Summer School in Quantitative Fisheries Stock Assessment

Day 3: An introduction to Virtual Population Analysis (VPA) & Extended Survivors Analysis (XSA)

Cóilín Minto, Chato Osio, Alessandro Orio, Alessandro Ligas, Alessandro Mannini, Danai Mantopoulou Palouka

Food and Agriculture Organization of the United Nations
General Fisheries Commission for the Mediterranean
Commission générale des pêches pour la Méditerranée
European Commission

Outline

Outline

Aristotle's ruminations

.., almost every serious intellectual advance has had to begin with an attack on some Aristotelian doctrine;

Bertrand Russell, *History of Western Philosophy*

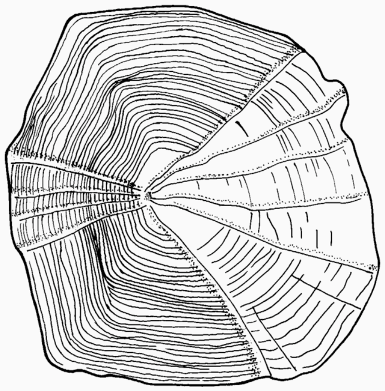
Aristotle $\sim$ 350 B.C.

Murex spp.

Both the murex and the ceryx are long lived. The murex lives for about six years, and the yearly increase is indicated by a distinct interval in the spiral convolution of the shell.

History of Animals V

van Leeuwenhoek $\sim$1800

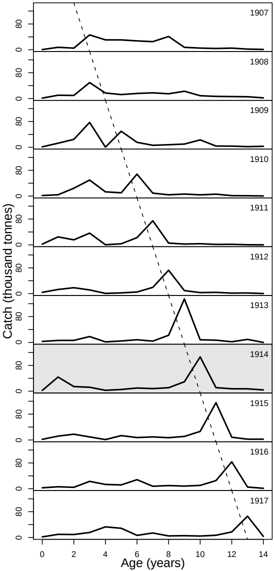


- Hoffbauer (1899)
- Thomson (1900-1903)
- Johnston (1904)

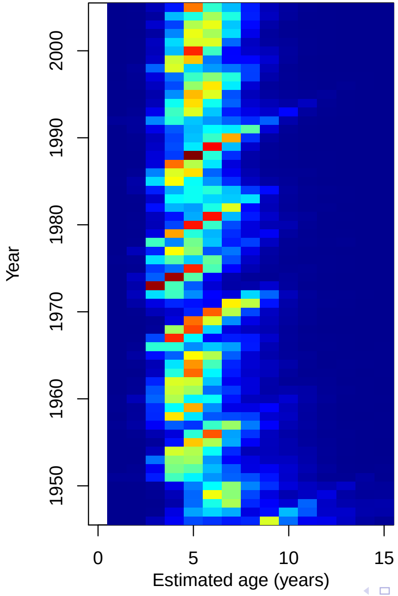
Hjort 1914

Applied the methods of human *vital statistics* to herring

- Cohort progression from samples
- Variability in cohort strength
- Critical period & aberrant drift hypotheses



Progression of a multiple cohorts



Age structure

Numbers at age matrix

$$\begin{array}{c|ccc} & & \text{Age} \longrightarrow & \\ \hline \text{Time} \downarrow & N_{a1,t1} & N_{a2,t1} & N_{a3,t1} \\ & N_{a1,t2} & N_{a2,t2} & N_{a3,t2} \\ & N_{a1,t3} & N_{a2,t3} & N_{a3,t3} \\ & N_{a1,t4} & N_{a2,t4} & N_{a3,t4} \\ & N_{a1,t5} & N_{a2,t5} & N_{a3,t5} \end{array}$$

Cohort

Cohort progression

$$\begin{array}{c|ccc} & \multicolumn{3}{c}{\text{Age} \longrightarrow} \\ \hline \text{Time} \downarrow & N_{a1,t1} & N_{a2,t1} & N_{a3,t1} \\ & N_{a1,t2} & N_{a2,t2} & N_{a3,t2} \\ & N_{a1,t3} & N_{a2,t3} & N_{a3,t3} \\ & N_{a1,t4} & N_{a2,t4} & N_{a3,t4} \\ & N_{a1,t5} & N_{a2,t5} & N_{a3,t5} \end{array}$$

