SCIE3250 – Logbook

**Research question:** How often does the Hockey-Stick model out-perform the Beverton-Holt model when predicting the biomass of fish using data from the RAM stock assessment database?

**Prediction:** Hockey-Stick will ‘fit better’ 25-30% of the time.

**Significance:** The Hockey-Stick model suggests a more conservative harvest rate for growth rates less that 2.618. This means that when the Beverton Model is incorrectly used there is a risk of overexploitation.

Week 1

Meeting: Friday 25th Nov

|  |  |  |
| --- | --- | --- |
| Task | Tasks assigned. | Notes |
| 1.1 | Create biomass time series plots for all species from ‘Bio.csv’ (available on RAM stock assessment database). | RAM DB v4.495 (latest version) downloaded  Code written to generate graphs from ‘Bio.csv’ (SCIE3250\_task\_1.1\_code.R)  Graphs compiled into single PDF (SCIE3250\_task\_1.1\_biomass\_time\_series\_plots.pdf)  There were 999 different species listed in ‘Bio.csv’. Of these 479 included data points (520 columns contained only ‘NA’). |
| 1.2 | Classify each plot as either ‘deterministic or ‘non-deterministic and discard all the ‘deterministic’ plots. | The column number for each ‘deterministic looking’ plot was recorded.  See below example of ‘deterministic’ and ‘non-deterministic’ plots:  Deterministic Non-deterministic  If I was unsure, then it was classed as deterministic.  Of the 479 plots, 160 were classed as non-deterministic. |
| 1.3 | Read literature by Costello regarding modelling methods using the ram stock assessment database with the hope to determine the most meaningful way to plot R\_{t} vs R\_{t+1}. | It is now obvious that this study is an over-simplification of real life as all environmental and interaction factors are being discarded. Regardless, a lot of models are based off, or are variations of the Beverton-Holt (BH) model and the Hockey-Stick (HS) is rarely used. |
| 1.4 | Determine if ‘Bio.csv’ (total biomass) includes the ‘catch’. |  |
| 1.5 | Plot R\_{t} vs R\_{t+1}. | To this I think I will have to plot:  Bio\_{t}-Mcatch\_{t} vs Bio\_{t+1}  Bio – TBbest time series – Total biomass (metric tons)  Mcatch – TCbest time series – Total catch (metric tons)  These two data sets contain entries in different cells so I will sort through and find which cells have entries from both fields. |

Next meeting: Mon 5th Dec

Questions for next meeting:

Some questions, assumptions, and things to note:

The data provided in the database is not independently reviewed for its ‘correctness’.

Total biomass:

* What technique is typically used to estimate biomass?

Total catch:

* By-catch could be included
* Under-reporting

RAM Stock Assessment Database:

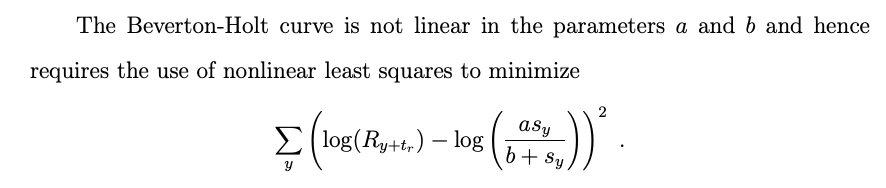
* Is biased towards large commercially important fisheries.
* Captures 35% of world catch.
* Biased towards wealthy nations with management systems in place.
* (<https://academic.oup.com/icesjms/article/71/5/1040/648075>)

What do we have to assume about the data?

* IID?
* Following what distribution?

Parameter estimation and likelihood functions.

* We want, L(r,k | Y)
* <https://www.jstor.org/stable/3100085?seq=3#metadata_info_tab_contents>



When fitting models I had to remove 18 more data sets (not exactly sure for the reason why yet) meaning only 142 data sets were plotted

Week 2

Meeting: Tuesday 6th Dec

|  |  |  |
| --- | --- | --- |
| Task | Tasks assigned. | Notes |
| 2.1 | Read Costello – find out how many data sets he used and more about his methods. |  |
| 2.2 | Classify deterministic/non-deterministic data sets from R\_{t} vs R\_{t+1} plots. |  |
| 2.3 | Estimate parameters using optim. |  |
| 2.4 | Compare optim to my other numeric techniques. |  |
| 2.5 | Find sources for literature review. | Costello, |