Australian Government
Bureau of Rural Sciences

Southern Bluefin Tuna. Exploration of catch per unit effort standardisation models

Veronica Boero Rodriguez

June 2010

Boero Rodriguez, V 2010, *Southern Bluefin Tuna. Exploration of catch per unit effort standardisation models*, Bureau of Rural Sciences, Canberra.

© Commonwealth of Australia 2010

This work is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced by any process without prior written permission from the Commonwealth. Requests and inquiries concerning reproduction and rights should be addressed to the Commonwealth Copyright Administration, Attorney General's Department, Robert Garran Offices, National Circuit, Barton ACT 2600 or posted at http://www.ag.gov.au/cca.

The Australian Government acting through the Bureau of Rural Sciences has exercised due care and skill in the preparation and compilation of the information and data set out in this publication. Notwithstanding, the Bureau of Rural Sciences, its employees and advisers disclaim all liability, including liability for negligence, for any loss, damage, injury, expense or cost incurred by any person as a result of accessing, using or relying upon any of the information or data set out in this publication to the maximum extent permitted by law.

Postal address:
Bureau of Rural Sciences
GPO Box 858
Canberra, ACT 2601

Ph: 1800 020 157
Fax: 02 6272 2330
Email: info@brs.gov.au
Internet: http://www.brs.gov.au

Table of contents

Figures

.

Introduction

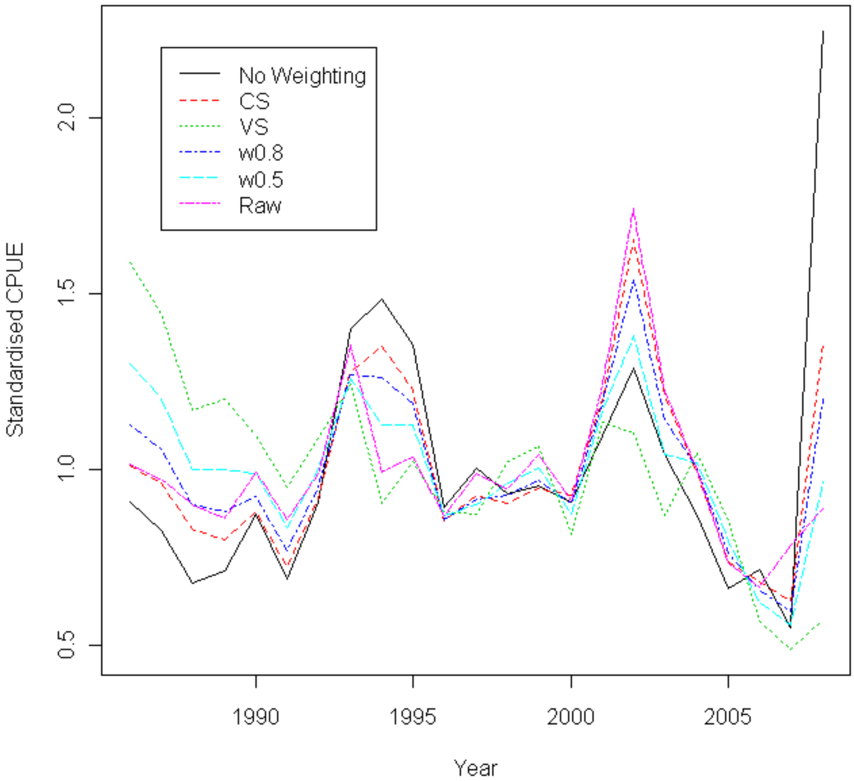
The Southern Bluefin Tuna (SBT) standardised catch per unit effort (CPUE) series is used as an index of abundance in the operating model. This series presents a large trend change in 2008, especially in the unweighted series (Figure 1). However, in the raw CPUE series (without standardisation or weightings) a much smaller trend change is produced in 2007 and then another increase in 2008 (Figure 1). As the standardised CPUE is considered an index of abundance, it is expected to be fairly smooth without large changes in trend from one year to the next but this is not the case in this standardised CPUE series. The fact that this large change in trend is present in the standardised series, but not in the raw series indicates that the model may be producing the trend rather than the data.

This paper covers a data exploration to find an appropriate model. A basic analysis of fleet dynamics is produced to see if there is a change in fleet behaviour, and a formulation of alternative models to the one currently used. The R code to produce the analysis included in this paper is detailed in Appendix A.

The data used for this analysis is Japan's logbook data from 1986 to 2007 aggregated into 5 by 5 degrees squares and the RTMP data in 2008 multiplied by a factor to make it equivalent to the logbook data, also aggregated into 5 by 5 degrees squares **as supplied in March 2010**. The RTMP data does not include New Zealand joint venture operations. The variables in the data set are:

- Year
- Month (from April to September)
- Area (CCSBT areas 4 to 9, areas 5 and 6 amalgamated)
- Lat5 (latitude 5 degrees)
- Long5 (longitude 5 degrees)
- Hook (number of hooks in shot)
- Age4p (number of SBT 4 years old plus caught in shot)
- Tuna1
- Tuna2
- N_BET (number of Big Eye tuna caught in shot)
- N_YFT (number of Yellow Fin tuna caught in shot)
- CPUE (age4p per 1000s hooks)
- logCPUE
- BETcpue5
- YFTcpue5

Figure 1: CPUE series 1986-2008



Data exploration to find variables interactions

Plots of mean CPUE by year and area, Lat5 and month might help to provide a better insight on the year interaction with the other possible effects, although it is not possible to see how all these factors interact together. Also, a plot of area by month might help to understand the area and month interaction.

Looking at the plot of mean CPUE by year and area, it is possible to see that in 2008 with respect to 2007, there is a large increase in mean CPUE in Area 7 and to a lesser extent in Area 4, at the same time that there is a considerable decrease in Area 56 (Figure 2). The decrease in 2008 in Area 56 could be due to the New Zealand data not being included in the dataset analysed. However, over the years analysed there is large variability in mean CPUE by year in each area and the differences between areas are within that variability which might indicate that there is no strong interaction between area and year.

