On the use of the De Finetti ternary diagrams to show the species composition of free and FAD-associated tuna schools in the Atlantic and Indian oceans

By Alain Fonteneau, Emmanuel Chassot, Sofia Ortega-García, Alicia Delgado de Molina and Nicolas Bez

Summary

This paper shows the difficulty to visually to illustrate the variability of the species composition of sampled sets landed by purse seiners in the Indian and Atlantic Oceans. When large numbers of sets have been sampled during the last 30 years, it is still quite difficult to visualize the heterogeneity of their species compositions. This paper proposes to use the ternary plots proposed by De Finetti to solve this problem. These ternary plots have been drawn as examples on the Free and FAD schools sets sampled in the Indian and Atlantic oceans during the 2002-2008 period. These ternary plots have been drawn as examples using 3 different plotting methods, but other alternate improved methods could also be envisaged following the discussion of this working document

Résumé

Cet article rapelle la difficulté de montrer visuellement l'hétérogénéité de la composition spécifique des bancs de thons échantillonnés sur les senneurs dans les océans Indien et Atlantique. Si de très nombreux échantillons ont été réalisés sur ces flottilles depuis près de 30 ans, il s'avère difficile de bien visualiser la variabilité des composition spécifiques ainsi mises en évidence. Cet article propose d'employer à cet effet les diagrammes ternaires de De Finetti. Ces diagrammes sont ainsi réalisés à titre d'exemple sur des échantillons provenant de bancs libres et de bancs associés à des DCP dans les océans Indien et Atlantique durant la période 2002-2008. Ces diagrammes sont dessinés à titre d'exemple selon 3 méthodes graphiques différentes, d'autres méthodes de dessin pouvant être aussi envisagées suite à la discussion de ce document de travail.

1- Introduction

EU scientists and their scientific partners from Cote d'Ivoire, Senegal, Seychelles and Madagascar have been conducting since the early 1980s a systematic sampling of the species composition of tuna landings by purse seiners. The main goal of this sampling is to obtain an accurate estimation of the species composition of tuna landings by purse seiners (Bard & Vendeville 1986; Pianet et al. 2000). The sampling results have been routinely used to correct the species composition of the EU purse seiner log books in order to provide better estimates

of annual catch by species to the ICCAT and the IOTC. These data form the basis of stock assessment models and other statistical analyses. However, while the data processing in use combines all fish samples taken in a specific quarter and for large fishing areas, the results of such a sampling programme can also be used for estimating the species composition variability among fishing sets. The 'SPECIES' file that describes the set by set tuna species composition is a very interesting dataset that allows to synthesize much of the information provided by raw sampling data (Fonteneau et al. 2007). However, when scientists want to analyze the variability of species composition, they are facing the difficulty of showing a complex variability of at least 3 components, i.e. the 3 major tuna species targeted by purse seiners in equatorial waters: yellowfin (*Thunnus albacares*), skipjack (*Katsuwonus pelamis)*, and bigeye (*Thunnus obesus*). Other tuna species such as albacore (*Thunnus alalunga*) are also landed and sampled but in very low proportion.

We consider that the historical method proposed by the Italian statistician Bruno de Finetti (1906–1985) based on ternary plots (De Finetti 1937) and has been used to selection problems mainly en genetic area, would be a convenient tool to show the variability of these observed species composition:

"A ternary plot (or ternary graph, triangle plot, simplex plot) or de Finetti diagram is a barycentric plot on three variables which sum to a constant. It graphically depicts the ratios of the three variables as positions in an equilateral triangle. It is used in petrology, mineralogy, metallurgy, and other physical sciences to show the compositions of systems composed of three species" (Wikipedia).

The main aim of this paper is to describe a few technical methods that can be used to plot such De Finetti diagrams in order to show the tuna species composition of sampled schools.

2- Materials and method

The basic data used are from the so-called "SPECIES" files, i.e. the detailed species composition of all the sampled sets collected in the Indian and Atlantic oceans. This study is based on the subset of recent data collected during the 2002-2008 period, on the landing or transhipment operation of the EU and associated purse seiners. This multispecies sampling scheme on purse seiners and the framework of this 'SPECIES' file have been described and analyzed by document SCRS/2009/131.

