FLsp - a Surplus Production model in FLR

Finlay Scott <finlay.scott@cefas.co.uk
Cefas, Lowestoft, UK

June 2011

1 Introduction

This package implements the surplus production described in Polacheck REF. At the moment only the model including observation error is implemented. The Pella Tomlinson shape is used (which defaults to Schaeffer) Tested against the three data sets in the paper. Accurate gradients and hessian are returned using automatic differentiation (implemented using ADOLC REF)

2 The Model

Following the Polacheck paper Assumptions: Observation error Only r and k are estimated sigma and q are approximated as described in the paper The inital biomass is k

The general equation for the biomass through time is:

$$B_{y+1} = B_y + g(B_y) - C_y \tag{1}$$

where B is the stock biomass at the start of year y, C is the catch during the year and g is surplus production as a function of biomass.

Here we implement the Pella-Tomlinson form of surplus production:

$$g(B) = \frac{r}{p}B(1 - (B/K)^p) \tag{2}$$

where r is the intrinsic growth rate parameter and K is the average biomass level prior to exploitation. By default, p is set to 1 making the surplus prodution formulation the same as a Schaefer model.

The biomass is related to an index of abundance:

$$I_y = qB_y \tag{3}$$

Where I is an index of relative abundance in year y and q is the catchability coefficient.

Here we assume that errors are introduced through observation. The population dynamics are assumed to be deterministic and all of the error occurs in the relationship between stock biomass and the index of abundance. It is assumed that the error is multiplicative and log-normal with a constant coefficient of variation. The estimates of the model parameters are (B_0, r, K and q) are obtained by maximising the likeilhood function:

$$L = \prod exp(\hat{v}_y^2/(2\hat{\sigma}_v^2))/(\sqrt{2\pi}\hat{\sigma}_v) \tag{4}$$

where the product is over all years for which CPUE data are available and:

$$\hat{v}_y = log(C/E)_y - log(\hat{C/E})_y \tag{5}$$

$$\hat{\sigma}_v^2 = \sum \hat{v}_y^2/n \tag{6}$$

where n is the number of data points.

The value of q which maximises the likelihood is given by:

$$\hat{q} = exp\left(\frac{1}{n}\sum_y log(I_y/\hat{B}_y)\right) \tag{7}$$

Following Polacheck et al B_0 is set to K. This means that only two parameters need to be estimated: r and K. In $FLsp$ the estimation is performed using the *DEoptim* package REF.

3 The FLsp class

The $FLsp$ class extends the $FLModel$ class by including slots to store the catch and index time series. Catch is represented as an $FLQuant$ and index is represented as an $FLQuants$ object. This allows multiple indices to be used (not yet implemented).

To estimate the parameters r and K, an $FLsp$ object must be created with catch and index data. The method $fitsp()$ is then called. Once the object has been fitted, the biomass trajectory and other variables of interest (e.g. $sigma^2$ and $\hat{q}$ can be calculated).

4 Creating and fitting FLsp objects

Here we show how to create and fit a surplus production model using $FLsp$. The data set is New Zealand Rock Lobster, taken from Polcheck REF.

```
> # Load the library
> library(FLsp)
> # Load the New Zealand Rock Lobster data set
> data(nzrl)
> # This is a dataframe with year, catch and cpue
> # Take a look at the top of it
> head(nzrl)
> # Make FLQuant objects of the catch and cpue series
> catch <- FLQuant(nzrl$catch, dimnames=list(year=nzrl$year))
> index <- FLQuant(nzrl$cpue, dimnames=list(year=nzrl$year))
> # Create the FLsp object
> nzrl <- FLsp(catch=catch,index=index)
```

After creating our object we are ready to fit the parameters.

```
> nzrl <- fitsp(nzrl)
```

