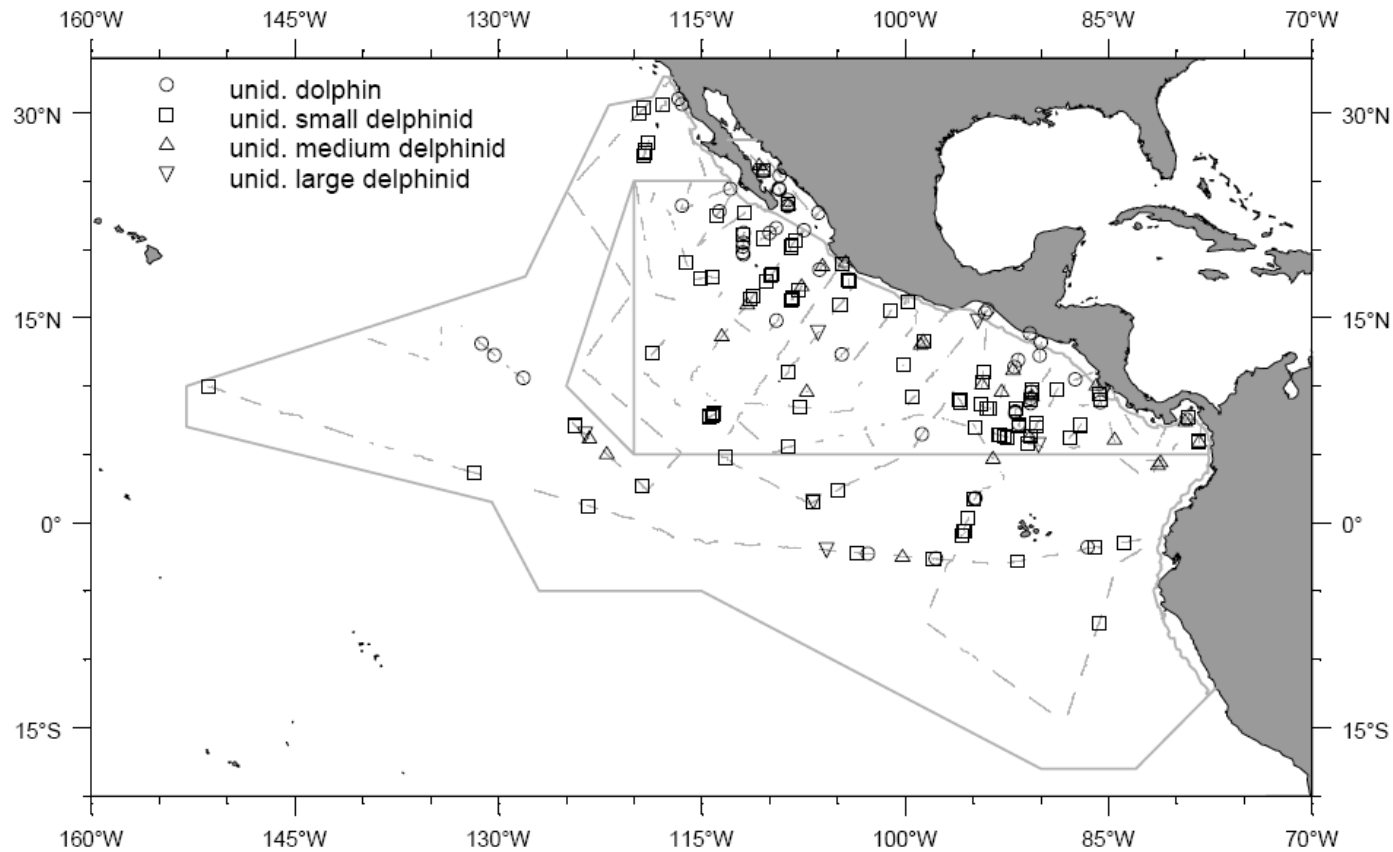 - If not included you may under-estimate variability

?