The average species composition types (showing the percentage of each species in the purse seine catches) estimated during the 2002-2008 period on the FAD and free school sets in the Atlantic and Indian Ocean are shown in figure 1.
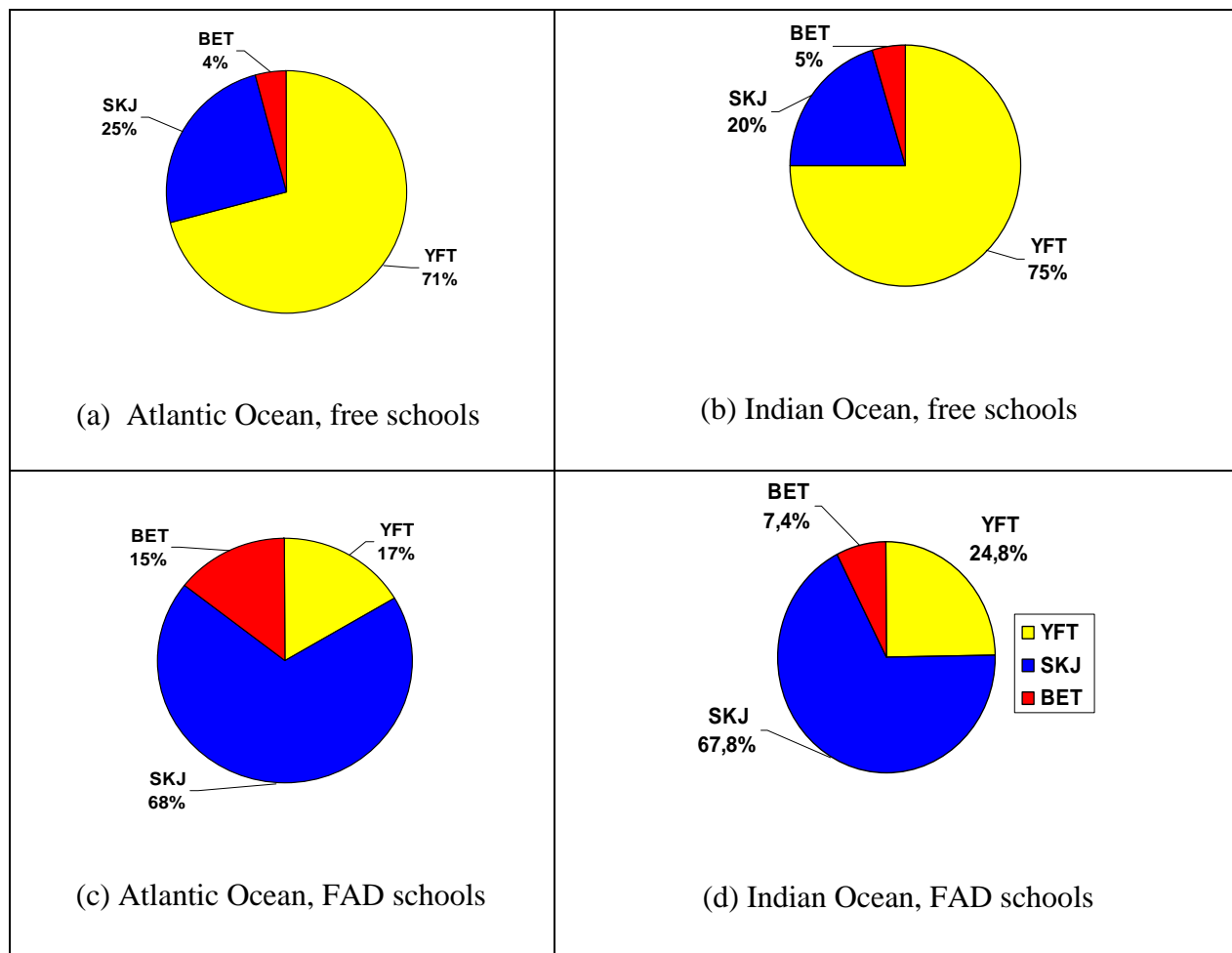
Figure 1 a-d: Estimated species composition of FADs and free school catches in the Atlantic and Indian oceans during the 2002-2008 period. YFT = yellowfin; SKJ = skipjack; BET = bigeye