Figure 2: CPUE by year and area

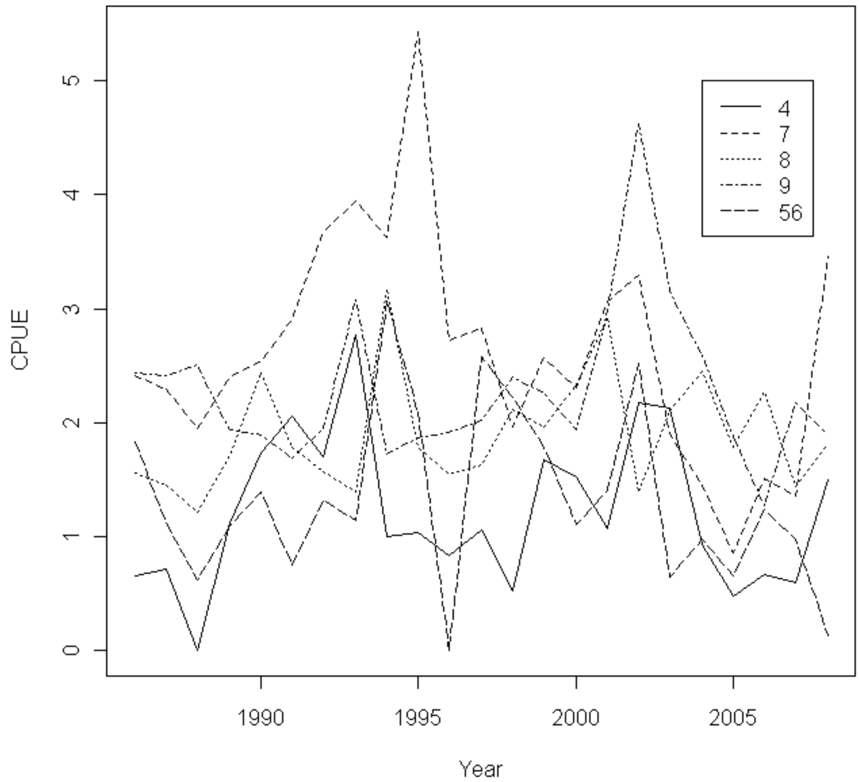
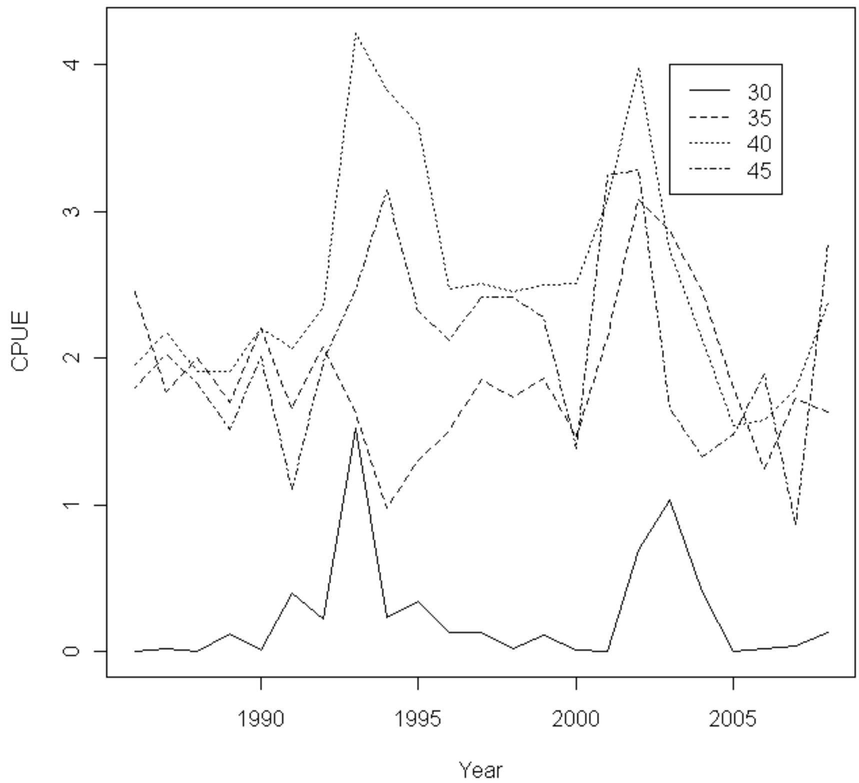


Figure 3 shows a large increase in mean CPUE in 2008 with respect to 2007 in Lat5 45 and to a lesser extent in Lat5 40. However, since 1990 mean CPUE in Lat 5 35, 40 and 45 has been variable and the change between the last two years in Lat5 45 was within that variability.

Figure 3: CPUE by year and Lat5



Also, there is large variability in mean CPUE by year and month (Figure 4). The largest difference between 2007 and 2008 is a pronounced decrease in mean CPUE in April, but again this decrease is well within the variability.

Figure 4: CPUE by year and month

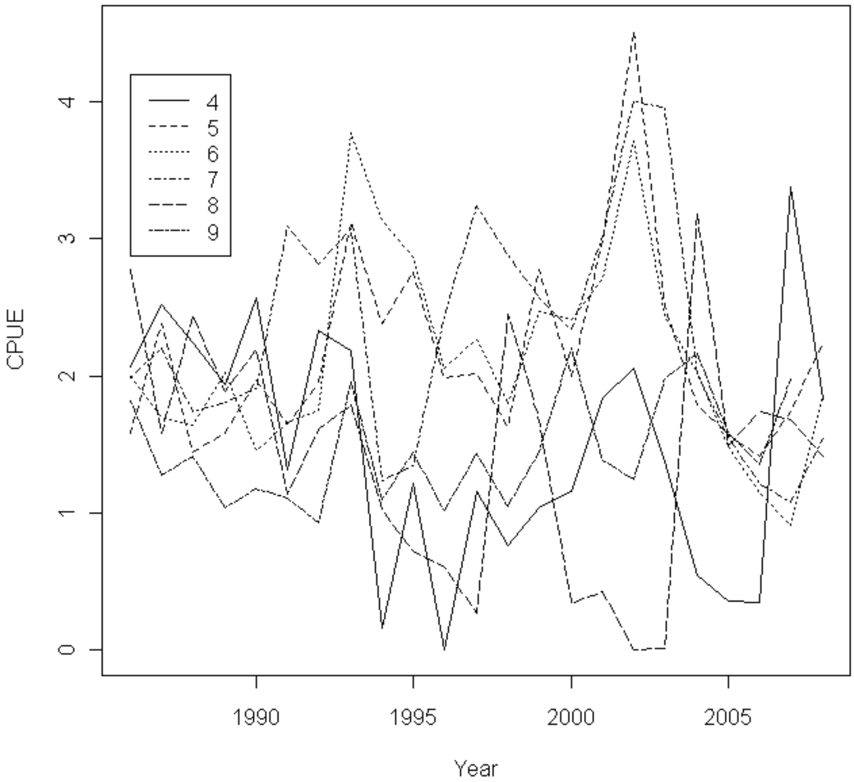
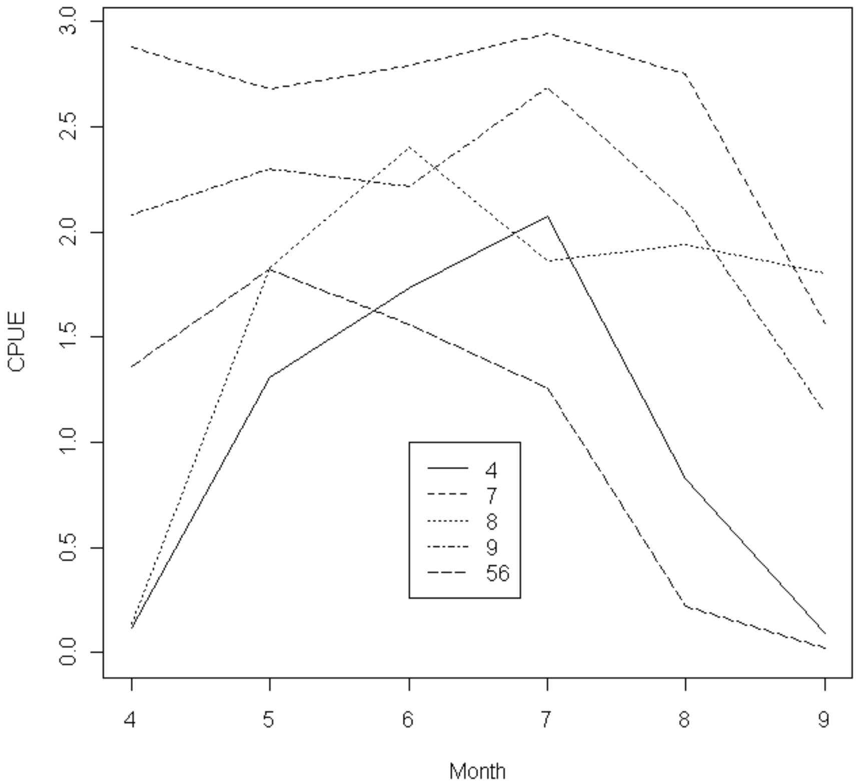