How many?
60, 80, 100?

?

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Unidentified Sightings – Sightings Category Hierarchy

Unidentified Sightings – ETP Data

Only a problem if there are unidentified sightings on or close to the trackline!

Unidentified Sightings

$$\hat{N}_{ij} = \hat{N}_{ij}^* + \hat{N}_{uj}\left(\frac{\hat{N}_{ij}^*}{\sum_{k=1\dots n}\hat{N}_{kj}^*}\right)$$

i refers to the ith stock / species code,
u refers to the unidentified code
j denotes the stratum,
n gives the number of species / stocks
   associated with the unidentified code u
* denotes abundance estimates based on
   only identified sightings

$\hat{N}_{IdESpin} = 40$

$\hat{N}_{IdWBSpin} = 60$

$\hat{N}_{UnidSpin} = 10$

$\hat{N}_{ESpin} = 40 + 40\% \, of \, 10 = 44$
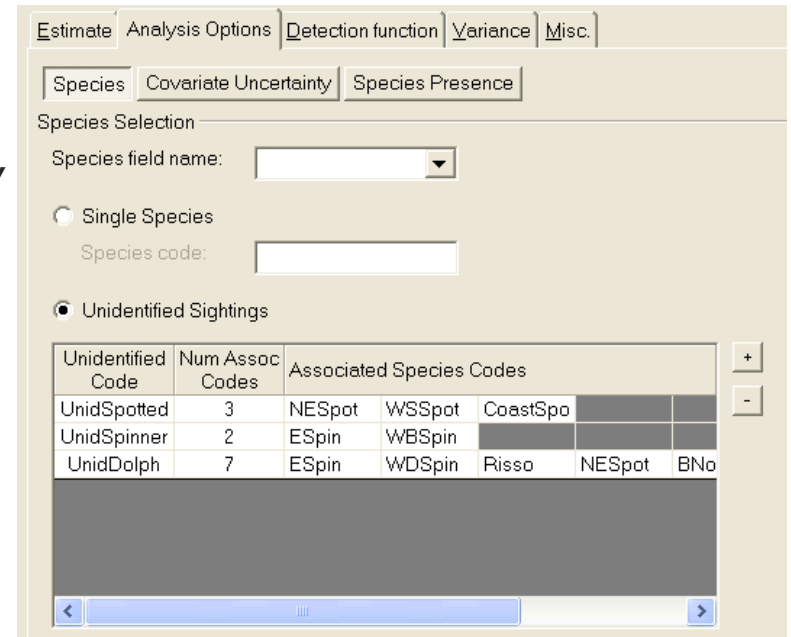
$\hat{N}_{WBSpin} = 60 + 60\% \, of \, 10 = 66$

Gerrodette, T. and Forcada, J. (2005) Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. *Marine Ecology Progress Series*, 291:1-21.

CREEM
Centre for Research into Ecological
and Environmental Modelling

University of
St Andrews

Unidentified Sightings

mads

Distance 7

```
unid.defs <- list("UnidSpotted" = c("NESpot","WSSpot",
                                    "CoastSpot),
            "UnidSpinner" = c("ESpin","WBSpin"),
            "UnidDolph" = c("ESpin","WBSpin",
                            "NESpot","WSSpot",
                            "CoastSpot", "Risso",
                            "Bnose"))
```

Estimate | Analysis Options | Detection function | Variance | Misc.

Species | Covariate Uncertainty | Species Presence

Species Selection

Species field name: [▼]

○ Single Species

   Species code: [ ]

⊙ Unidentified Sightings

Unidentified Code	Num Assoc Codes	Associated Species Codes				
UnidSpotted	3	NESpot	WSSpot	CoastSpo		
UnidSpinner	2	ESpin	WBSpin			
UnidDolph	7	ESpin	WDSpin	Risso	NESpot	BNo

Species Presence

mads   Distance 7

```
presence <- list("Core" = c("NESpot","ESpin","WBSpin","Bnose", "Risso"),
          "Core2" = c("WSSpot","ESpin","WBSpin","Bnose", "Risso"),
          "Outer" = c("WSSpot","WBSpin","Bnose", "Risso"),
          "Coastal North" = c("CoastalSpot", "WBSpin", "Bnose",
                             "Risso"),
          "Coastal Sorth" = c("CoastalSpot", "WBSpin", "Bnose",
                             "Risso"))
```

Covariate Uncertainty

- If you do not incorporate covariate uncertainty then you may under-estimate the variability in your density/abundance estimates!