Cohort

Cohort progression

$$
\begin{array}{cccc}
 & & \text{Age} \longrightarrow & \\
\text{Time} \downarrow & N_{a1,t1} & N_{a2,t1} & N_{a3,t1} \\
 & N_{a1,t2} & N_{a2,t2} & N_{a3,t2} \\
 & N_{a1,t3} & N_{a2,t3} & N_{a3,t3} \\
 & N_{a1,t4} & N_{a2,t4} & N_{a3,t4} \\
 & N_{a1,t5} & N_{a2,t5} & N_{a3,t5}
\end{array}
$$

Cohort change

Fish live and die in continuous time. How can we describe the abundance of a cohort over time?

Cohort change in time

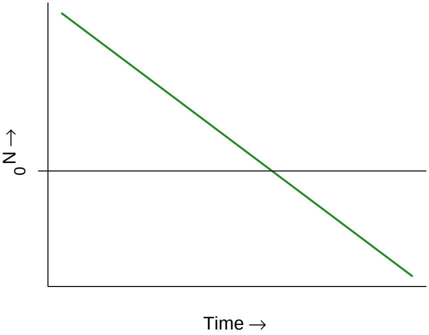
First guess: Cohort declines at a constant rate

$$\frac{dN}{dt} = -Z$$

with initial cohort size $N(0)$ the solution is

$$\int \frac{dN}{dt} dt = \int -Z dt$$

$$N(t) = N(0) - Zt$$



Can't be correct as abundance goes negative!

Cohort change in time

Second guess: Cohort declines at a rate proportional to abundance

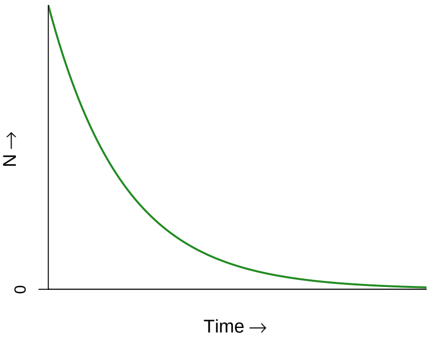
$$\frac{dN}{dt} = -ZN$$

$$\frac{dN}{N} = -Zdt$$

$$\int \frac{1}{N} dN = \int -Zdt$$

$$\ln N = C - Zt$$

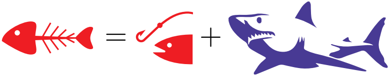
$$N(t) = e^C e^{-Zt} = \boxed{N(0)e^{-Zt}}$$



Just an exponential decay but central to what follows!

Fishing and natural mortality

Let the rate of decline depend on both fishing and natural mortality



$$Z = F + M$$

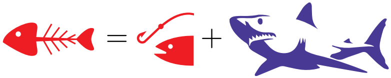
Say for a yearly time-step, $t = 1$, we have

$$N_{t+1} = N_t e^{-Z} = N_t e^{-(F+M)}$$

that is the number surviving.
For $F = 0.3$, $M = 0.1$ and $N_t = 100$, how many survive to next year?

Fishing and natural mortality

Let the rate of decline depend on both fishing and natural mortality



$$Z = F + M$$

Say for a yearly time-step, $t = 1$, we have

$$N_{t+1} = N_t e^{-Z} = N_t e^{-(F+M)}$$

that is the number surviving.
For $F = 0.3$, $M = 0.1$ and $N_t = 100$, how many survive to next year?

$$N_{t+1} = 100e^{-(0.3+0.1)} = 67$$

How many died?

Fishing and natural mortality

How many died?

$$\begin{aligned} 33 &= 100\left(1 - e^{-(0.3+0.1)}\right) \\ &= N_t\left(1 - e^{-(F+M)}\right) \end{aligned}$$

How many died because of fishing (catch)?