Figure 5: CPUE by month and area



There seems to be an interaction between area and month as the CPUE trends are fairly smooth and different in different areas (Figure 5).

Analysis of effort distribution

A thorough analysis of fishing effort distribution was done by Giannini (2009) and Tomoyuki (2009). The main conclusions from these papers were that the variability in fishing effort distribution is rather large so it is not possible to conclude clearly that the fleet behaviour has changed. However, the number of 5 by 5 and 1 by 1 degree squares fished each year presents a decreasing trend. This paper includes a more basic analysis of effort distribution and proposes models which might be able to account for this in the case that fleet behaviour change is present.

Plots of mean year effort by area, Lat5 and month might give an insight on fleet fishing behaviour. Effort by area is quite variable and does not seem to be a clear change over the years (Figure 6). However, in Area 9 effort is fairly cyclic and in 2007 and 2008 was among its lowest levels.

There is large variability in effort in Lat5 35 and 40 over the years, while in Lat5 30 and 45 effort is fairly stable (Figure 7).

There is also large variability over the years in the mean effort by month, especially in May, June and July (Figure 8).

The large variability in effort distribution makes it difficult to recognise any clear change in effort distribution and hence a change in fishing fleet behaviour.

Figure 6: Mean year effort by area

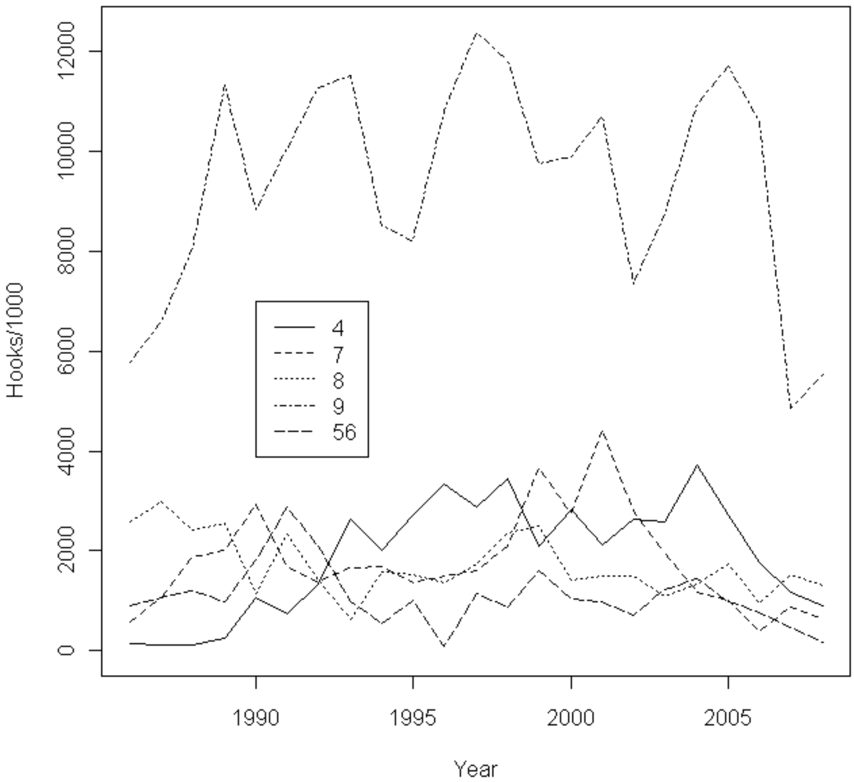


Figure 7: Mean year effort by Lat5

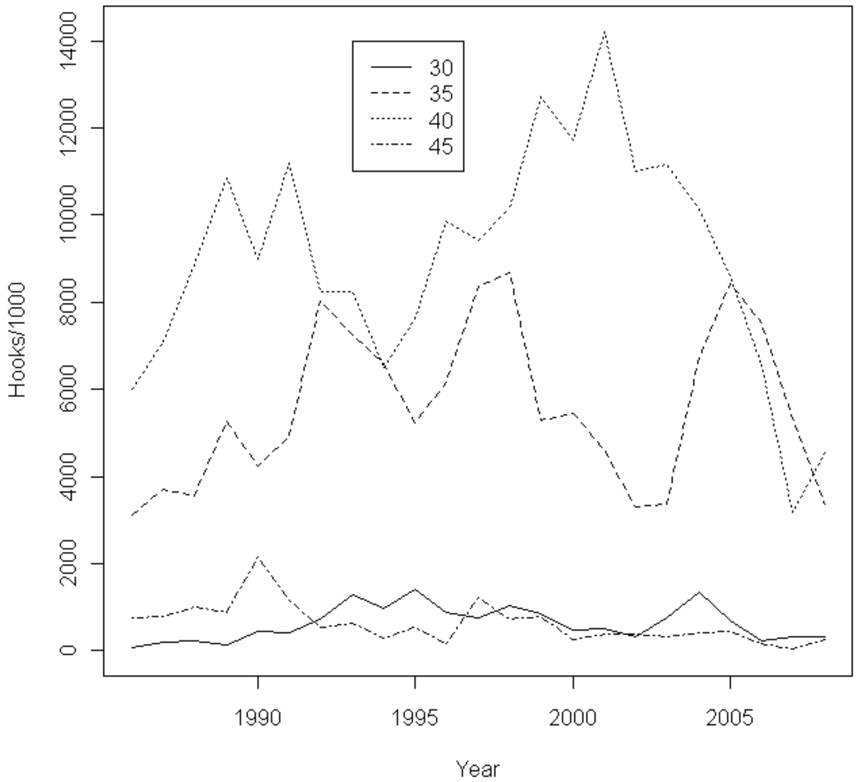
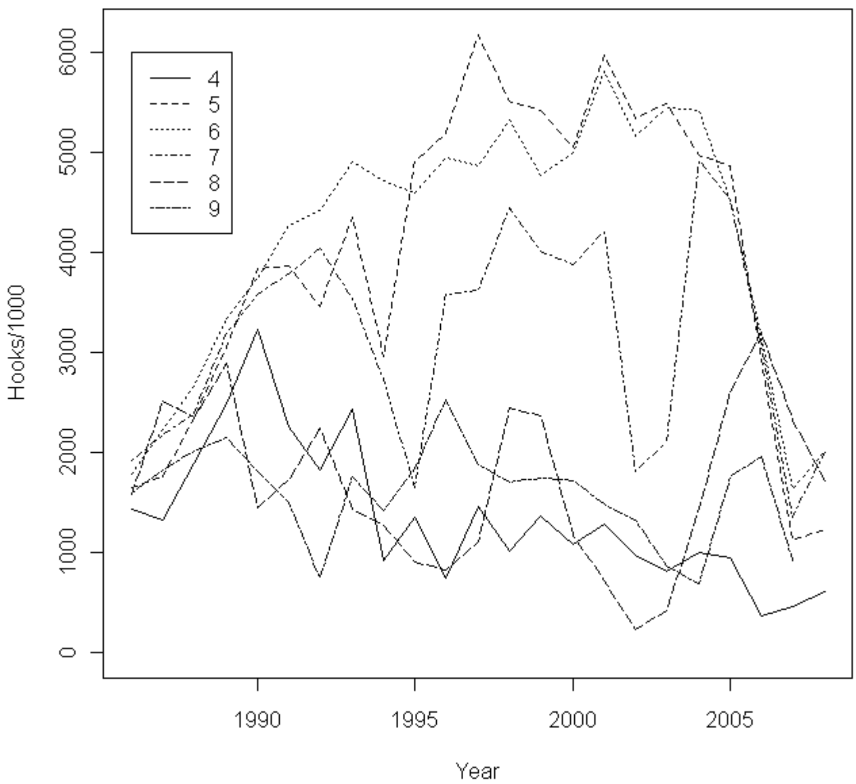


