Simulated underwater glider sampling for Antarctic krill

Doug Kinzey, NOAA Southwest Fisheries Science Center

10/28/2022

Introduction

Two functions, gldrs() and gldry(), simulate autonomous underwater gliders equipped with echosounders sampling in a diagonal up and down swath through the water column. A dataset of krill acoustic densities collected during US AMLR ship surveys from two strata in CCAMLR Subarea 48.1 from 1992 to 2011 is sampled using this sawtooth pattern. The ability of the samples obtained to reproduce vertically-integrated acoustic energy returns (nautical area scattering coefficients, or 'NASC') attributed to Antarctic krill is assessed. The gliders sample at a vertical angle of 22.7 degrees, or an aspect ratio of 2.22. Each descent and ascent of a completed dive is called a 'yo'. The echosounder in this simulation ensonifies the water column 100 m below the gliders as they descend to a maximum depth of between 150 or 1000 m per yo. The gliders ensonify only the first 250 m of the water column (where most krill are found) regardless of maximum yo depth. These gliders do not sample on the ascending portion of their dive.

Input arguments

```
NASC.yrs = c(2001:2009,2011)
AMLR.area = 'SA'
n.rep = 6
n.gldr = c(1,3,5)
save.tables = 1
max.NASC.m = 250
depths = c(150,500,1000)
azfp.off = c(150,150,150)
qntl.vals = c(0.97,0.98,0.99,0.999,1)
smpl.st = 0
```

There are ten input parameters to the simulation:

'NASC.yrs' are the years to be sampled, with a maximum range of 'c(2001:2009,2011)' (the sampling strata in the dataset were unsampled by the ships in 2010).

'AMLR.area' can be one of two values 'SA' (the southern area, Bransfield Strait) or 'WA' (the western area, Cape Shirreff).

'n.rep' is the number of replicate glider samples to collect. Each replicate can represent one or multiple gliders. Run-time increases with increasing replicates.

'n.gldr' is the number of gliders sampling the stratum during each replicate. It can be a scalar or a vector of values representing sequentially increasing numbers of gliders. When 'n.gldr' is a vector it will overwrite the results with each new grouping of gliders so the simulation results in this case need to be saved as tables (save.tables = 1) and then accessed for analysis using read.table(). Run-time increases with increasing gliders/replicate. Run times for 500 replicates of 5 gliders/replicate can take several weeks and run out of memory on a 24 thread 32 GB RAM PC.

'max.NASC.m' is the maximum depth ensonified by the gliders. In these simulations the echosounder is turned off at 150 m regardless of the maximum yo depth, so this is always 250m.

'depths' are the maximum yo depths to be simulated. For a given dataset of krill densities, deeper yos will obtain fewer samples.

'azfp.off' is a vector of the same length as 'depths'. This allows the option to ensonify deeper during deeper dives up to the maximum depth of krill densities recorded in the dataset, 500 m in this example.

'qntl.vals' allow the proportions of krill in these highly skewed distributions to be assessed.

'smpl.st' allows either random (smpl.st = 0) starting positions in the dataset for each replicate sample, or using the same starting positions (= 1) for each replicate as a previous simulation. If this option is desired, save.tables = 1 for the previous simulation, and then the 'gldr_strt_ . . . ' files from the first run are placed in the working directory and called using smpl.st=1. This option can be useful in evaluating the effects of coding changes on identical glider sampling patterns.

Load the datasets

There are two datasets: (1) 'NASC_leg1.RData' (dimensions 100 x 199772), the krill acoustic values binned into 100 m horizontal by 5 m vertical cells (transposed in the simulation functions for easier visualization), and (2) 'leg1.csv' (dimensions 199772 x 5), the unique identifiers, strata and year for each column of 'NASC_leg1.RData'. Two additional files, 'AMLR_SA_transects.csv' and 'AMLR_WA_transects.csv', contain the original ship transect waypoints of the Bransfield Strait and Cape Shirreff strata, respectively, for Figure 1.

```
load('input_files/NASC_leg1.RData')
leg1 <- read.csv('input_files/leg1.csv')
```

Source the R code

Two R scripts, 'gldrs.r', and 'gldry.r' contain the two functions, 'gldrs()' and 'gldry(), respectively. 'gldrs()' reads the input arguments and calculates four values that are passed to 'gldry()'. 'gldry()' calculates the glider sampling pattern and returns this pattern and three other arguments to 'gldrs()', which then samples the acoustic dataset following the glider sampling pattern.

```
source('input_files/gldry.r')
source('input_files/gldrs.r')
```

Run the simulation

Supply the 'Input arguments' above to 'gldrs()'.

```
Sys.time() # keep track of runtime
```

```
## [1] "2022-10-28 07:40:01 PDT"
```

```
gldrs(NASC.yrs=NASC.yrs,AMLR.area=AMLR.area,n.rep=n.rep,n.gldr=n.gldr,
                      save.tables=save.tables,max.NASC.m=max.NASC.m,depths=depths,
                      azfp.off=azfp.off,qntl.vals=qntl.vals,smpl.st=smpl.st)
Sys.time()
```

```
## [1] "2022-10-28 08:22:43 PDT"
```

Next steps

After the 'Gldr_simulation…rmd' files have been rendered, the working directory will contain output files from the simulations in a 'Tables' subdirectory. These output files can be plotted using 'Figures1-10_SA.rmd', and 'Figures1-10_WA.rmd'. Note that both 'Glider_simulation_WA.rmd' and 'Glider_simulation_SA.rmd' must finish before 'Figures1-10_SA.rmd' and 'Figures1-10_WA.rmd' can produce the coverage plots in Figure 9.