The current issue of interest is to show the heterogeneity of these species compositions at any level of time and area stratum as well as for each fishing mode (FAD or free schools). Here, we propose that the ternary plots proposed by De Finetti (1937) can be an efficient way to visually illustrate the complexity and variability of these observed species compositions. However, as the ternary plots can only handle 3 parameters, the catches of the 3 dominant target species, i.e. yellowfin, skipjack and bigeye, have been kept as the 3 basic parameters of the triangles drawn. In a ternary diagram, the reading of the percentage of each component is given based on the lines parallel to the triangle sides (Figure 2).
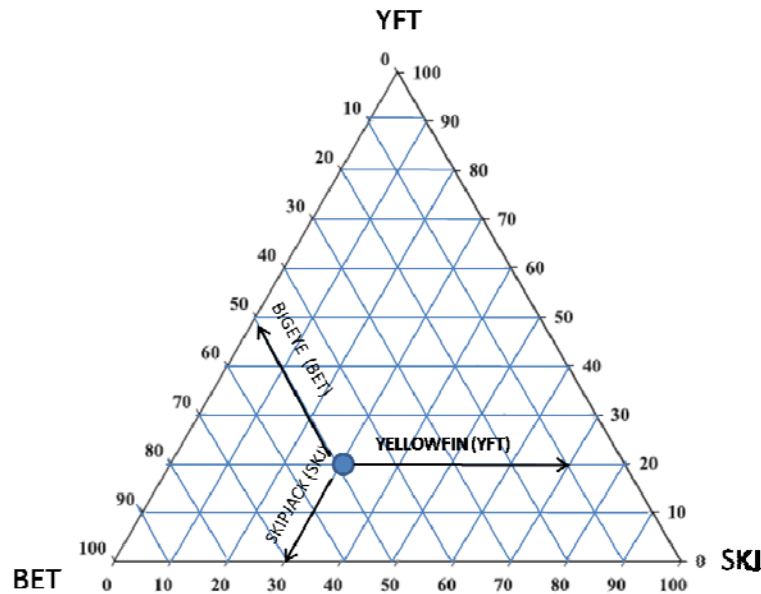
Figure 2: Example of ternary diagram illustrating the reading of the percentage of each tuna in the species composition of a fishing set or tuna school.

Plotting functions of ternary diagrams have been implemented in some specific libraries of statistical software. For instance the 'triax.plot' function of the 'plotrix' package (Lemon et al. 2009) and the 'ternaryplot' function of the vcd package (Meyer et al. 2009) are available in R (R Development Core Team 2008). Such plotting functions can be used to plot 1 observed species composition but fail to adequately represent a large dataset of species compositions. Here, we propose 2 different methods based on De Finetti ternary diagrams to fully include the range of species compositions observed in a single plot:

(i) Drawing the lines of estimated iso-percentage of the frequency of observed species composition. These lines can be estimated by interpolation methods such as kernel density or kriging. Each range of percentage can be then correspondingly coloured, e.g. using a range of colours from light blue to dark blue, pink, and red. An example of *ad hoc* plot function is given in appendix.

(ii) Showing the relative proportion or percentage of the samples observed for each cell of species composition, the percentage in each circle being calculated by classes of 10% in the species composition of each species.

3-Results

Four figures, giving the frequency of species composition sampled in the Atlantic and Indian oceans for free schools (Figures 3-4) and FAD-associated schools in the same oceans (Figures 5-6) during 2001-2008 have been drawn to show typical results.

Figure 2 and 3 based on estimated iso-percentages and figures 4 and 5 showing the frequency of samples by 10% classes of species composition illustrate the great similarities of free school samples between the Indian and Atlantic Oceans during 2002-2008, and the general predominance of pure yellowfin schools and numerous schools with a high percentage of yellowfin.
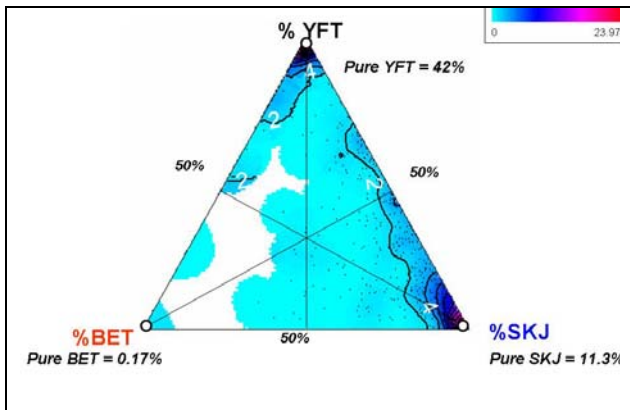
Figure 3: Percentage of occurrence in the observed species composition of the Free schools sets, Atlantic Ocean, shown by iso lines
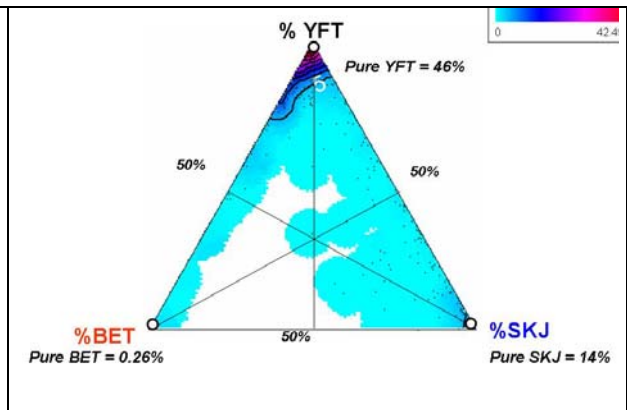