Fishing and natural mortality

How many died?

$$\begin{aligned} 33 &= 100\left(1 - e^{-(0.3+0.1)}\right) \\ &= N_t\left(1 - e^{-(F+M)}\right) \end{aligned}$$

How many died because of fishing (catch)?

$$\begin{aligned} 25 &= \frac{0.3}{0.3+0.1} \times 33 \\ C_t &= \frac{F}{F+M}\left(1 - e^{-(F+M)}\right) N_t \end{aligned}$$

which is the famous Baranov catch equation!

Outline

Goal of VPA

Re-construct past cohort abundance based on past catches: but how?

VPA derivation

$$
\begin{aligned}
N_{t+1} &= N_t e^{-(F+M)} \\
\ln\left(\frac{N_{t+1}}{N_t}\right) + M &= -F \\
F &= -\ln\left(\frac{N_{t+1}}{N_t}\right) - M
\end{aligned}
$$

The Baranov equation gives can be re-arranged

$$
\begin{aligned}
C_t &= \frac{F}{F+M}\left(1 - e^{-(F+M)}\right) N_t \\
&= \frac{F}{F+M}(N_t - N_{t+1})
\end{aligned}
$$

VPA derivation

Substitute in the expression for F

$$C_t = \frac{-\ln\left(\frac{N_{t+1}}{N_t}\right) - M}{-\ln\left(\frac{N_{t+1}}{N_t}\right) - M + M}(N_t - N_{t+1})$$

$$C_t = \frac{-\ln\left(\frac{N_{t+1}}{N_t}\right) - M}{-\ln\left(\frac{N_{t+1}}{N_t}\right)}(N_t - N_{t+1})$$

$$C_t = \left(1 - \frac{M}{\ln(N_t) - \ln(N_{t+1})}\right)(N_t - N_{t+1})$$

Use this VPA equation[1] to solve for N_t given C_t, M and N_{t+1}

[1]See also Pope's method

VPA notes

Note

- *Virtual* as the population is inferred, not observed
- Popualtion is re-constructed from assumptions on deaths

Example: Hake GSA 9–11

Catch-at-age for 2008 cohort

Age	Year	Catch
0	2008	65121
1	2009	12360
2	2010	1452
3	2011	285
4	2012	54
5	2013	12
6	2014	20

Example: Hake GSA 9–11

Catch-at-age for 2008 cohort

Age	Year	Catch
0	2008	65121
1	2009	12360
2	2010	1452
3	2011	285
4	2012	54
5	2013	12
6	2014	20

Need to guess N_t for the oldest age group via a guess on the terminal F

Example: Hake GSA 9–11

Assume $M = 0.3$ and pick a terminal $F = 1$ to first get terminal N

$$C_{6,2014} = \frac{F_{6,2014}}{F_{6,2014} + M}\left(1 - e^{-(F_{6,2014}+M)}\right) N_{6,2014}$$

$$20 = \frac{1}{1.3}\left(1 - e^{-(1.3)}\right) N_{6,2014}$$

$$N_{6,2014} = 35.7$$

Age	Year	Catch	F	N
0	2008	65121		
1	2009	12360		
2	2010	1452		
3	2011	285		
4	2012	54		
5	2013	12		
6	2014	20	1.00	35.7

Example: Hake GSA 9–11

Now use VPA equation to solve for previous N values, given catches[2]

$$C_t = \left(1 - \frac{M}{\ln(N_t) - \ln(N_{t+1})}\right)(N_t - N_{t+1})$$

Age	Year	Catch	F	N
0	2008	65121		95648.9
1	2009	12360		16898.9
2	2010	1452		2332.3
3	2011	285		518.4
4	2012	54		145.4
5	2013	12		62.0
6	2014	20	1.00	35.7

[2] Can use `uniroot` solver in R to do this

Example: Hake GSA 9–11

Fill in the fishing mortalities

$$F = -\ln\left(\frac{N_{t+1}}{N_t}\right) - M$$

Age	Year	Catch	F	N
0	2008	65121	1.43	95648.9
1	2009	12360	1.68	16898.9
2	2010	1452	1.20	2332.3
3	2011	285	0.97	518.4
4	2012	54	0.55	145.4
5	2013	12	0.25	62.0
6	2014	20	1.00	35.7