- mads incorporates covariate uncertainty by resampling covariate values during a bootstrap based on a chosen distribution.

  - *Normal*
  - *Lognormal (bias corrected)*
  - *Poisson*
  - *Zero-truncated Poisson (zero truncated)*

	object	distance	size	totsize	totsizesd
1	1	5.04	1	1	0.39787
2	2	6.01	218	218	0.62944
3	3	0.21	2	2	0.32486
4	4	0.67	24	24	0.28934
5	5	0.02	13	13	0.42235
6	6	0.25	254	254	0.35072

Covariate Uncertainty

mads

Distance 7

```
cov.uncert <- data.frame(variable.layer = "observation",
            variable.name = "size",
            cor.factor.layer = "numeric",
            cor.factor.name = 1,
            uncertainty.layer = "observation",
            uncertainty.name = "sizeSD",
            uncertainty.measure = "sd",
            sampling.distribution = "Lognormal.BC")
```
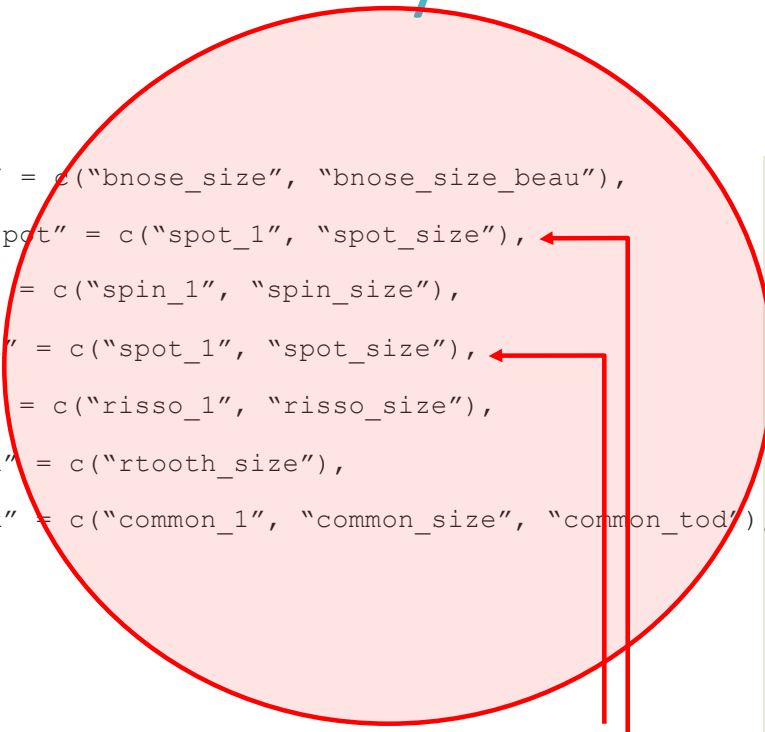
Model Uncertainty

- What to do when you have two or more models providing an equally good fit to the data but giving substantially different estimates of density / abundance?

- Not incorporating model uncertainty may lead you to under-estimate the variability in your density/abundance estimates!

- mads resamples the data during a bootstrap, refits the models and selects the one with the minimum selection criteria; AIC, AICc or BIC. You can either resample observations or samplers (i.e. transects) the latter option is the usual resampling unit.