Figure 8: Mean year effort by month



Vessel and grid effects

Even though Tomoyuki (2009a) showed that including a fixed vessel effect did not significantly change the series trend, a vessel effect, either random or fixed, might contribute to model change in fishing fleet behaviour as it incorporates in the models the differences between vessels. In the case of a random effect, it models the between vessel variability separately from the within vessel variability. Also, to model the shrinkage in number of squares fished (5 by 5 or 1 by 1 degree), the factors Area and Lat5 could be replaced by a grid square factor.

The analysis including a grid effect model for the 5 by 5 degrees squares factor is included in this paper. R code to perform this analysis is included in Appendix A. Appendix A also includes R code for models including vessel and grid effects at the 1 by 1 degree square resolution. However, these models might fail due to issues that are not foreseeable without exploring the dataset at a finer spatial resolution.

Models

The previous sections suggest that:

1. No interactions with year should be included in the model
2. An interaction between month and area should be included in the model
3. A random vessel effect should be included in the model
4. A fixed grid square effect could replace the area and Lat5 factors.

The standardised and raw CPUE series from 1986 to 2008 using a model with fixed effects year, area, month and Lat5, no year interactions and an interaction between month and area are shown in Figure 9. These standardised series smooth to different degrees (depending on the weightings) the peaks in the raw series.

Figure 10 shows the raw CPUE series together with standardised series using models that include year, month and 5 by 5 degrees squares as fixed effects and in the other 5 by 5 degrees squares are included as random effects. There is almost no difference in the final series between the two models.

In Appendix A, the R code is provided to fit models including vessel effects (fixed and random) and also 1 by 1 degree squares as fixed or random. The most logical model for interpretability would be a model with squares as fixed effects and vessels as random effects. However, this model might not work, as it requires many degrees of freedom and might arise singularity in the variance matrix.

Figure 9: Raw and standardised CPUE series using a model with no year interactions.

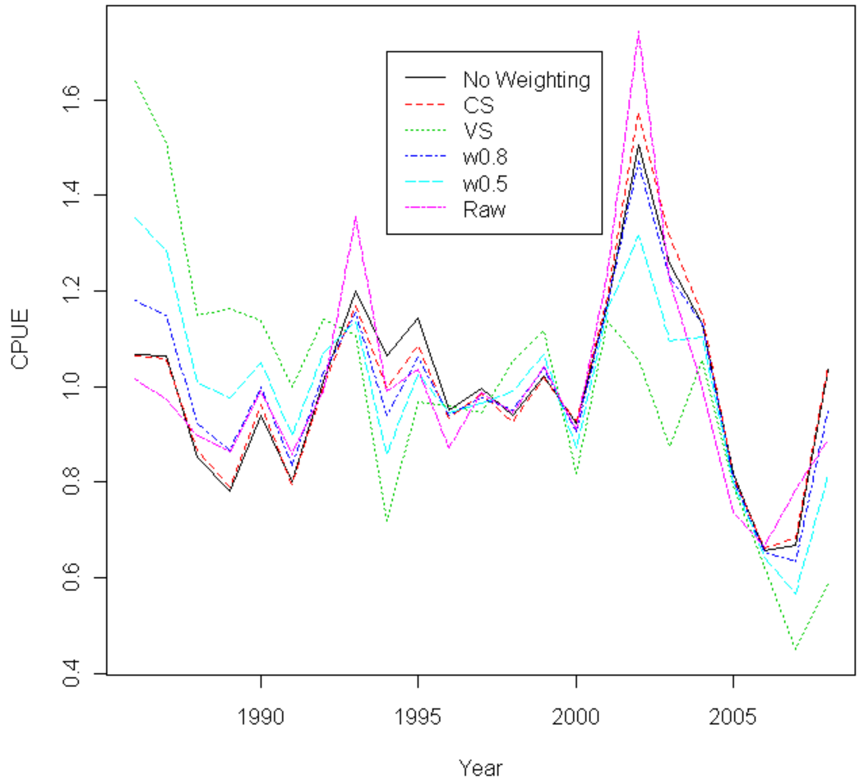
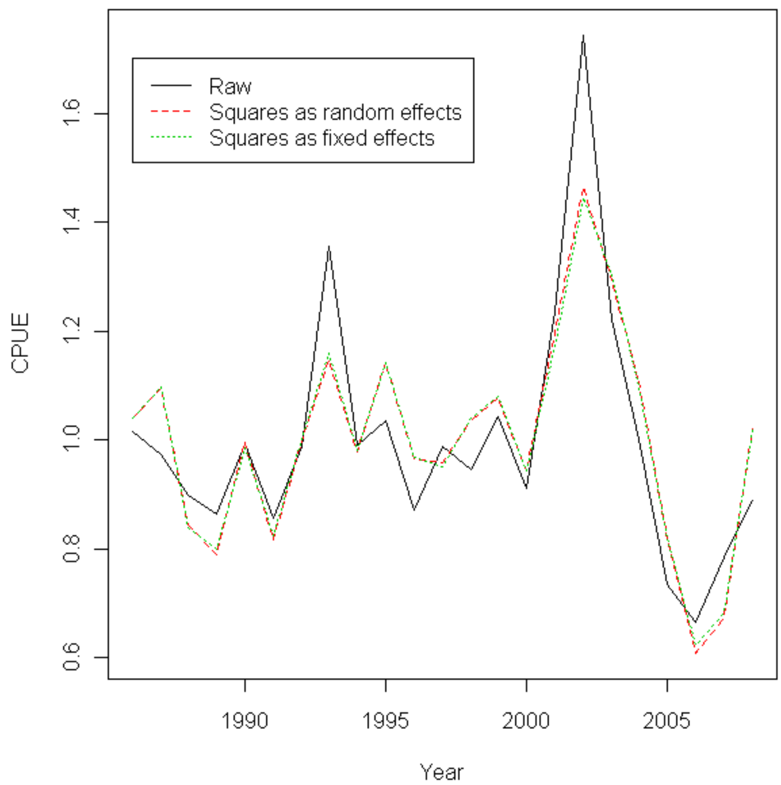


Figure 10: Raw and standardised CPUE series using 5 by 5 degrees squares in the models.



