

Antarctic_Rarefaction

Trevor Eakes

August 6, 2016

Rarefaction methods applied to Antarctic phytoplankton cell abundance counts taken from the raw data analyzed in Yamamoto, 1986. This analysis is faulty at the moment because it is too computationally challenging for the computer to manage, and the methods do not apply for how large the sample size is. See notes in script below.

Note Diatoms in Antarc are tentatively from column 3 to 50

Preparing the data

```
source('NiftyTrevFunctions.R')
Antarc <- read.csv("AntarcticOceandatasheet.csv") #Pull up datasheet
head(Antarc)
```

```
##      Transects      Location Diatoms Actinocyclus..1.sp..
## 1    PF transect    Antarctic      NA                    0
## 2    PF transect      Middle      NA                    10
## 3    PF transect Subantarctic      NA                    0
## 4    STC-1 transect Subantarctic      NA                    0
## 5    STC-1 transect      Middle      NA                    0
## 6    STC-1 transect Subtropical      NA                    0
## Asteromphalus.hookerii A..pervulus Chaetoceros.atlanticum
## 1              30          190              20
## 2              65           40              90
## 3               4            8               8
## 4               0            4               0
## 5               0            0               0
## 6               0            0               0
## C..atlanticum.v..audax C..atlanticum.v..neapolitanum C..boreale
## 1               0              0              15
## 2               0              0              20
## 3               0              0               0
## 4               0              0               0
## 5               0              0               0
## 6               0              0               0
## C.bulbosus C..compressum C..concavicornis C..decipiens C..glaciles
## 1          15              0           40              0              0
## 2           0              0           70             10              0
## 3           0              0            4              0              0
## 4           0              0            0              0              0
## 5           0              0            0              0              0
## 6           0              0            0              0              0
## C..laciniosum C..pendulum C..peruvianum C..shimperi C..shimperi Chaetoceros.spp.
## 1           0              0            0              0              0
## 2           0              0           20              0             35
## 3           0              0            8              0              4
## 4           0              0           20             84             52
## 5           0              0           60             20             40
```

## 6	0	0	60	190	180
##	Corethron.criophilum	Coscinodiscus.bullatus	C..furcatus	C..gyratus	
## 1	335		0	0	55
## 2	300		0	10	160
## 3	8		2	2	56
## 4	30		0	0	24
## 5	35		0	0	0
## 6	35		0	0	0
##	C..inflatus	C.lentiginosus	C..oculoides	C..ritscherii	C..tumidus
## 1	10	115	40	200	10
## 2	30	50	0	205	70
## 3	0	16	2	16	2
## 4	0	0	0	8	0
## 5	0	0	0	0	0
## 6	0	0	0	0	0
##	C..subtilis.	C..variolatus.	Coscinodiscus.spp.	Dactyliosolen.antarcticus	
## 1	0	0	0		75
## 2	0	0	0		100
## 3	0	2	0		0
## 4	4	0	0		0
## 5	0	0	0		0
## 6	0	0	80		0
##	Eucampia.balaustium	Navicula.spp.	Nitzschia.spp.	Planktoniella.sol	
## 1	15	720	42410	0	
## 2	20	600	101580	0	
## 3	10	145	11293	0	
## 4	3	83	1441	0	
## 5	0	27	1189	5	
## 6	0	80	1345	10	
##	Rhizosolenia.alata	R..fragilissima	R..hebetata.semispina	R..styliformis	
## 1	90	30	105	10	
## 2	120	20	340	0	
## 3	25	10	43	0	
## 4	3	0	10	0	
## 5	0	0	5	0	
## 6	5	0	0	0	
##	Thalassionema.elegans	Thalassiosira.spp.	Thalassiothrix.antarctica		
## 1	275	4345	2845		
## 2	5520	2840	4440		
## 3	262	423	138		
## 4	180	213	0		
## 5	280	140	0		
## 6	260	85	0		
##	Thalassiothrix.delicatula	Tropidoneis.antarctica	T..belgicae		
## 1	0	200	70		
## 2	0	170	225		
## 3	0	12	38		
## 4	0	3	32		
## 5	0	0	5		
## 6	0	0	0		
##	T..medierranea	Tropidoneis.belgicae	T..fusiformis	Unidentified.diatoms	
## 1	0	0	0	25	
## 2	0	0	0	40	
## 3	0	0	0	10	