The published values for this data set are: $r = 0.0659$, $K = 129000$, $\hat{q} = 2.461\text{e-}5$, $\sigma = 0.207$, $B_{current} = 21150$. These can be compared to our results by interrogating the $FLsp$ object.

```
> # Look at the fitted parameters
> params(nzrl)

An object of class "FLPar"
params
         r          k
5.1428e-02 1.4064e+05
units:  NA

> # qhat
> qhat(nzrl)

An object of class "FLQuant"
, , unit = unique, season = all, area = unique


     year
quant 1
  all 2.237e-05

units:  NA

> # sigma2
> sqrt(sigma2(nzrl))

An object of class "FLQuant"
, , unit = unique, season = all, area = unique


     year
quant 1
  all 0.20698

units:  NA

> # returns the full biomass timeseries
> biomass(nzrl)

An object of class "FLQuant"
, , unit = unique, season = all, area = unique


     year
quant 1945   1946   1947   1948   1949   1950   1951   1952   1953   1954   1955   1956   1957   1958
  all 140640 139831 139018 138182 136946 135259 132853 130397 127562 124012 119225 115149 109676 1058
     year
quant 1962   1963   1964   1965   1966   1967   1968   1969   1970   1971   1972   1973   1974   1975
  all  95382  92377  89453  86531  83259  79711  76705  73523  70542  67651  64978  63281  61287  594
     year
quant 1979   1980   1981   1982   1983   1984   1985   1986   1987   1988   1989   1990
  all  53453  51107  48590  46168  43432  40591  37165  33715  30376  27101  25098  22841

units:  NA
```

	New Zealand Rock Lobster		South Atlantic Albacore		Northern Namibian Hake	
Measure	FLsp	Published	FLsp	Published	FLsp	Published
r	r[1]	0.0659	r[2]	0.328	r[3]	0.379
K ('000 t)	k[1]	129.0	k[2]	239.6	k[3]	2772.6
σ	sigma[1]	0.207	sigma[2]	0.111	sigma[3]	0.124
$\hat{q}$	qhat[1]	2.461e-05	qhat[2]	0.2671	qhat[3]	4.360e-04
$B_{current}$	bc[1]/1000	21.15	bc[2]	75.51	bc[3]	1646.3
MSY	0	2133.74	0	19.65	0	263.2

Table 1: Comparing the published results with those from *FLsp* for three data sets.

It can be seen that there is good agreement between the published results and those generated with *FLsp*. The differences are likely caused by the precision of the printed data set in the Polcheck paper (REF) and the fitting method used.

5 Testing FLsp against the other data sets

The results fitted with *FLsp* are in good agreement with the published results.

6 Plotting results

There is no generic plot for FLsp at the moment. However, it is possible to look at the fitted index and residuals using relatively simple code. For example, to plot the indices with the fitted indices you can use (see Figure 1):

```
> fitted <- cbind(as.data.frame(nzrl@fitted_index),type="fitted")
> index <- cbind(as.data.frame(nzrl@index),type="index")
> index <- rbind(index,fitted)
> print(xyplot(data ~ year | qname, group=type, data=index, type="b",auto.key=TRUE))
```

To look at the residuals and put a loess function through them use (see Figure 2):

```
> residuals <- as.data.frame(nzrl@residuals_index)
> print(xyplot(data ~ year | qname, data = residuals, panel = function(x, y) {
+     panel.xyplot(x, y)
+     panel.loess(x, y, span = 1)
+ }))
```

7 Profiling the fit

You can explore how good the fit is by looking at the likelihood profile. This is easily done by using the *profile*() method (see Figure 3).

Notice that the profile plot has a banana shaped flat section which contains the optimum solution. This is because the parameters r and K are correlated, making them difficult to estimate unless there is sufficient information in the data. The profile plot also includes the gradient of the log likelihood as r and K change (keeping K and r fixed at the estimated value found by the optimiser respectively). The dashed line is at a gradient of 0. If the fitting has worked, the gradient should be 0 at the estimated parameter values. It is just another simple way to check that the results of from the fitting process are sensible.