Model Uncertainty

mads

```
mod.uncert <- list("Bnose" = c("bnose_size", "bnose_size_beau"),

                   "CoastSpot" = c("spot_1", "spot_size"),

                   "ESpin" = c("spin_1", "spin_size"),

                   "NESpot" = c("spot_1", "spot_size"),

                   "Risso" = c("risso_1", "risso_size"),

                   "RTooth" = c("rtooth_size"),

                   "ShBeak" = c("common_1", "common_size", "common_tod"),

                   …

                   … )
```

These are all models fitted previously
for each of our species codes

Distance 7

Detection function models table showing Species Code, Num Models, and Model Names columns.

Other multi-analysis input…

- Need to supply a vector of all species codes.

- Need to supply the models which have already been fitted to the data using mrds, mads then obtains the observation data from these model object. Note – if running in Distance make sure Distance retains the fitted models in the R workspace!

- Supply the model options – the model selection criteria

- Supply the bootstrap options

- Supply a region, sample and observation tables

Other multi-analysis input...

```
species.codes <- c("bnose", "CoastSpot", "Espin", "NESpot", "Risso", "Rtooth", ... )
models <- list("bnose_size" = bnose_S,
               "bnose_size_beau" = bnose_SB,
               "spot_1" = spot_IO,
               ...,
               ...)
model.opts <- list(criterion = "AIC")
bootstrap.opts <- list(resample = 'samples', n=999)
```

Other multi-analysis input…

- Region, Sample and Observation tables

```
> region.table
  Region.Label Area
1            A 1680
```

```
> sample.table
   Sample.Label Region.Label Effort
1        1. T1            A    210
2        2. T2            A    210
3        3. T3            A    210
4        4. T4            A    210
5        5. T5            A    210
6        6. T6            A    210
7        7. T7            A    210
8        8. T8            A    210
9        9. T9            A    210
10     10. T10            A    210
```

```
> obs.table
   object Region.Label Sample.Label
1       1            A        1. T1
2       2            A        1. T1
3       3            A        1. T1
4       4            A        1. T1
5       5            A        1. T1
6       6            A        2. T2
7       7            A        3. T3
8       8            A        4. T4
9       9            A        4. T4
10     10            A        4. T4
```

Running the multi-analysis...

```
results<- execute.multi.analysis( species.code = species.codes,
                 unidentified.sightings = unid.defs,
                 models.by.species.code = mod.uncert,
                 ddf.model.objects = models,
                 ddf.model.options = model.opts,
                 region.table = region.table,
                 sample.table = sample.table,
                 obs.table = obs.table,
                 bootstrap = TRUE,
                 bootstrap.option = bootstrap.opts)
```

Output – Analysis Summary

```
Bootstrap resample implemented: TRUE

Number of resamples: 100 #could to with a few more!


Unidentified species codes were included in these analyses.

They were prorated as follows:

    Unidentified code UnID was prorated to species codes CD, WD


Model uncertainty was included in these analyses. See species results for convergence, selection and
model summaries.


Data details:

    Clusters: FALSE
```

Output – Species Summary

```
Bootstrap summary for species  :  CD


Detection function model summary


Model Selection:

                   df.all.hn df.all.hr

Converged               100       100

Failed to Converge        0         0

Selected                 71        29


(followed by a summary of the models and parameter estimates for selected
models)
```

Output – Species Summary

Summary for individuals

Bootstrap summary statistics:

	Region	Area	CoveredArea	Effort	n	ER	se.ER	cv.ER
1	Total	1e+06	6e+05	10000	186.51	0.018651	0.001392802	0.07467708

Abundance:

	Label	Estimate	se	cv	lcl	ucl	df	pctUnid	pctUnid.se	pctUnid.cv	pctUnid.lcl	pctUnid.ucl
1	Total	663.6778	94.17961	0.1419056	537.3726	869.0931	104.7939	10.66939	1.570144	0.1471634	7.997528	14.08607

Density:

	Label	Estimate	se	cv	lcl	ucl	df
1	Total	0.0006636778	9.417961e-05	0.1419056	0.0005373726	0.0008690931	104.7939a