## 4	0	0	0	0
## 5	0	0	0	60
## 6	0	0	0	20
##	Dinoflagellates	Amphidinium.hadai.	Amphidinium.turbo.	Amphidinium.sp..1
## 1	NA	0	0	0
## 2	NA	0	0	0
## 3	NA	0	4	36
## 4	NA	228	0	124
## 5	NA	160	0	480
## 6	NA	40	0	460
##	Amphidinium.sp..2	Amphidinium.sp..3	Amphidoma.sp.1	
## 1	0	0	0	
## 2	0	0	0	
## 3	0	0	0	
## 4	4	0	24	
## 5	20	0	1660	
## 6	0	0	260	
##	Ceratium.arietinum.v..glacilentun.	C.furca	Ceratium.fusus	C..lineatum
## 1		0	0	5
## 2		0	0	20
## 3		0	0	36
## 4		0	0	193
## 5		5	0	720
## 6		0	5	560
##	C.massiliense	C..pentagonum	Cladopyxis.brachiolata	Cochlodinium.sp..1
## 1	0	0	0	0
## 2	0	35	0	0
## 3	0	0	0	4
## 4	0	0	0	0
## 5	10	0	0	0
## 6	0	0	0	0
##	Dinophysis.hastata	D..meteori	D.ovum	Dinophysis.parvula.
## 1	0	0	0	10
## 2	0	0	0	0
## 3	0	0	0	4
## 4	0	5	0	0
## 5	0	0	0	0
## 6	5	0	20	0
##	D..ruudii.	D..schroederi	Dinophysis.sp.1	Dissodinium.bicorne
## 1	0	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	20	0	60	14
## 6	0	10	0	5
##	Gonyaulax.minuta	G..polygramma	Gymnodinium.guttula.	G..minor
## 1	0	0	0	0
## 2	0	0	0	0
## 3	6	0	0	0
## 4	4	8	8	3140
## 5	0	0	0	2960
## 6	50	0	0	1100
##	Gymnodinium.spp.	Gyrodinium.pp.	Gyrodinium.glaciale.	Oxytoxum.curvatum.
## 1	0	0	0	0