Conclusions

The fact that the large change in trend is present in the standardised series but not in the raw series indicates that the model may be producing the results rather than the data.

There is not much evidence of interactions between year and the other fixed effects.

To model to some extent the variability in effort it is recommended that a random vessel effect is included in the model. Also, to model the shrinkage in the number of squares fished, it is recommended that the area and Lat5 effects are replaced by a fixed grid square effect.

References

Giannini, F 2009,*Working paper b3.1: Examining concentration patterns of SBT CPUE*, Bureau of Rural Sciences, Canberra.

Tomoyoki, I 2009, *Working paper: CPUE2009-b3.2. Number of 5x5 and 1x1 degree squares operated*, NRIFSF, May.

Tomoyoki, I 2009a, *Working paper: CPUE2009-b10.1. Including fixed vessel effects in CPUE standardisation and comparison by two data sets between 5x5 and shot by shot*, NRIFSF, May.

Appendix A

R code

```
#set the appropriate directory
setwd("P:\\Cross_Programme\\FisheryStats\\Southern Bluefin Tuna\\SBT CPUE
Standardisation\\CPUE 2009_Japan")
#libraries needed
library(nlme)
library(MASS)
#read the data
#logbook data
logData<-read.csv("5x5Core_02112009log.csv")
#RTMP data
RTMPdata<-read.csv("5x5core02112009RTMP.csv")
#CS—read constant squares weightings
aiCS <- read.table("aridx_cs_2009SAG.prn" ,header=T)
#VS—read variable squares weightings
aiVS <- read.table("aridx_vs_2009SAG.prn" ,header=T)
#delta constant
log.add <- 0.2
#factor variables
logData$year<-factor(logData$Year)
logData$month<-factor(logData$Month)
logData$area<-factor(logData$Area)
logData$lat5<-factor(logData$Lat5)
RTMPdata$year<-factor(RTMPdata$Year)
RTMPdata$month<-factor(RTMPdata$Month)
RTMPdata$area<-factor(RTMPdata$Area)
RTMPdata$lat5<-factor(RTMPdata$Lat5)
#create 5x5 square variable
RTMPdata$sq<-paste(RTMPdata$Lat5,RTMPdata$Lon5,sep="")
RTMPdata$sq<-factor(RTMPdata$sq)
#raw cpue RTMP
cpue.year<-aggregate(RTMPdata$CPUE,by=list(RTMPdata$year),mean)
cpue.r<-cpue.year$x/mean(cpue.year$x)
yearv<-c(seq(from=1986,to=2008,by=1))
month<-c(seq(from=4,to=9,by=1))
lat5<-c(seq(from=30,to=45,by=5))

#plots of mean cpue by year and area, month, Lat5
l5c<-aggregate(RTMPdata$CPUE,by=list(RTMPdata$year,RTMPdata$lat5),mean)
plot(yearv,l5c$x[1:23],ylim=c(0,max(l5c$x)),type="l",lty=1,xlab="Year",ylab="Raw CPUE")
```

```
lines(yearv,l5c$x[24:46],lty=2)
lines(yearv,l5c$x[47:69],lty=3)
lines(yearv,l5c$x[70:92],lty=4)
legend(2003,4,c("30","35","40","45"),lty=c(1:4))
areafc<-aggregate(RTMPdata$CPUE,by=list(RTMPdata$year,RTMPdata$area),mean)
plot(yearv,areafc$x[1:23],ylim=c(0,max(areafc$x)),type="l",lty=1,xlab="Year",ylab="Raw
CPUE")
lines(yearv,areafc$x[24:46],lty=2)
lines(yearv,areafc$x[47:69],lty=3)
lines(yearv,areafc$x[70:92],lty=4)
lines(yearv,areafc$x[93:115],lty=5)
legend(2004,5,c("4","7","8","9","56"),lty=c(1:5))
moc<-aggregate(RTMPdata$CPUE,by=list(RTMPdata$year,RTMPdata$month),mean)
plot(yearv,moc$x[1:23],ylim=c(0,max(moc$x)),type="l",lty=1,xlab="Year",ylab="Raw CPUE")
lines(yearv,moc$x[24:46],lty=2)
lines(yearv,moc$x[47:69],lty=3)
lines(yearv,moc$x[70:92],lty=4)
lines(yearv,moc$x[93:115],lty=5)
lines(yearv,moc$x[116:138],lty=6)
legend(1986,4.2,c("4","5","6","7","8","9"),lty=c(1:6))
#plot of mean cpue by month and area
moac<-aggregate(RTMPdata$CPUE,by=list(RTMPdata$month,RTMPdata$area),mean)
plot(month,moac$x[1:6],ylim=c(0,max(moac$x)),type="l",lty=1,xlab="Year",ylab="Raw CPUE")
lines(month,moac$x[7:12],lty=2)
lines(month,moac$x[13:18],lty=3)
lines(month,moac$x[19:24],lty=4)
lines(month,moac$x[25:30],lty=5)
legend(6,1,c("4","7","8","9","56"),lty=c(1:5))

#plots of mean effeort by year and area, month, Lat5
RTMPdata$effort<-RTMPdata$hook/1000
l5e<-aggregate(RTMPdata$effort,by=list(RTMPdata$year,RTMPdata$lat5),sum)
plot(yearv,l5e$x[1:23],ylim=c(0,max(l5e$x)),type="l",lty=1,xlab="Year",ylab="Hooks/1000")
lines(yearv,l5e$x[24:46],lty=2)
lines(yearv,l5e$x[47:69],lty=3)
lines(yearv,l5e$x[70:92],lty=4)
legend(1993,14000,c("30","35","40","45"),lty=c(1:4))
areafe<-aggregate(RTMPdata$effort,by=list(RTMPdata$year,RTMPdata$area),sum)
plot(yearv,areafe$x[1:23],ylim=c(0,max(areafe$x)),type="l",lty=1,xlab="Year",ylab="Hooks/1000
")
lines(yearv,areafe$x[24:46],lty=2)
lines(yearv,areafe$x[47:69],lty=3)
lines(yearv,areafe$x[70:92],lty=4)
lines(yearv,areafe$x[93:115],lty=5)
```

```
legend(1990,7000,c("4","7","8","9","56"),lty=c(1:5))
moe<-aggregate(RTMPdata$effort,by=list(RTMPdata$year,RTMPdata$month),sum)
plot(yearv,moe$x[1:23],ylim=c(0,max(moe$x)),type="l",lty=1,xlab="Year",ylab="Hooks/1000")
lines(yearv,moe$x[24:46],lty=2)
lines(yearv,moe$x[47:69],lty=3)
lines(yearv,moe$x[70:92],lty=4)
lines(yearv,moe$x[93:115],lty=5)
lines(yearv,moe$x[116:138],lty=6)
legend(1986,6000,c("4","5","6","7","8","9"),lty=c(1:6))
```