Figure 4: Percentage of occurrence in the observed species composition of the free schools sets, Indian Ocean, shown by iso lines
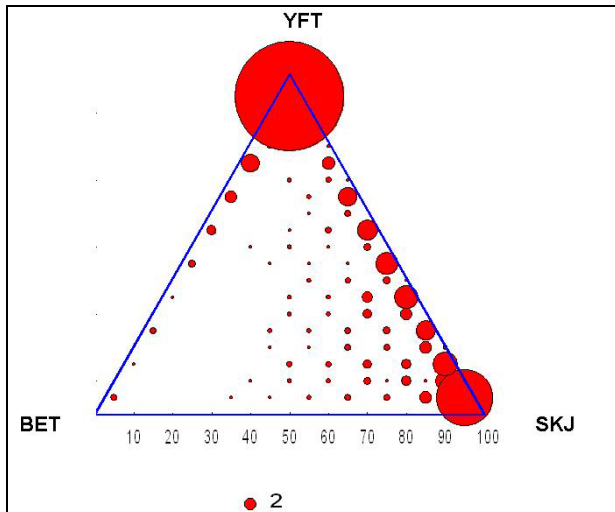

Figure 5: Percentage of occurrence in the observed species composition of the free school sets, Atlantic Ocean, shown by pies with an area proportional to frequency of the sampling composition
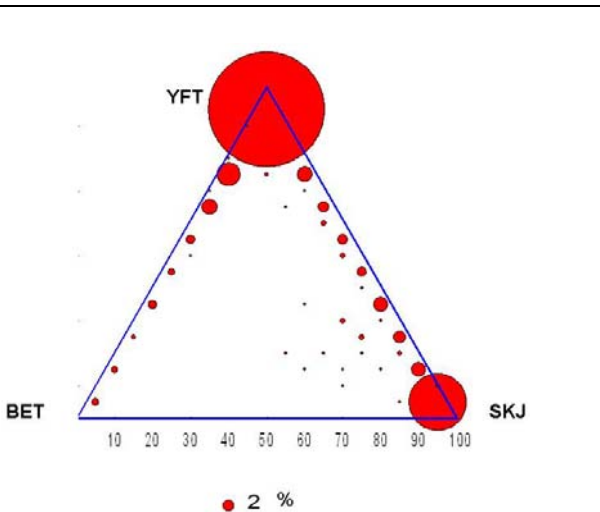

Figure 6: Percentage of occurrence in the observed species composition of the free school sets, Indian Ocean, shown by pies with an area proportional to frequency of the sampling composition

However it can also be noted that when pure SKJ schools are in equivalent percentage in both oceans (11 and 14%), schools with high % of skipjack have been commonly observed in the Atlantic Ocean (northern area), when they have been seldom observed in the Indian Ocean.