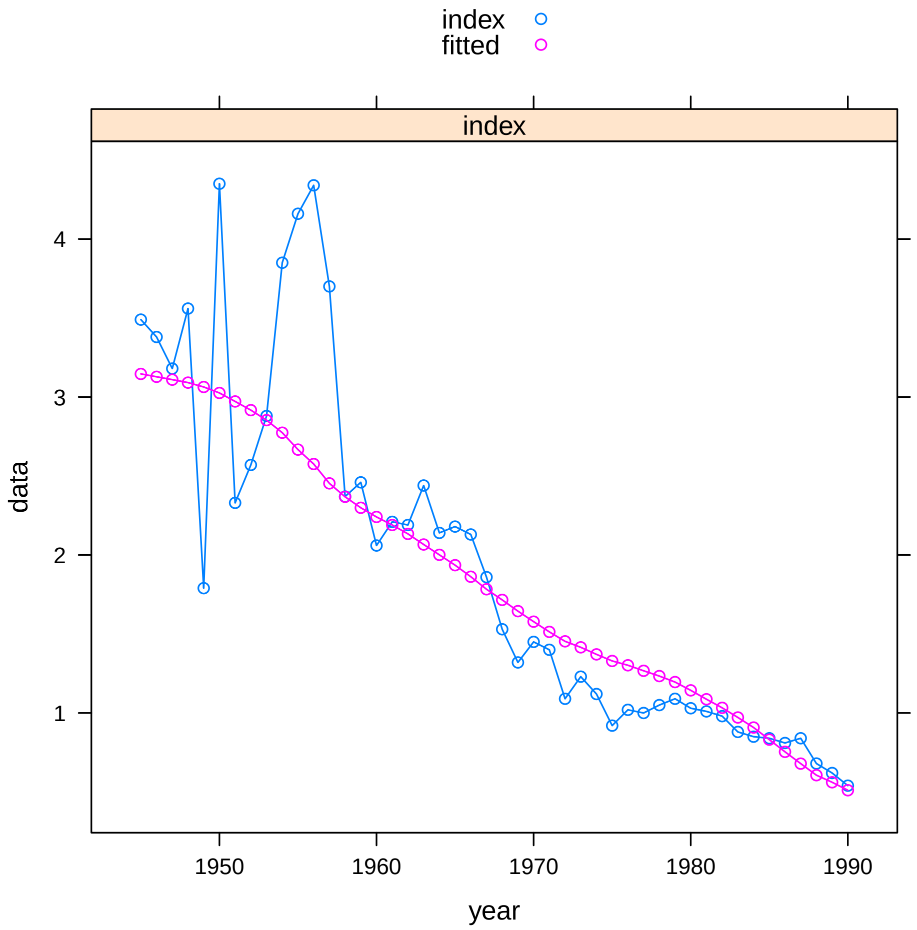

Figure 1: Indices and fitted indices for New Zealand rock lobster

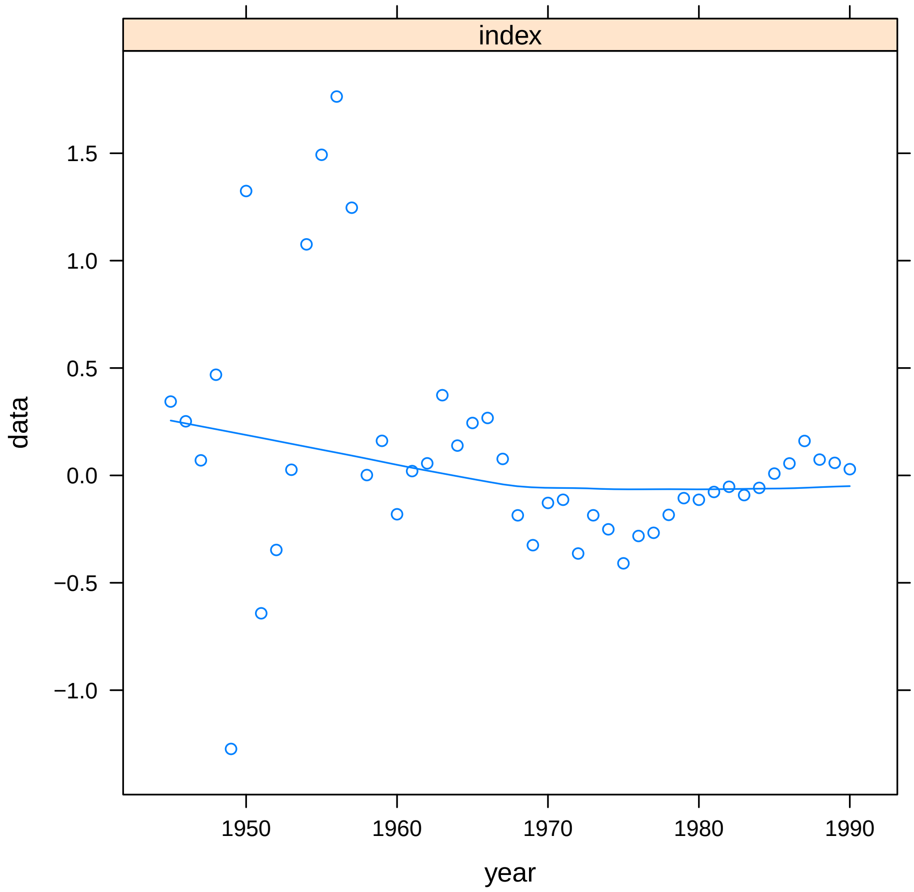

Figure 2: Residuals for New Zealand rock lobster