#current model

```
RTMPm<-
lm(logCPUE~year+month+area+lat5+BETcpue5+YFTcpue5+month*area+year*lat5+year*area,
data=RTMPdata)
#check residuals
par(mfrow=c(1,2), pty="s")
plot(RTMPm$fitted,RTMPm$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
#calculate mean BET cpue for all combinations of year, month, area and Lat5 levels
BETav<-aggregate(RTMPdata$BETcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$area, RTMPdata$lat5), mean)
BET <- data.frame(
  year = BETav$Group.1,
  month = BETav$Group.2,
  area = BETav$Group.3,
  lat5 = BETav$Group.4,
  BETcpue5 = BETav$x)
#as per BET for YFT
YFTav<-aggregate(RTMPdata$YFTcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$area, RTMPdata$lat5), mean)
YFT <- data.frame(
  year = YFTav$Group.1,
  month = YFTav$Group.2,
  area = YFTav$Group.3,
  lat5 = YFTav$Group.4,
  YFTcpue5 = YFTav$x)
#create the matrix for prediction and convert NAs to zeros
pred_data<-
expand.grid(year=levels(RTMPdata$year),month=levels(RTMPdata$month),area=levels(RTMPdat
a$area),lat5=levels(RTMPdata$lat5))
pred<-merge(pred_data,BET, all=T)
pred<-merge(pred,YFT, all=T)
```

```
for(i in 1:length(pred$BETcpue5)){
if(is.na(pred$BETcpue5[i])) pred$BETcpue5[i]<-0
if(is.na(pred$YFTcpue5[i])) pred$YFTcpue5[i]<-0}
#calculate predictions
pred$cpue<-exp(predict(RTMPm,newdata=pred)+((0.6715)^2)/2)-log.add
#calculate unweighted series
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpueUW<-cpue$x/mean(cpue$x)
#merge CS weightings and convert NAs to zeros
CS <- data.frame(
    month = aiCS$MO,
    area = aiCS$A,
    lat5 = aiCS$LAT,
    cs = aiCS$AI)
pred<-merge(pred,CS,all=T)
for(i in 1:length(pred$cs)){
if(is.na(pred$cs[i])) pred$cs[i]<-0}
#calculate CS series
pred$cpuec<-pred$cpue*pred$cs
cpuec<-aggregate(pred$cpuec,by=list(pred$year),mean)
cpueCS<-cpuec$x/mean(cpuec$x)
#merge VS weightings and convert NAs to zeros
aiVS <- subset(aiVS, X.YR > 1985)
aiVS <- subset(aiVS, LAT < 50)
VS <- data.frame(
    year = aiVS$X.YR,
    month = aiVS$MO,
    area = aiVS$A,
    lat5 = aiVS$LAT,
    vs = aiVS$AI)
VS <- subset(VS, year > 1985)
pred<-merge(pred,VS,all=T)
for(i in 1:length(pred$vs)){
if(is.na(pred$vs[i])) pred$vs[i]<-0}
#calculate VS series
pred$cpuev<-pred$cpue*pred$vs
cpuev<-aggregate(pred$cpuev,by=list(pred$year),mean,na.rm=T)
cpueVS<-cpuev$x/mean(cpuev$x)
#calculate w0.8 and w0.5 series
cpue08 <- 0.8*cpueCS+(1-0.8)*cpueVS
cpue05 <- 0.5*cpueCS+(1-0.5)*cpueVS
#plot all series
par(mfrow=c(1,1))
```

```
plot(yearv,cpueUW,ylim=c(min(cpueUW,cpueCS,cpueVS,cpue08,cpue05,
cpue.r),max(cpueUW,cpueCS,cpueVS,cpue08,cpue05,cpue.r)),ylab="Standardised CPUE",
xlab="Year",type="l",lty=1,col=1)
lines(yearv,cpueCS,lty=2,col=2)
lines(yearv,cpueVS,lty=3,col=3)
lines(yearv,cpue08,lty=4,col=4)
lines(yearv,cpue05,lty=5,col=5)
lines(yearv,cpue.r,lty=6,col=6)
legend(1987,2.2, c("No Weighting","CS","VS","w0.8","w0.5","Raw"), lty=c(1:6) ,col=c(1:6))
```

#no year interactions

```
RTMPm1<-lm(logCPUE~year+month+area+lat5+BETcpue5+YFTcpue5+month*area,
data=RTMPdata)
par(mfrow=c(1,2), pty="s")
plot(RTMPm1$fitted,RTMPm1$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm1$resid, main=" ")
abline(a=0,b=1)
BETav<-aggregate(RTMPdata$BETcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$area, RTMPdata$lat5), mean)
BET <- data.frame(
   year = BETav$Group.1,
   month = BETav$Group.2,
   area = BETav$Group.3,
   lat5 = BETav$Group.4,
   BETcpue5 = BETav$x)
YFTav<-aggregate(RTMPdata$YFTcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$area, RTMPdata$lat5), mean)
YFT <- data.frame(
   year = YFTav$Group.1,
   month = YFTav$Group.2,
   area = YFTav$Group.3,
   lat5 = YFTav$Group.4,
   YFTcpue5 = YFTav$x)
pred_data<-
expand.grid(year=levels(RTMPdata$year),month=levels(RTMPdata$month),area=levels(RTMPdat
a$area),lat5=levels(RTMPdata$lat5))
pred<-merge(pred_data,BET, all=T)
pred<-merge(pred,YFT, all=T)
for(i in 1:length(pred$BETcpue5)){
if(is.na(pred$BETcpue5[i])) pred$BETcpue5[i]<-0
if(is.na(pred$YFTcpue5[i])) pred$YFTcpue5[i]<-0}
pred$cpue<-exp(predict(RTMPm1,newdata=pred)+((0.7345)^2)/2)-log.add
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpueUW1<-cpue$x/mean(cpue$x)
CS <- data.frame(
```

```
    month = aiCS$MO,
    area = aiCS$A,
    lat5 = aiCS$LAT,
    cs = aiCS$AI)
pred<-merge(pred,CS,all=T)
for(i in 1:length(pred$cs)){
if(is.na(pred$cs[i])) pred$cs[i]<-0}
pred$cpuec<-pred$cpue*pred$cs
cpuec<-aggregate(pred$cpuec,by=list(pred$year),mean)
cpueCS1<-cpuec$x/mean(cpuec$x)
aiVS <- subset(aiVS, X.YR > 1985)
aiVS <- subset(aiVS, LAT < 50)
VS <- data.frame(
    year = aiVS$X.YR,
    month = aiVS$MO,
    area = aiVS$A,
    lat5 = aiVS$LAT,
    vs = aiVS$AI)
VS <- subset(VS, year > 1985)
pred<-merge(pred,VS,all=T)
for(i in 1:length(pred$vs)){
if(is.na(pred$vs[i])) pred$vs[i]<-0}
pred$cpuev<-pred$cpue*pred$vs
cpuev<-aggregate(pred$cpuev,by=list(pred$year),mean,na.rm=T)
cpueVS1<-cpuev$x/mean(cpuev$x)
cpue081 <- 0.8*cpueCS1+(1-0.8)*cpueVS1                    #w0.8
cpue051 <- 0.5*cpueCS1+(1-0.5)*cpueVS1
par(mfrow=c(1,1))
plot(yearv,cpueUW1,ylim=c(min(cpueUW1,cpueCS1,cpueVS1,cpue081,cpue051,
cpue.r),max(cpueUW1,cpueCS1,cpueVS1,cpue081,cpue051,cpue.r)),ylab="CPUE",
xlab="Year",type="l",lty=1,col=1)
lines(yearv,cpueCS1,lty=2,col=2)
lines(yearv,cpueVS1,lty=3,col=3)
lines(yearv,cpue081,lty=4,col=4)
lines(yearv,cpue051,lty=5,col=5)
lines(yearv,cpue.r,lty=6,col=6)
legend(1993,1.75, c("No Weighting","CS","VS","w0.8","w0.5","Raw"), lty=c(1:6) ,col=c(1:6))
```