For all cohorts

Catch-at-age matrix

	Age						
Year	0	1	2	3	4	5	6
2006	82424	14603	2299	299	103	29	0
2007	62020	13976	1314	209	55	12	3
2008	65121	12123	870	251	81	38	8
2009	91660	12360	812	172	86	25	9
2010	42494	12523	1452	225	68	37	12
2011	66748	12820	1204	285	101	38	8
2012	29969	10271	880	209	54	14	2
2013	26054	12899	744	134	53	12	4
2014	42564	10589	1331	187	40	24	20

For all cohorts

Catch-at-age matrix

	Age						
Year	0	1	2	3	4	5	6
2006	82424	14603	2299	299	103	29	0
2007	62020	13976	1314	209	55	12	3
2008	65121	12123	870	251	81	38	8
2009		12360	812	172	86	25	9
2010			1452	225	68	37	12
2011				285	101	38	8
2012					54	14	2
2013						12	4
2014							20

Terminal F assumption

	Age						
Year	0	1	2	3	4	5	6
2006							1
2007							1
2008							1
2009							1
2010							1
2011							1
2012							1
2013							1
2014							1

VPA N values

Apply VPA equation to each cohort back in time to get population numbers at age ($\times 1,000$)

	Age						
Year	0	1	2	3	4	5	6
2006	190441	22118	3454	570	164	40	0
2007	154442	19932	2416	495	154	34	5
2008	166070	17813	1658	558	175	65	14
2009		19813	1703	408	185	59	16
2010			2551	479	144	62	22
2011				548	150	47	14
2012					152	26	4
2013						65	7
2014							36

Terminal F assumption

Need to also make an assumption for the F across all ages in the final year, e.g.

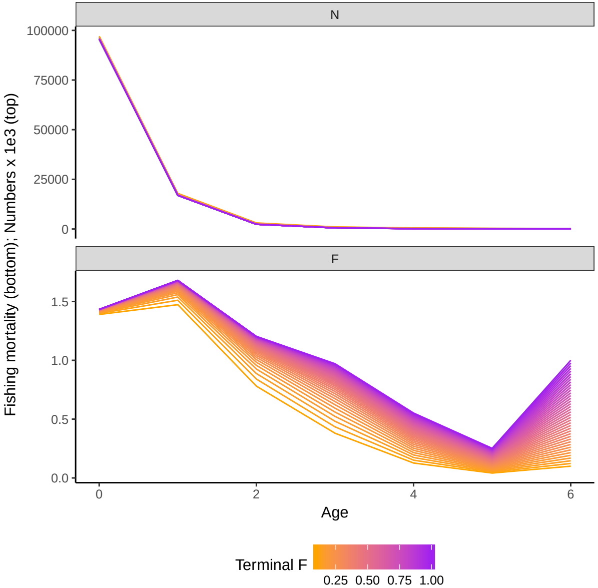
	Age						
Year	0	1	2	3	4	5	6
2006							1
2007							1
2008							1
2009							1
2010							1
2011							1
2012							1
2013							1
2014	1	1	1	1	1	1	1

VPA N values

Apply VPA equation to each cohort back in time to get population numbers at age ($\times 1,000$)

	Age						
Year	0	1	2	3	4	5	6
2006	190439	22116	3452	570	164	40	0
2007	154439	19931	2415	494	153	33	5
2008	166055	17812	1658	558	174	65	14
2009	202354	19808	1703	408	185	58	16
2010	125110	18912	2549	479	144	61	21
2011	151388	18061	2058	547	150	47	14
2012	114385	14953	1429	416	151	26	4
2013	110899	20344	1406	259	121	64	7
2014	104876	21066	2488	342	72	43	36

Sensitivity to terminal F assumption



Vanilla VPA in a nutshell

- Accounting procedure - given N_{t+1}, C_t and M solve the VPA equation for N_t, repeat
- How to get the terminal N? Assume a terminal F.
- Results sensitive to assumption on terminal F (often the cohorts we're <u>most</u> interested in!)
- Need a better solution

Outline

Instead of assume, let's tune!