On the opposite, figures 7-10 show that the frequency in the observed species composition of FAD samples has been very similar in the 2 oceans over 2002-2008. Figures 4 and 5 also indicate that FAD-associated sets are most often composed of a mixture of the 3 species; bigeye tuna being most often present in the samples, but always taken in relatively low proportion of the catches.
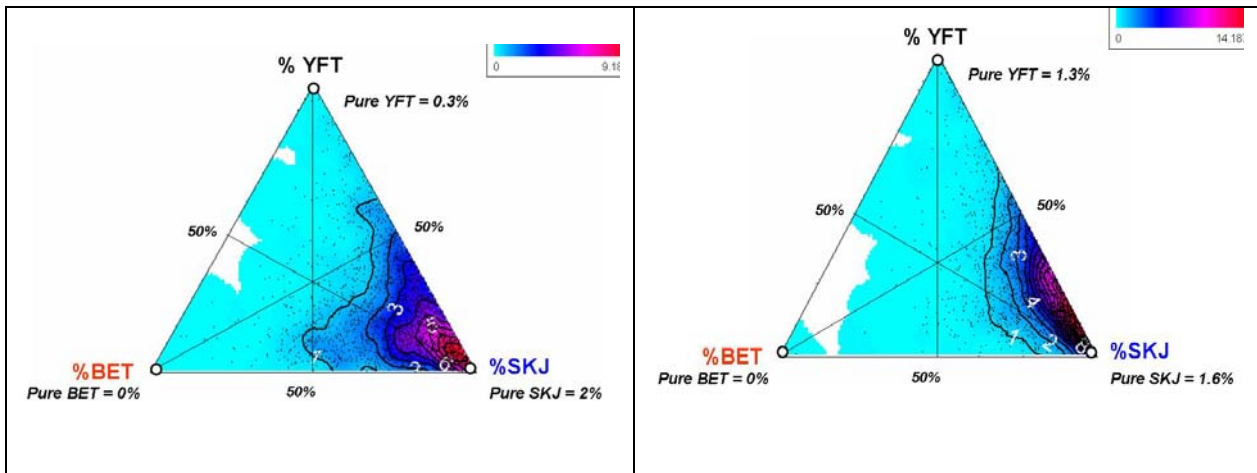
Figure 7: Percentage of occurrence in the observed species composition of the FAD-associated school sets, Atlantic Ocean, shown by iso lines

Figure 8: Percentage of occurrence in the observed species composition of the FAD-associated school sets, Indian Ocean, shown by iso lines
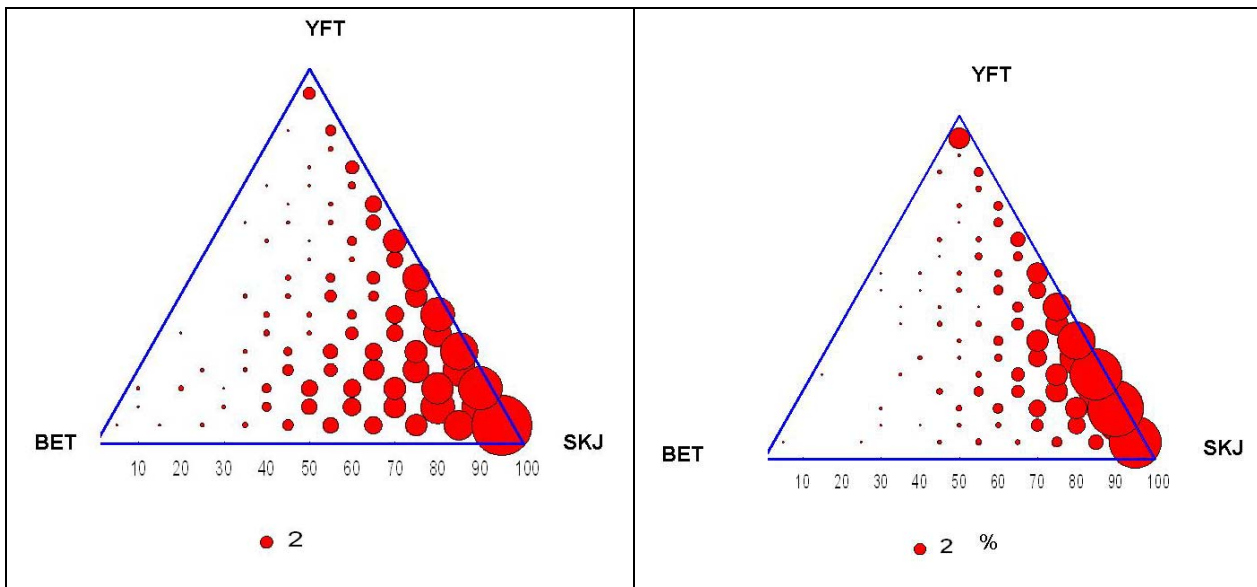

Figure 9: Percentage of occurrence in the observed species composition of the FAD-associated school sets, Atlantic Ocean, shown by pies with an area proportional to the frequency of sampling composition

Figure 10: Percentage of occurrence in the observed species composition of the FAD-associated school sets, Indian Ocean, shown by pies with an area proportional to the frequency of sampling composition

Figure 7 to 9 also show that bigeye has been observed in most samples and in both oceans, but most often showing low percentages in its relative weight.