## 2	70	0	0	0		
## 3	446	0	154	10		
## 4	2472	0	816	0		
## 5	6520	0	3500	0		
## 6	8200	0	4950	20		
##	0..coronatum	0..glacile	0..gracile	0..laticeps	0..longiceps	0..nanum
## 1	0	0	0	0	0	0
## 2	0	0	0	0	0	135
## 3	0	0	0	20	0	1468
## 4	0	0	0	296	4	2280
## 5	0	0	920	3980	0	12660
## 6	0	0	410	5680	0	11230
##	0..parvum	0..scolopax	0..sublatum	0..turbo	0..variabile	0..viride
## 1	0	0	0	0	0	0
## 2	0	0	0	0	0	0
## 3	0	0	0	0	48	0
## 4	0	0	4	0	70	4
## 5	60	20	0	0	2300	0
## 6	50	0	0	20	5460	0
##	Oxytoxum.sp.1	Podalampas.spinifera	Pronoctiluca.pelagica			
## 1	0	0	0			
## 2	0	0	10			
## 3	0	0	14			
## 4	0	0	4			
## 5	1780	5	5			
## 6	1210	0	0			
##	Prorocentrum.micans.	Prorocentrum.obtusidens	P..rostratum	P..triestinum.		
## 1	0	0	0	0		
## 2	0	0	0	0		
## 3	2	0	0	36		
## 4	0	285	0	132		
## 5	0	3210	60	1060		
## 6	0	2140	0	550		
##	Protoperidinium.conicum	Protoperidinium.elegantissimum.	P..minusculum			
## 1	0	0	0			
## 2	0	10	0			
## 3	0	0	0			
## 4	0	0	0			
## 5	20	0	20			
## 6	20	0	10			
##	P..orientale	P..pallidum	P..pellcidum	P..sourniai	P.tenuissimum	
## 1	0	0	0	0	0	
## 2	0	20	0	0	0	
## 3	0	10	0	0	0	
## 4	16	0	0	16	8	
## 5	20	9	32	40	5	
## 6	20	0	85	40	25	
##	P..tristylum.	P..trochoideum	Protoperidinium.spp.	Scrippsiella.spp.		
## 1	0	0	0	0		
## 2	0	0	0	30		
## 3	0	0	20	16		
## 4	4	0	24	38		
## 5	0	470	260	460		
## 6	0	260	80	1020		

```
## Unidentified.dinoflagellates Silicoflagellates.Distephanus.speculus
## 1 10 310
## 2 60 220
## 3 44 48
## 4 116 0
## 5 440 5
## 6 90 10

Antarc[,3] <-NULL #Get rid of diatoms column
Antarc[is.na(Antarc)] <- 0 #Get rid of pesky NA's.
AntarcDiversity <- Create.Diversity.I(Antarc, 3, ncol(Antarc))#Use nifty function to
#make diversity indices
View(AntarcDiversity)
SplitData(Antarc, "Location", "") #Splits up data into the 4 regions

## <environment: R_GlobalEnv>

#Antarctic
Antarctic <- Antarctic[, 3:ncol(Antarctic)] #take only the phytoplankton here
Antarctic <- Antarctic[-which(Antarctic==0)] #get rid of 0's
Antarctic <- Antarctic[!is.na(Antarctic)] #get rid of NA's
#Middle
Middle <- Middle[, 3:ncol(Middle)]
Middle <- colSums(Middle) #add columns
Middle <- Middle[-which(Middle==0)]
Middle <- Middle[!is.na(Middle)]
#Subantarctic
Subantarctic <- Subantarctic[, 3:ncol(Subantarctic)]
Subantarctic <- colSums(Subantarctic)
Subantarctic <- Subantarctic[-which(Subantarctic==0)]
Subantarctic <- Subantarctic[!is.na(Subantarctic)]
#Subtropical
Subtropical <- Subtropical[, 3:ncol(Subtropical)]
Subtropical <- colSums(Subtropical)
Subtropical<- Subtropical[-which(Subtropical==0)]
Subtropical<- Subtropical[!is.na(Subtropical)]
```