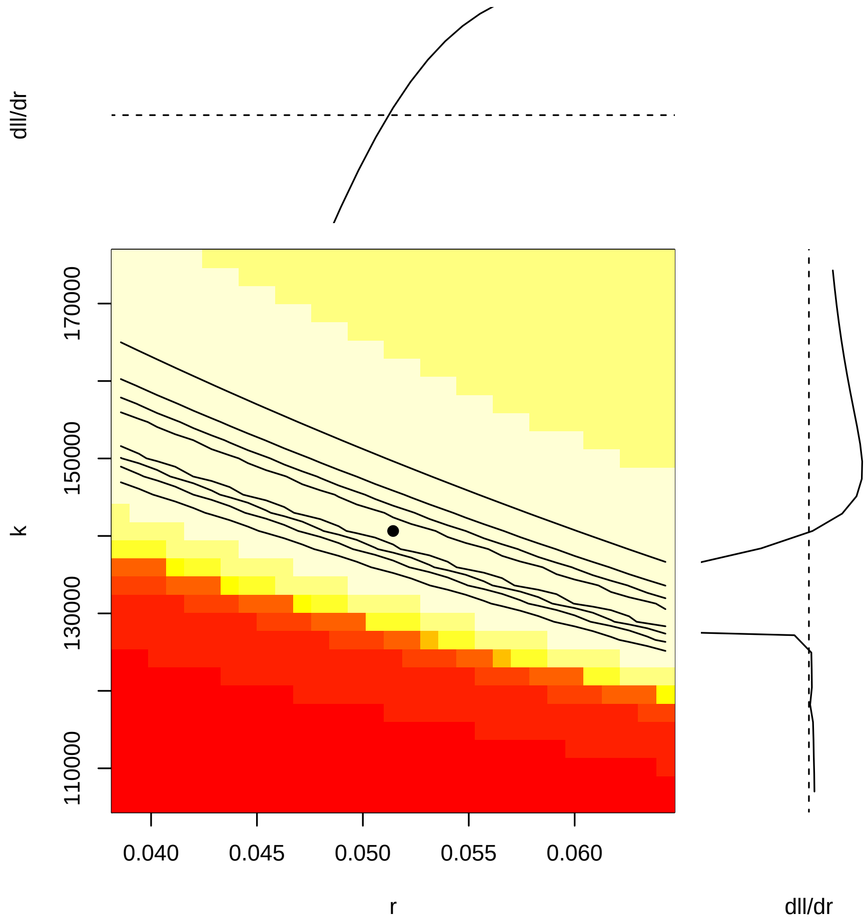

Figure 3: Profile plot for New Zealand rock lobster

```
> profile(nzrl, maxsteps = 31, range = 0.25)
```

8 Uncertainty