The same type of De Finetti diagram could also be plotted by time and area stratum, for instance by quarters and in the sampling areas that are used in the data processing of the species and size composition of purse seine catches (Pallares and Petit 1998, Pallarés and Hallier 1997). Examples of such species composition heterogeneity in these strata are for instance shown for the Somalia area in the third quarter 3 of 2006 (Figure 7). Such in-depth

further analysis should *inter alia* allow to better explore the seasonal and between years variability of the observed species composition of the sampled schools.

4-Conclusion

This first work would allow concluding that the De Finetti ternary plots appear to be an interesting way to synthetize the complex variability of the species composition of the sampled tuna schools. The present method and results are still provisional, but they are already well showing the great analogy and the minor differences in the species composition of the FAD and free schools sampled in the Atlantic and Indian oceans as well as the great heterogeneity of the species composition between free and FAD associated schools. Further work would be needed to simplify and to improve the building and the presentation of these De Finetti triangles used to show the species composition of tuna samples. This further study should also generalize the use of these ternary plots in a wide range of strata, allowing to analyze better and to illustrate the time and space variability of tuna schools in the Indian and Atlantic oceans. The same illustrative method could also be used to compare the heterogeneity of species composition observed at a worldwide scale in the purse seine fisheries.

Acknowledgements

References

Bard, F. & Vendeville, P. (1986), Note sur l'échantillonnage plurispécifique des thons tropicaux au port d'Abidjan, Col. Vol. Sci. Pap. ICCAT 25, 37-45

B. de Finetti. La Prévision: Ses Lois Logiques, ses Sources Subjectives. In Annales de *l'Institut Henri Poincaré 7*, pages 1-68. Paris, 1937. Note: Translated into English by Henry E. Kyburg Jr., Foresight: Its Logical Laws, its Subjective Sources. In Henry E. Kyburg Jr. and Howard E. Smokler (1964, Eds.), Studies in Subjective Probability, 53-118, Wiley, New York

Fonteneau A., J. Ariz, A. Delgado de Molina, J. Dorizo, V. Lucas and R. Pianet 2007. Species composition of FAD and free swimming schools fished by purse seiners in the Western Indian Ocean during the period 1991-2006. Document IOTC-2007-WPTT-05

Lemon, J., Bolker, B., Oom, S., Klein, E., Rowlinsgon, B., Wickham, H., Tyagi, .A., Eterradossi, O., Grothendieck, G., Toews, M., Kane, J., Cheetham, M., Turner, R., Whittoft, C., Stander, J., Petzoldt, T., Duursma, R., Biancotto, E., and Levy, O. 2009. Plotrix: Various plotting functions. R package version 2.7.

Meyer, D., Zeileis, A., and Hornik, K. 2009. vcd: Visualizing Categorical Data. R package version 1.2-4.

Pallarés, P. & Hallier, J-P. (1997), Analyse du schéma d'échantillonnage multispécifique des thonidés tropicaux (Programme n° 95/37), EU technical report, IEO/ORSTOM

Pallarés, P. & Petit, C. (1998), Tropical tunas: new sampling and data processing strategy for estimating the composition of catches by species and sizes, Col. Vol. Sci. Pap. ICCAT 48(2), 230-246