A heatmap plot of the abundance of different species in this data set.

```
#Create big dataframe
require(ggplot2)

## Loading required package: ggplot2
require(gplots)

## Loading required package: gplots

##
## Attaching package: 'gplots'

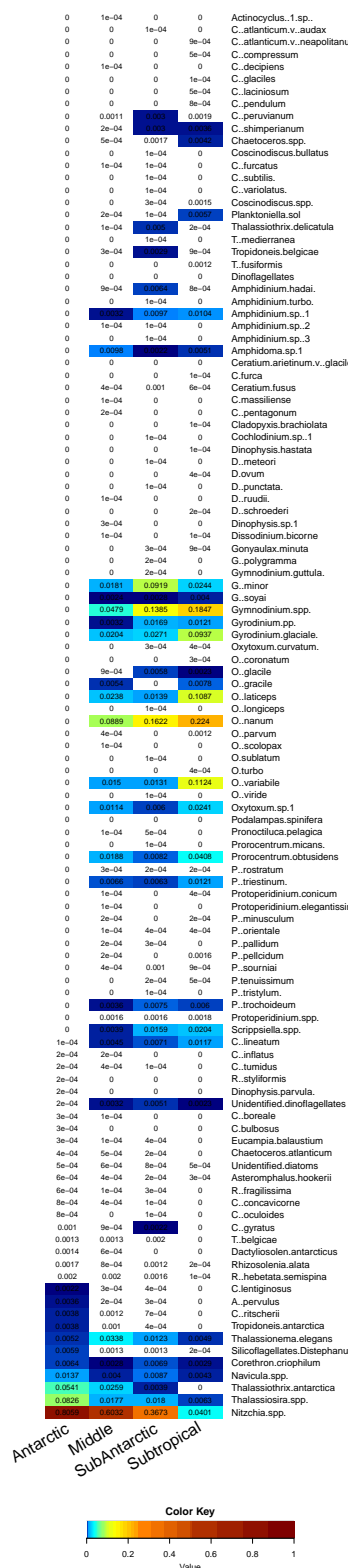
## The following object is masked from 'package:stats':
##
## lowess
```

```
require(iNEXT) #Rarefaction package by Chao
```

```
## Loading required package: iNEXT
```

```
#
Antarc2 <- Antarc[,3:ncol(Antarc)]
Antarc2<- rbind(Antarc2[1,], colSums(Antarc2[c(2,5,8),]), colSums(Antarc2[c(3,4,7),]), colSums(Antarc2[
Antarc2<- t(round(apply(Antarc2, 1, function(x) x/sum(x)), digits=4)) #simplified for easy plotting,
#now Antarc3 each abundance is a percentage of the total cells found per region
row.names(Antarc2) <- c("Antarctic", "Middle", "SubAntarctic", "Subtropical")
Antarc2 <- t(Antarc2)
Antarc2 <- Antarc2[order(Antarc2[,1]),decreasing=TRUE]
Antarc2[is.na(Antarc2)] <- 0
jet.colors <- colorRampPalette(c("#00007F", "#007FFF", "cyan", "#7FFF7F", "yellow",
"#FF7F00", "#7F0000"))
col <-jet.colors(48)
col <-c("white", col)
# (optional) defines the color breaks manually for a "skewed" color transition
col_breaks = c(seq(0, .001, length=1),
                seq(.002, .01, length=10),
                seq(.011,.05,length=10),
                seq(.051,.1,length=10),
                seq(.11, .25, length=9),
                seq(.251, .5, length=5),
                seq(.51, 1, length=5))
#Heatmap plotting function, just add matrix!
AbundHeatmap <- function(matrix, col, breaks) {
heatmap.2(matrix,
           cellnote = matrix, # same data set for cell labels
           offsetCol = 0.5,
           notecol="black",      # change font color of cell labels to black
           trace="none",        # turns off trace lines inside the heat map
           margins =c(9,9),# widens margins around plot
           cexCol=2.2,
           cexRow=1.3,
           srtCol=30,
           col=col,             # use on color palette defined earlier
           breaks= breaks,      # enable color transition at specified limits
           dendrogram="none",    # only draw a row dendrogram
           Colv=NA,              # turn off column clustering
           key=TRUE,
           key.title = NULL,
           symbreaks=TRUE,
           symkey=FALSE,
           density.info="none",
           densadj = 0.25,
           Rowv = FALSE,
           #denscol="black",
           keysize=.2,
           #( "bottom.margin", "left.margin", "top.margin", "left.margin" )
           #key.par=list(mar=c(7,.5,3,1)),
           #lmat -- added 2 lattice sections (5 and 6) for padding
           lmat=rbind( c(0, 3, 0), c(2,1,0), c(0,4,0) ),
           lhei=c(.5, 5, .3),
```

```
      lwid = c(.5,3,.5))}  
par(mar = rep(1, 4))  
AbundHeatmap(Antarc2, col=col, breaks=col_breaks)
```



##Rarefaction of the prepared data

Note: rarefaction seems problematic here

This script takes a very long time to run because of the high number of cells counted in the Antarctic samples. The rarefaction index may be working but likely does not give meaningful results at this scale. Asymptote of the rarefaction curve is reached very quickly relative to the total number of individuals. Consideration of the method, scaling, and applicability should be made here. Try running the script overnight, if it is still not working a subsampling approach of the population may need to be designed.

Script turned off for functionality

```
AntarcR <- list("Antarctic"=Antarctic, "Subantarctic"=Subantarctic, "Middle"=Middle, "Subtropical"=Subtropical) # #iNEXT creates rarefaction dataframe and metadata for a group of samples
A0 <- iNEXT(AntarcR, q=0, datatype="abundance") #richness A1 <- iNEXT(AntarcR, q=1, datatype="abundance") #Shannon based diversity
```

Rarefaction of abundance data showing richness

```
AntarcRareRichness <- ggiNEXT(A0, type = 1, facet.var = "none", color.var = "site") AntarcRareRichness+ylab("Richness") #Rarefaction of abundance data showing diversity (H1)
AntarcRareH <- ggiNEXT(A1, type = 1, facet.var = "none", color.var = "site") AntarcRareH+ylab("1st order Hill #") #Coverage plotted for 1st order hill number
AntarcRareCoverage <- ggiNEXT(A1, type = 2, facet.var = "none", color.var = "site") AntarcRareCoverage+ylab("Sample coverage H1") “
```