Scaling factors for multi-region stock assessments:

Indian Ocean tropical tunas

Simon Hoyle and Adam Langley

Indian Ocean Tuna Commission
Commission des Thons de l'Océan Indien
iotc ctoi

Introduction

- In stock assessments with multiple regions, it is important to determine relative abundances among regions.
  - Can estimate using CPUE data.
- ‘Regional scaling’ developed and used in WCPO tropical tuna assessments since 2005 (Langley et al 2005).
- Another version used in IO YFT assessments since 2008.
- Methods not published in peer reviewed literature.
- Need to compare versions and explore improvements.

Introduction 2

- Relative catch rates among regions are a proxy for density.
- Abundance = density x area, so we sum cell abundances by region.
  - $\kappa_r = \sum_{\forall i \; in \; region \; r} density_i \times area_i$
- Regional scaling parameters
  - $\nu_r = \frac{\kappa_r}{\max(\kappa_{r,\forall r})} \sim$ relative abundance by region, during the scaling period.
- In the stock assessment, the scaling parameters v_r are used to adjust the CPUE series so that mean($CPUE_{scaling\ period,\ r}$) = ν_r

Introduction 3: the choices required

- Scaling period requirements
  - Spatial coverage should be good
  - CPUE should represent abundance (esp. wrt target change)
  - Trends by region should be consistent
- Density estimation
  - Standardization versus raw CPUE
  - Select modelling approach and covariates (using AIC)
- Area estimation
  - Assume equal areas (old method) vs. estimate ocean areas (proposed method)

Methods 1: data selection

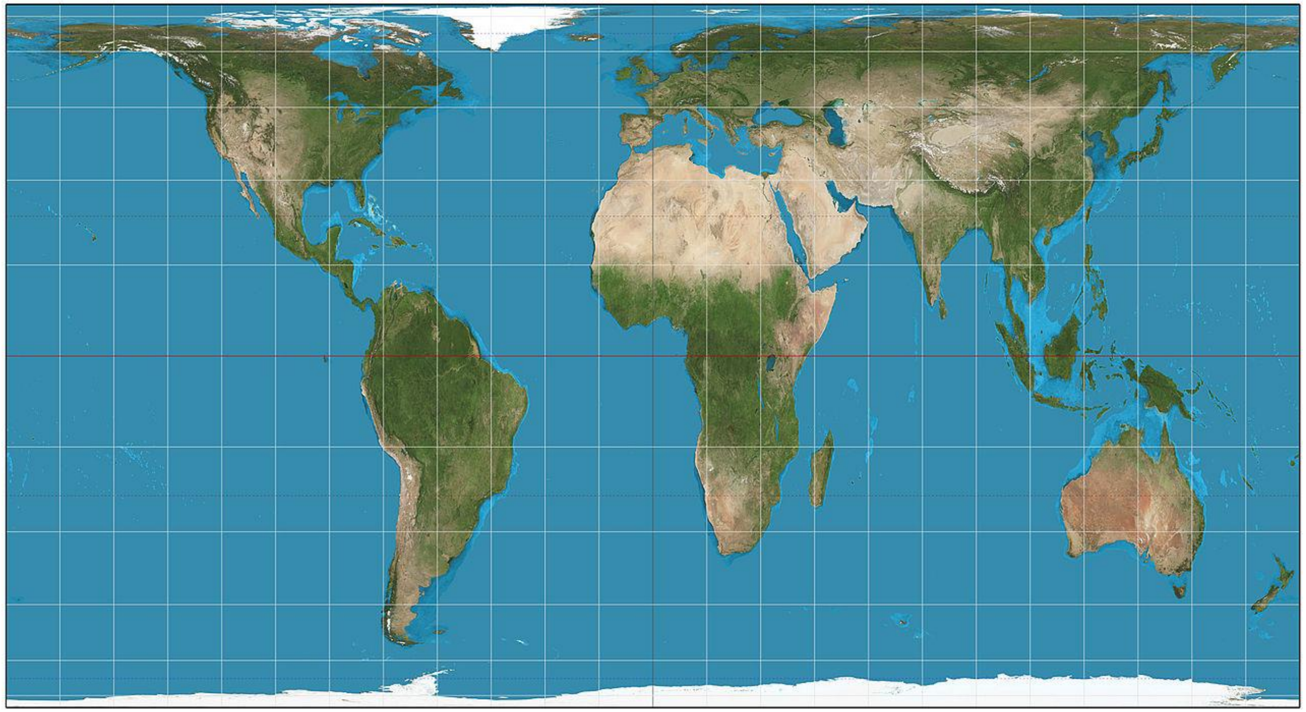
- Data
  - Public domain aggregated catch and effort held by IOTC for all fleets
- Data selection
  - Japanese and Korean fleets only
  - Similar methods, i.e. compatible catch rates
- Time period
  - 1960-1975 (used in YFT before 2016)
  - 1963-1975 (used in YFT 2016)
  - 1975-1994, 1979-1994, 1980-2000 (new alternatives)