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org

Appendix 1. R function to plot a ternary diagram based on the kernel density of tuna species composition

```
##############################################################################
### The dataset is available on authors request and includes 3 columns of catch by species,
            i.e. yellowfin, skipjack, and bigeye
### This is an ad hoc method to plot a ternary diagram based on a re-expression of the (a,b,c)
            values of species composition into a (x,y) regular bi-plot. The triangle shape
            is made through the use of white polygons, which is not very satisfying
##############################################################################
### Libraries
require(MASS)
require(fields)

### Path and data reading
rep.data <- "/home/data"
setwd(rep.data)
fad_0208 <- read.csv("data-example-Finetti.csv", sep = ",", header = T)

#Simple function to converts composition into (x,y) coordinates
# a = %SKJ
# b = %BET
# c = %YFT
duo.coordinate.f = function(a,b,c){
y <- sin(pi/3) * c
x <- y/tan(pi/3) + a
res <- as.data.frame(cbind(x,y))
res
}

#Estimates proportions and kernel density
fadsby <- as.data.frame(cbind(fad_0208$SKJ,fad_0208$BET,fad_0208$YFT))
fadsby.prop <- fadsby/rowSums(fadsby)
names(fadsby.prop) <- c("SKJ","BET","YFT")
dataset <- duo.coordinate.f(fadsby.prop$SKJ,fadsby.prop$BET,fadsby.prop$YFT)
density.data <- kde2d(dataset[,1],dataset[,2], lims = c(-0.2,1.2,-0.1,.9), n = 100)

#Plot the density
image.plot(density.data,nlevel = 500,axes = F, bty = "n", horizontal = T, legend.lab = "",
            legend.mar = 6, cex = 1.5, legend.shrink = .7)

#Mask to delete all values out of the Triangle
polygon(x = c(0,0.5,0.5,-0.4,-0.4),y = c(0,sin(pi/3),sin(pi/3)+0.3,sin(pi/3)+0.3,0), col =
            "white", border = NA)
polygon(x = c(1.5,1,0.5,0.5,1.4), y = c(0,0,sin(pi/3),sin(pi/3)+0.2,sin(pi/3)+0.2), col =
            "white", border = NA)
polygon(x = c(-0.4,1.4,1.4,-0.4), y = c(0,0,-0.12,-0.12), col = "white", border = NA)

#Grid
values <- seq(0,1,by = 0.1)
```

```
for (i in (1:11)){
    abline(h = i * sin(pi/3)/10, col = "white")