Tuned VPA use additional information (other than catches) to infer the terminal F values

Can use:

- Survey index
- Effort data

Laurec-Shepherd tuning method

Ad-hoc tuning method to derive terminal F in the final year

- Analysis of (fleet disaggregated) catchability at age

Adapted Laurec-Shepherd tuning method

- Make assumption on terminal F, say $F = 1$

Adapted Laurec-Shepherd tuning method

- Make assumption on terminal F, say $F = 1$
- Run the VPA to get $N_{a,t}$
- Calculate $F_{a,t}$

Adapted Laurec-Shepherd tuning method

- Make assumption on terminal F, say $F = 1$
- Run the VPA to get $N_{a,t}$
- Calculate $F_{a,t}$
- Using effort (potentially by fleet), calculate the average catchability $q_{a,t}$ over a reference period (not too long) excluding the terminal values

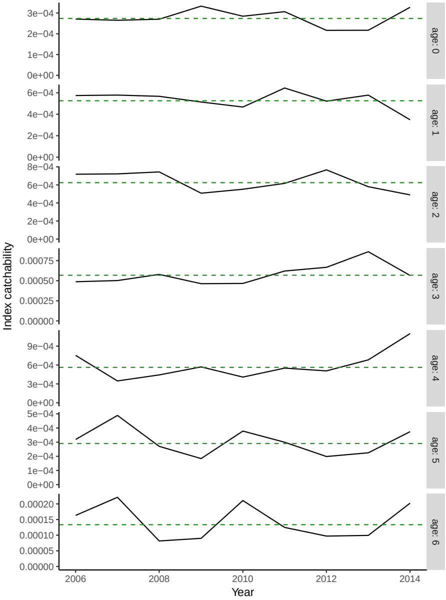
Adapted Laurec-Shepherd tuning method

- Make assumption on terminal F, say $F = 1$
- Run the VPA to get $N_{a,t}$
- Calculate $F_{a,t}$
- Using effort (potentially by fleet), calculate the average catchability $q_{a,t}$ over a reference period (not too long) excluding the terminal values
- Multiply catchability by the effort in the terminal year to get the F in the terminal year (potentially by fleet)

Adapted Laurec-Shepherd tuning method

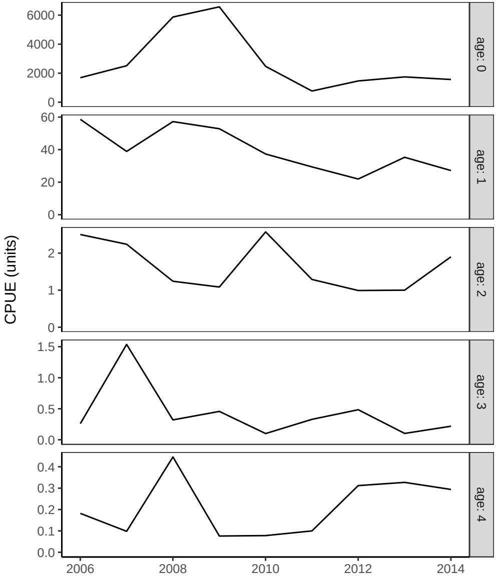
- Make assumption on terminal F, say $F = 1$
- Run the VPA to get $N_{a,t}$
- Calculate $F_{a,t}$
- Using effort (potentially by fleet), calculate the average catchability $q_{a,t}$ over a reference period (not too long) excluding the terminal values
- Multiply catchability by the effort in the terminal year to get the F in the terminal year (potentially by fleet)
- Repeat with new terminal F values until they converge