Methods 2: density estimation

- Density estimation: progressive changes
  1. Mean unstd CPUE/cell, assume each cell has equal area (means method).
  2. log*(CPUE + c) ~ cell + year-qtr* (standardize with glm).
  3. As above, but estimate cell abundance using ocean area.
  4. As above, with statistical weighting by ocean area (stat wts).
  5. log*(CPUE + c) ~ cell + fleet + year-qtr.*
  6. $\log(CPUE + c) \sim cell + fleet + year + reg.qtr$.
  7. Replace glm with gam, and use spatial smoother.
     $\log(CPUE + c) \sim te(lat5, lon5) + fleet + year + reg.qtr$.
  8. Fill gaps in 6. with cell estimates from 7.

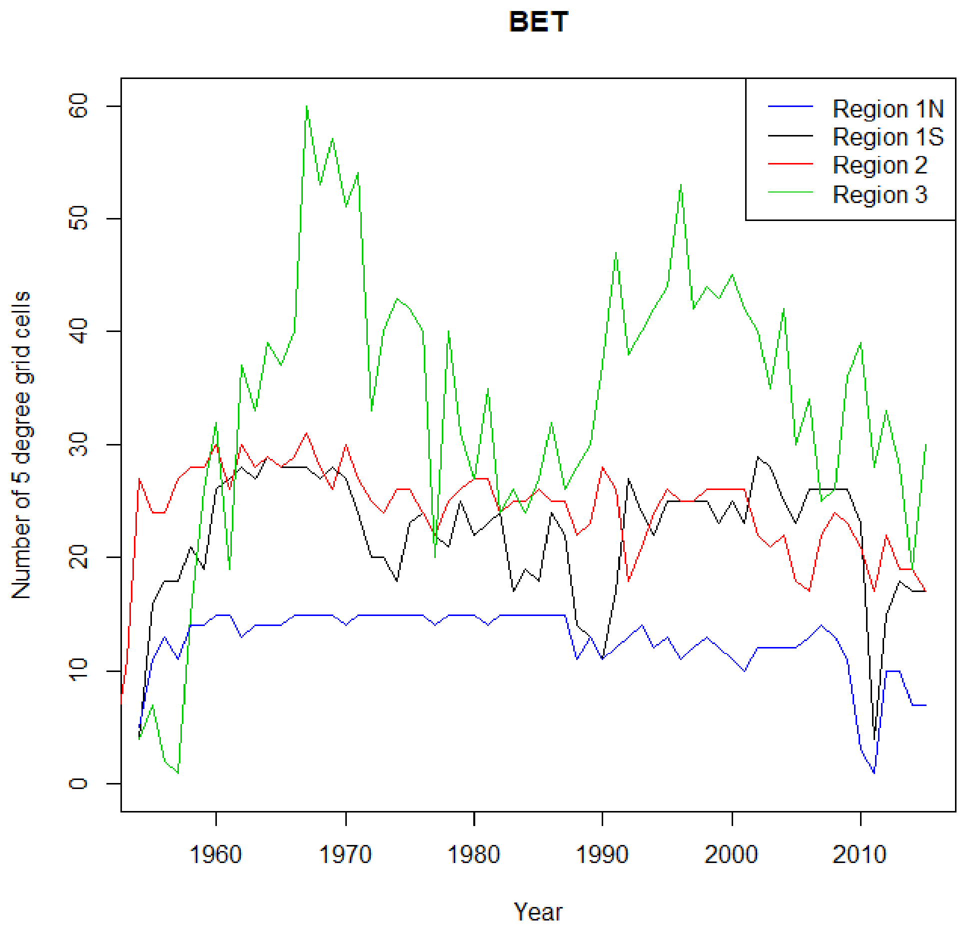
Methods 3: area estimation