    abline(b = sin(pi/3)*2, a = -2*sin(pi/3)*values[i], col = "white")
    abline(b = -2*sin(pi/3), a = 2*sin(pi/3)*values[i], col = "white")
    }

#Axes
axis(1, seq(0, 1, by = 0.1), tick = T, labels = seq(0,100,by = 10),cex.axis = 1.3, line = -2)
segments(-0.015,2*sin(pi/3)*-0.015+0.025,.49,2*sin(pi/3)*0.49+0.025)
segments(1.015,-2*sin(pi/3)*1.015+2*sin(pi/3)+0.025,.51,-
                2*sin(pi/3)*0.51+2*sin(pi/3)+0.025)

#Ticks
y <- seq(0,1,by = .1)*sin(pi/3)
x <- (y -0.025)/(2*sin(pi/3))
segments(x-.01,y,x,y,col = "black")
text(x = x-0.06, y = y, labels = as.character(rev(seq(0,100,by = 10))),cex =1.3)
x1 <- (y - (2*sin(pi/3)+0.025))/(-2*sin(pi/3))
segments(x1+.01,y,x1,y,col = "black")
text(x = x1+0.06,y = y, labels = as.character(seq(0,100,by = 10)), cex = 1.3)
text(x = 0.1, y = 0.5, labels = "BET", cex = 1.4)
text(x = .9, y = 0.5, labels = "YFT", cex = 1.4)
mtext("SKJ", 1, cex = 1.4, line = 1)
mtext("Species composition density", 1,cex = 1.2, line = 3.2)
```
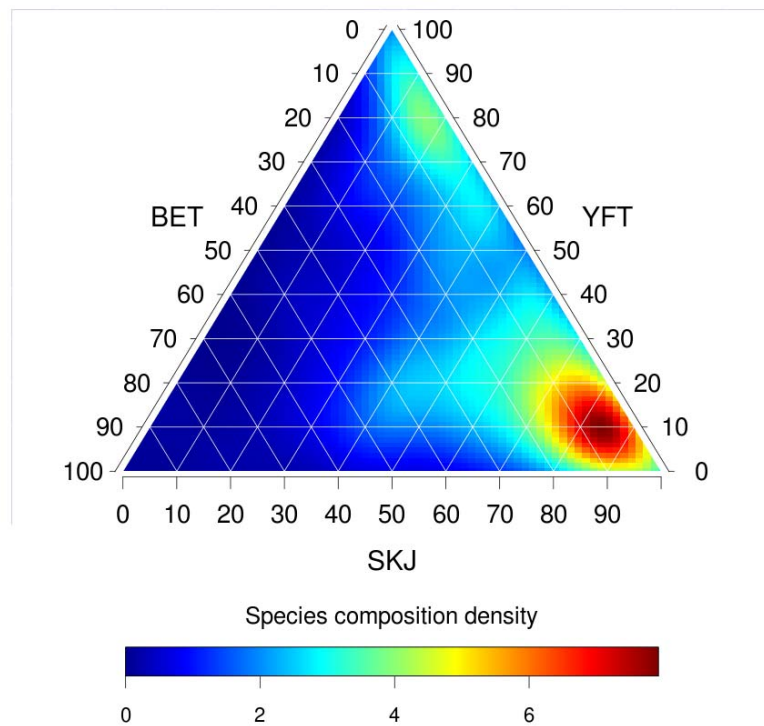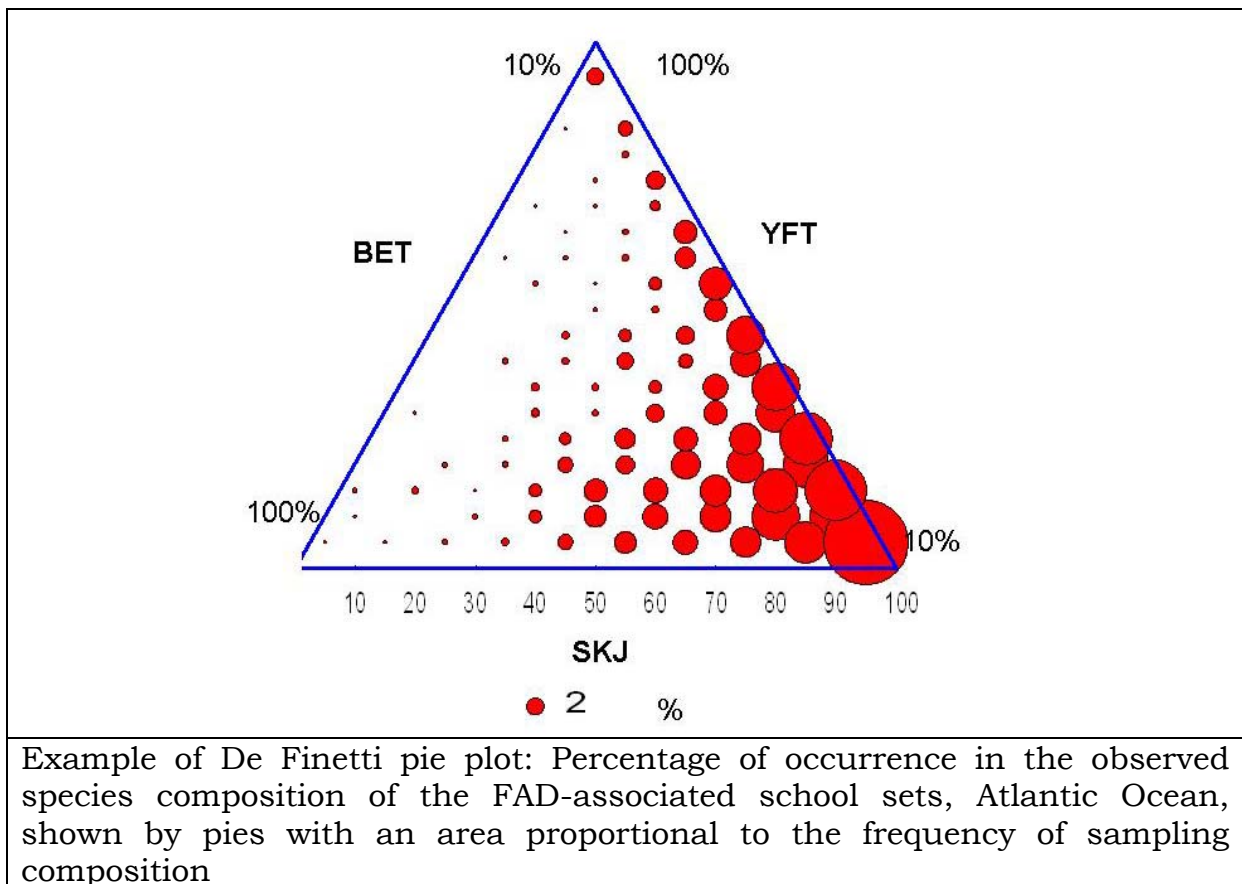
Appendix 2: Software creating the De Finetti pie diagrams

An ad hoc home made software has been prepared to do the De Finetti ternary pie diagram. This software has been written in 2 FORTRAN programs:

(1) The 1st program, BILFINETTI, reads the observed species composition of the sampled tuna sets, and it calculates by 10% classes of yellowfin, skipjack and bigeye, the numbers of observed relative weights for all the combinations of the 3 species by classes of 10% of relative weight.

(2) The 2nd program simply draws a De Finetti triangle where each circle corresponds to the observed frequencies of the 3 species, the scale of the pies being chosen by the user. The output of this program is an HPGL file that can easily be incorporated in a WORD or PPT text.

Example of De Finetti pie plot: Percentage of occurrence in the observed species composition of the FAD-associated school sets, Atlantic Ocean, shown by pies with an area proportional to the frequency of sampling composition

This small software can be obtained from the authors