Catchabilities

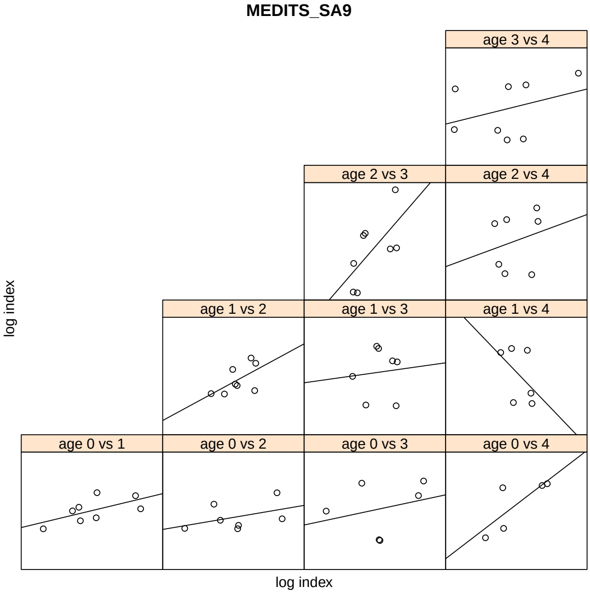


Outline

A closer look at the index: time series



A closer look at the index: cohort progression

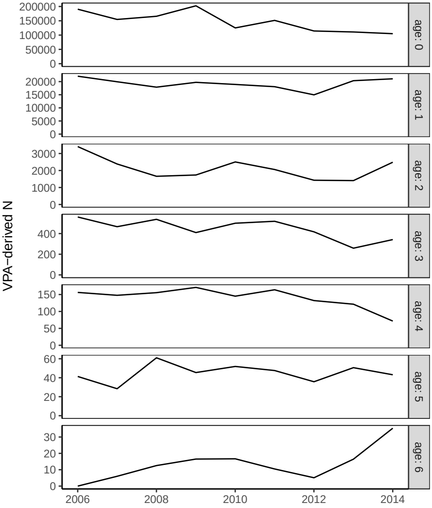


Extended survivors analysis (XSA)

XSA will use the survey index to update the numbers from the VPA . . . how?

XSA

Start with an initial VPA run to get number under terminal assumption



CPUE - N relationship

Yesterday we had (in numbers here)

$$C = qEN$$
$$C/E = CPUE = qN$$

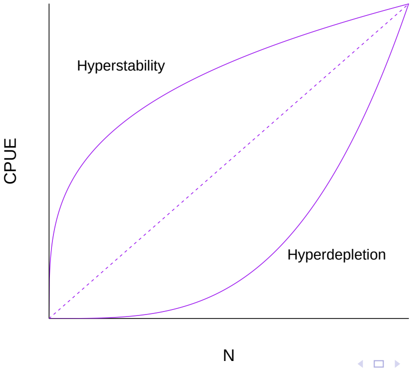
that is the CPUE is proportional to abundance - this typically does not hold

CPUE - N relationship

Yesterday we had (in numbers here)

$$C = qEN$$
$$C/E = CPUE = qN$$

that is the CPUE is proportional to abundance - this typically does not hold



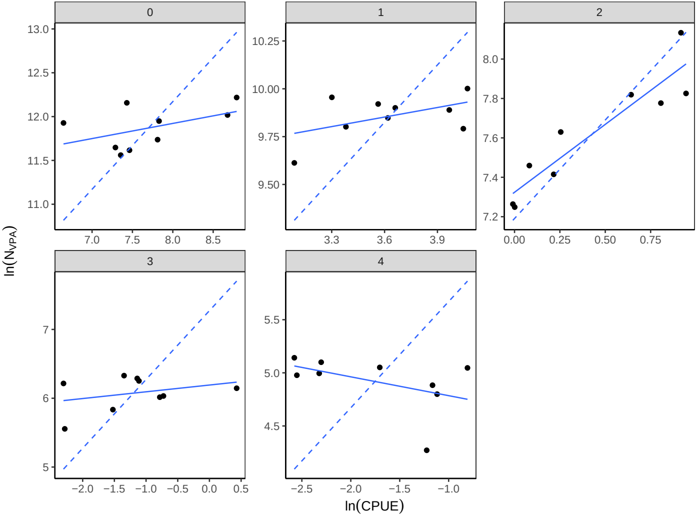
CPUE - N relationship

Better to use a non-linear relationship

$$\begin{aligned} CPUE &= qN^{\beta} \\ \ln(CPUE) &= \ln(q) + \beta \ln(N) \\ \ln(N) &= \frac{\ln(CPUE)}{\beta} - \frac{\ln(q)}{\beta} \end{aligned}$$