- Old approach: equal cell areas
- Proposed approach: estimate cell ocean areas with GIS
  - Ocean area = whole cell – land
  - Cells are smaller at higher latitudes, further from the equator
  - Some cells are mostly land, which affects some regions more than others



Equal area projection

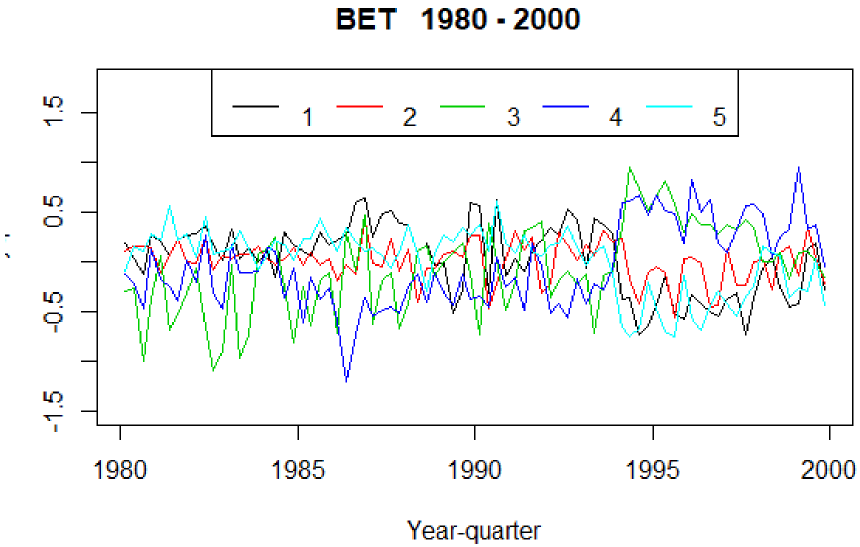
Results 1:
Period requirements & spatial coverage

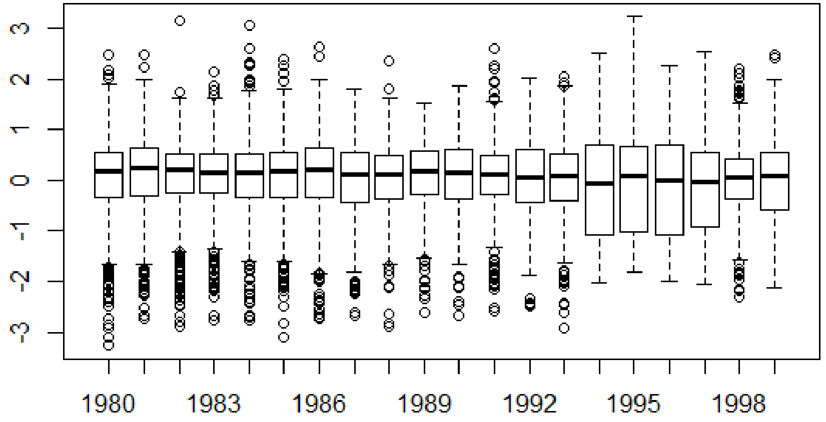


- Bigeye cells fished per year and region.
- Lower coverage 1970-1991 in Region 3, and post-2000 more broadly.
- Uncertainty whether data are representative 1960-1963, and about target change late 1960s - early 1970s.
- Break in CPUE time series in 1979 with introduction of vessel ids.