#model with 5x5 squares as random effect

```
#model
RTMPm2<-lme(logCPUE~year+month+BETcpue5+YFTcpue5,random=~1|sq, data=RTMPdata)
par(mfrow=c(1,2), pty="s")
plot(RTMPm2$fitted,RTMPm2$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm2$resid, main=" ")
```

```
abline(a=0,b=1)
#constructing the standardised series
BETav<-aggregate(RTMPdata$BETcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$sq), mean)
BET <- data.frame(
  year = BETav$Group.1,
  month = BETav$Group.2,
  sq = BETav$Group.3,
  BETcpue5 = BETav$x)
YFTav<-aggregate(RTMPdata$YFTcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$sq), mean)
YFT <- data.frame(
  year = YFTav$Group.1,
  month = YFTav$Group.2,
  sq = YFTav$Group.3,
  YFTcpue5 = YFTav$x)
pred_data<-
expand.grid(year=levels(RTMPdata$year),month=levels(RTMPdata$month),sq=levels(RTMPdata
$sq))
pred<-merge(pred_data,BET, all=T)
pred<-merge(pred,YFT, all=T)
for(i in 1:length(pred$BETcpue5)){
if(is.na(pred$BETcpue5[i])) pred$BETcpue5[i]<-0
if(is.na(pred$YFTcpue5[i])) pred$YFTcpue5[i]<-0}
pred$cpue<-exp(predict(RTMPm2,newdata=pred,level=1)+((0.66542)^2)/2)-log.add
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpuesqr<-cpue$x/mean(cpue$x)
```

#squares as fixed effects

```
RTMPm3<-lm(logCPUE~year+month+BETcpue5+YFTcpue5+sq, data=RTMPdata)
plot(RTMPm3$fitted,RTMPm3$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm3$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
#use same prediction matrix as for random effects
pred$cpue<-exp(predict(RTMPm3,newdata=pred)+((0.6653)^2)/2)-log.add
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpuesqf<-cpue$x/mean(cpue$x)
par(mfrow=c(1,1))
plot(yearv,cpue.r,ylim=c(min(cpuesqr, cpuesqf, cpue.r),max(cpuesqr, cpuesqf,
cpue.r)),ylab="CPUE", xlab="Year",type="l",lty=1,col=1)
lines(yearv,cpuesqr,lty=2,col=2)
lines(yearv,cpuesqf,lty=3,col=3)
```

```
legend(1986,1.7, c("Raw", "Squares as random effects", "Squares as fixed effects"), lty=c(1:3)
,col=c(1:3))
```

#include vessel effects

#current model plus vessel fixed effects

```
RTMPm4<-
lm(logCPUE~year+month+area+lat5+BETcpue5+YFTcpue5+vessel+month*area+year*lat5+year
*area, data=RTMPdata)
#check residuals
plot(RTMPm4$fitted,RTMPm4$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm4$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
#calculate mean BET cpue for all combinations of year, month, area and Lat5 levels
BETav<-aggregate(RTMPdata$BETcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$area, RTMPdata$lat5, vessel), mean)
BET <- data.frame(
   year = BETav$Group.1,
   month = BETav$Group.2,
   area = BETav$Group.3,
   lat5 = BETav$Group.4,
   vessel=BETav$Group.5,
   BETcpue5 = BETav$x)
#as per BET for YFT
YFTav<-aggregate(RTMPdata$YFTcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$area, RTMPdata$lat5, vessel), mean)
YFT <- data.frame(
   year = YFTav$Group.1,
   month = YFTav$Group.2,
   area = YFTav$Group.3,
   lat5 = YFTav$Group.4,
   vessel= YFTav$Group.5,
   YFTcpue5 = YFTav$x)
#create the matrix for prediction and convert NAs to zeros
pred_data<-
expand.grid(year=levels(RTMPdata$year),month=levels(RTMPdata$month),area=levels(RTMPdat
a$area),lat5=levels(RTMPdata$lat5), vessel=levels(RTMPdata$vessel)
pred<-merge(pred_data,BET, all=T)
pred<-merge(pred,YFT, all=T)
for(i in 1:length(pred$BETcpue5)){
if(is.na(pred$BETcpue5[i])) pred$BETcpue5[i]<-0
if(is.na(pred$YFTcpue5[i])) pred$YFTcpue5[i]<-0}
#calculate predictions
pred$cpue<-exp(predict(RTMPm4,newdata=pred)+((0.6715)^2)/2)-log.add #replace 0.6715 by
model's standard error
```

```
#calculate unweighted series
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpueUW<-cpue$x/mean(cpue$x)
#merge CS weightings and convert NAs to zeros
CS <- data.frame(
    month = aiCS$MO,
    area = aiCS$A,
    lat5 = aiCS$LAT,
    cs = aiCS$AI)
pred<-merge(pred,CS,all=T)
for(i in 1:length(pred$cs)){
if(is.na(pred$cs[i])) pred$cs[i]<-0}
#calculate CS series
pred$cpuec<-pred$cpue*pred$cs
cpuec<-aggregate(pred$cpuec,by=list(pred$year),mean)
cpueCS<-cpuec$x/mean(cpuec$x)
#merge VS weightings and convert NAs to zeros
aiVS <- subset(aiVS, X.YR > 1985)
aiVS <- subset(aiVS, LAT < 50)
VS <- data.frame(
    year = aiVS$X.YR,
    month = aiVS$MO,
    area = aiVS$A,
    lat5 = aiVS$LAT,
    vs = aiVS$AI)
VS <- subset(VS, year > 1985)
pred<-merge(pred,VS,all=T)
for(i in 1:length(pred$vs)){
if(is.na(pred$vs[i])) pred$vs[i]<-0}
#calculate VS series
pred$cpuev<-pred$cpue*pred$vs
cpuev<-aggregate(pred$cpuev,by=list(pred$year),mean,na.rm=T)
cpueVS<-cpuev$x/mean(cpuev$x)
#calculate w0.8 and w0.5 series
cpue08 <- 0.8*cpueCS+(1-0.8)*cpueVS
cpue05 <- 0.5*cpueCS+(1-0.5)*cpueVS
#plot all series
plot(yearv,cpueUW,ylim=c(min(cpueUW,cpueCS,cpueVS,cpue08,cpue05,
cpue.r),max(cpueUW,cpueCS,cpueVS,cpue08,cpue05,cpue.r)),ylab="Standardised CPUE",
xlab="Year",type="l",lty=1,col=1)
lines(yearv,cpueCS,lty=2,col=2)
lines(yearv,cpueVS,lty=3,col=3)
lines(yearv,cpue08,lty=4,col=4)
lines(yearv,cpue05,lty=5,col=5)
```

```
lines(yearv,cpue.r,lty=6,col=6)
legend(1987,2.2, c("No Weighting","CS","VS","w0.8","w0.5","Raw"), lty=c(1:6) ,col=c(1:6))
```