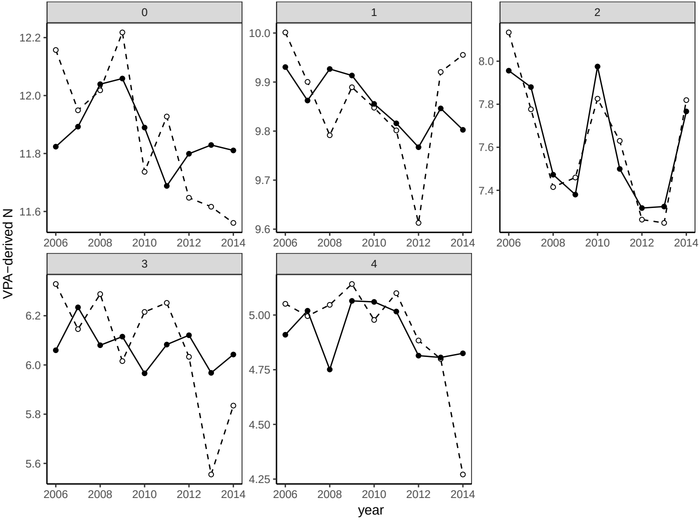
Can estimate $\{q, \beta\}$ parameters in a regression

ln(*CPUE*) - ln(*N*) relationship



Correct the numbers

Predict from the regression



Correct the numbers

- Summing down the cohort, the updated estimates of N are then used to calculate the number of **survivors** in the oldest age group
- Now repeat the VPA
- Re-correct the abundances
- Repeat until convergence (small change in the fishing mortalities between iterations for example)

Additional features

- Shrinkage (constrain differences between years in F or N; form of regularisation)
- Taper years to more relevant period to estimate catchability
- Much more in the tutorials!

Note

Diagnostics are again an essential first step, e.g., see catchability residuals from the linear regression

Outline

Preliminaries

VPA

Tuned VPA

XSA

Summary

Summary

- Age-based methods can allow us to track cohorts

Summary

- Age-based methods can allow us to track cohorts
- Catch equation can be derived from a linear rate of change in N

Summary

- Age-based methods can allow us to track cohorts
- Catch equation can be derived from a linear rate of change in N
- Methods developed for re-constructing population size based on deaths (catch and nat. mortality)

Summary

- Age-based methods can allow us to track cohorts
- Catch equation can be derived from a linear rate of change in N
- Methods developed for re-constructing population size based on deaths (catch and nat. mortality)
- VPA makes an assumption on terminal F

Summary

- Age-based methods can allow us to track cohorts
- Catch equation can be derived from a linear rate of change in N
- Methods developed for re-constructing population size based on deaths (catch and nat. mortality)
- VPA makes an assumption on terminal F
- Tuned-VPA makes use of index/CPUE data to get past terminal assumptions

Summary

- Age-based methods can allow us to track cohorts
- Catch equation can be derived from a linear rate of change in N
- Methods developed for re-constructing population size based on deaths (catch and nat. mortality)
- VPA makes an assumption on terminal F
- Tuned-VPA makes use of index/CPUE data to get past terminal assumptions
- XSA algorithm more sophisticated method

References

Darby, C. D., and Flatman, S. 1994. Virtual population analysis: Version 3.1 (Windows/DOS) User Guide. MAFF Directorate of Fisheries Research IT Report 1. 85 pp.

Hilborn, R., and Walters, C. J. (2013). Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Springer Science & Business Media. [Ch. 8]

Lassen, H.; Medley, P. Virtual population analysis. A practical manual for stock assessment. FAO Fisheries Technical Paper. No. 400. Rome, FAO. 2001. 129p.

Shepherd, J. G. 1999. Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. ICES Journal of Marine Science, 56: 584–591.