#current model plus vessels as random effect

```
RTMPm5<-
lm(logCPUE~year+month+area+lat5+BETcpue5+YFTcpue5+month*area+year*lat5+year*area,
random=~1|vessel, data=RTMPdata)
#check residuals
plot(RTMPm5$fitted,RTMPm5$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm5$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
#use same prediction matrix as for random effects
#calculate predictions
pred$cpue<-exp(predict(RTMPm5,newdata=pred,level=1)+((0.6715)^2)/2)-log.add #replace
0.6715 by model's standard error
#calculate unweighted series
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpueUW<-cpue$x/mean(cpue$x)
#merge CS weightings and convert NAs to zeros
CS <- data.frame(
    month = aiCS$MO,
    area = aiCS$A,
    lat5 = aiCS$LAT,
    cs = aiCS$AI)
pred<-merge(pred,CS,all=T)
for(i in 1:length(pred$cs)){
if(is.na(pred$cs[i])) pred$cs[i]<-0}
#calculate CS series
pred$cpuec<-pred$cpue*pred$cs
cpuec<-aggregate(pred$cpuec,by=list(pred$year),mean)
cpueCS<-cpuec$x/mean(cpuec$x)
#merge VS weightings and convert NAs to zeros
aiVS <- subset(aiVS, X.YR > 1985)
aiVS <- subset(aiVS, LAT < 50)
VS <- data.frame(
    year = aiVS$X.YR,
    month = aiVS$MO,
    area = aiVS$A,
    lat5 = aiVS$LAT,
    vs = aiVS$AI)
VS <- subset(VS, year > 1985)
pred<-merge(pred,VS,all=T)
for(i in 1:length(pred$vs)){
```

```
if(is.na(pred$vs[i])) pred$vs[i]<-0}
#calculate VS series
pred$cpuev<-pred$cpue*pred$vs
cpuev<-aggregate(pred$cpuev,by=list(pred$year),mean,na.rm=T)
cpueVS<-cpuev$x/mean(cpuev$x)
#calculate w0.8 and w0.5 series
cpue08 <- 0.8*cpueCS+(1-0.8)*cpueVS
cpue05 <- 0.5*cpueCS+(1-0.5)*cpueVS
#plot all series
plot(yearv,cpueUW,ylim=c(min(cpueUW,cpueCS,cpueVS,cpue08,cpue05,
cpue.r),max(cpueUW,cpueCS,cpueVS,cpue08,cpue05,cpue.r)),ylab="Standardised CPUE",
xlab="Year",type="l",lty=1,col=1)
lines(yearv,cpueCS,lty=2,col=2)
lines(yearv,cpueVS,lty=3,col=3)
lines(yearv,cpue08,lty=4,col=4)
lines(yearv,cpue05,lty=5,col=5)
lines(yearv,cpue.r,lty=6,col=6)
legend(1987,2.2, c("No Weighting","CS","VS","w0.8","w0.5","Raw"), lty=c(1:6) ,col=c(1:6))
```

#To add vessel effects to the model without interactions, just delete the interaction terms from RTMPm4 and RTMPm5

#model with squares fixed term and vessel random effects

```
RTMPm6<-lme(logCPUE~year+month+BETcpue5+YFTcpue5+sq,random=~1|vessel,
data=RTMPdata)
plot(RTMPm6$fitted,RTMPm6$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm6$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
BETav<-aggregate(RTMPdata$BETcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$sq, RTMPdata$vessel), mean)
BET <- data.frame(
   year = BETav$Group.1,
   month = BETav$Group.2,
   sq = BETav$Group.3,
   vessel = BETav$Group.4,
   BETcpue5 = BETav$x)
YFTav<-aggregate(RTMPdata$YFTcpue5, by=list(RTMPdata$year,RTMPdata$month,
RTMPdata$sq, RTMPdata$vessel), mean)
YFT <- data.frame(
   year = YFTav$Group.1,
   month = YFTav$Group.2,
   sq = YFTav$Group.3,
   vessel = YFTav$Group.4,
```

```
    YFTcpue5 = YFTav$x)
pred_data<-
expand.grid(year=levels(RTMPdata$year),month=levels(RTMPdata$month),sq=levels(RTMPdata
$sq), vessel=levels(RTMPdata$vessel))
pred<-merge(pred_data,BET, all=T)
pred<-merge(pred,YFT, all=T)
for(i in 1:length(pred$BETcpue5)){
if(is.na(pred$BETcpue5[i])) pred$BETcpue5[i]<-0
if(is.na(pred$YFTcpue5[i])) pred$YFTcpue5[i]<-0}
pred$cpue<-exp(predict(RTMPm6,newdata=pred,level=1)+((0.66542)^2)/2)-log.add #replace
0.66542 by model's standard error
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpuesqfvr<-cpue$x/mean(cpue$x)
```

#model with vessel fixed term and square random effect

```
RTMPm7<-lm(logCPUE~year+month+BETcpue5+YFTcpue5+vessel,
random=~1|sq,data=RTMPdata)
plot(RTMPm7$fitted,RTMPm7$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm7$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
#use same prediction matrix as for random effects
pred$cpue<-exp(predict(RTMPm7,newdata=pred,level=1)+((0.6653)^2)/2)-log.add #replace
0.6653 by model's standard error
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpuevfsqr<-cpue$x/mean(cpue$x)
```

#model with vessel fixed term and square fixed term

```
RTMPm8<-lm(logCPUE~year+month+BETcpue5+YFTcpue5+vessel+sq, data=RTMPdata)
plot(RTMPm8$fitted,RTMPm8$resid, xlab="Fitted values", ylab="Residuals")
qqnorm(RTMPm8$resid, main=" ")
abline(a=0,b=1)
#constructing the standardised series
#use same prediction matrix as for random effects
pred$cpue<-exp(predict(RTMPm8,newdata=pred)+((0.6653)^2)/2)-log.add #replace 0.6653 by
model's standard error
cpue<-aggregate(pred$cpue,by=list(pred$year),mean)
cpuevfsqr<-cpue$x/mean(cpue$x)
```