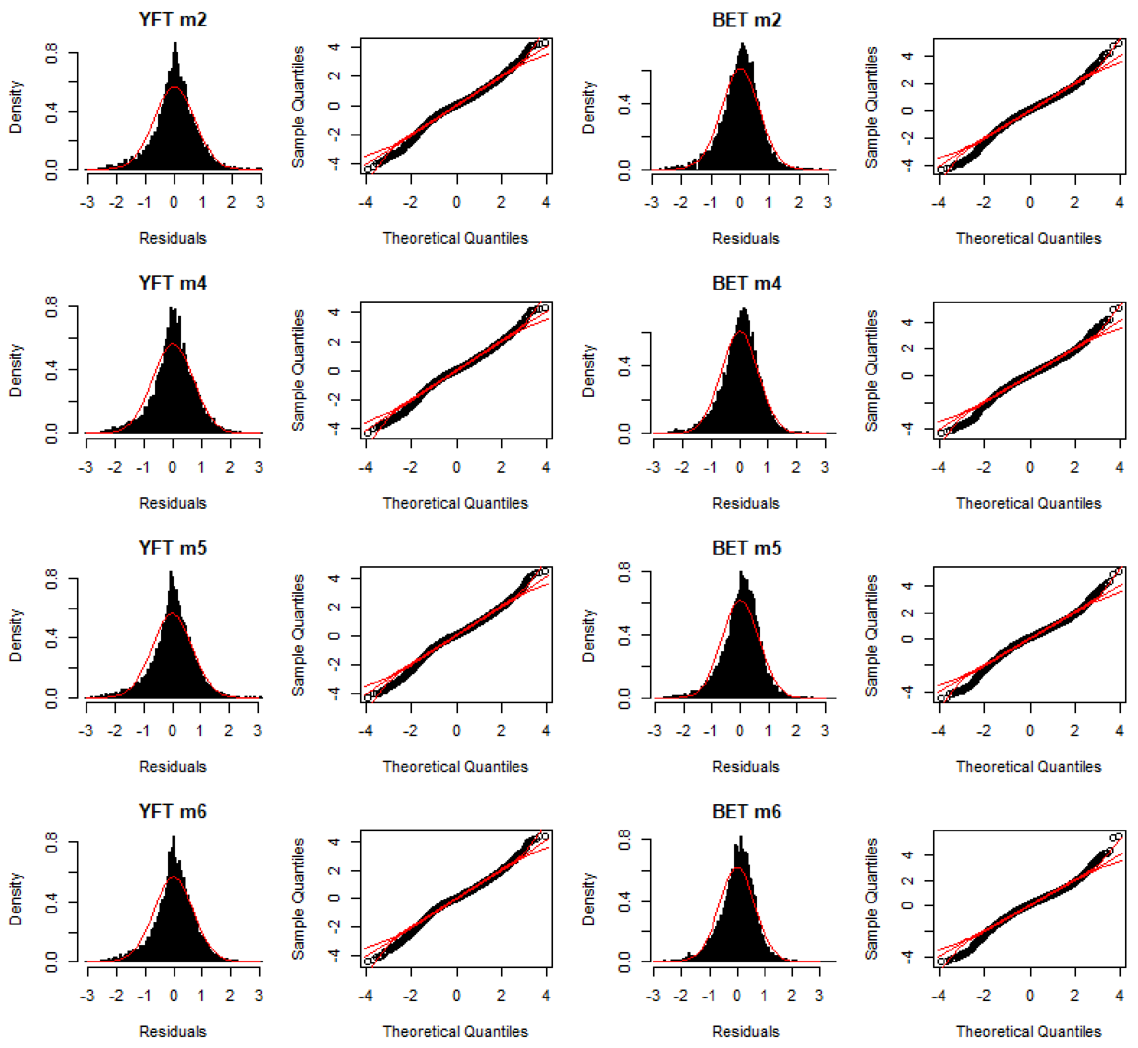
Results 2: period requirements

- Residuals indicate when data consistent with assumptions
  - Mean residuals stable 1979 – 94 (not shown)
  - Residuals markedly more variable 1994-7 (see figure)
  - Suggests that trends diverge among regions
  - Patterns similar in BET & YFT





Standardization diagnostics, 1979-1994



- Close to normally distributed
  - In aggregated data, rows made up of fewer sets are more variable, causing some non-normality
- 1980-2000 analysis has another peak on the left due to excess zero sets

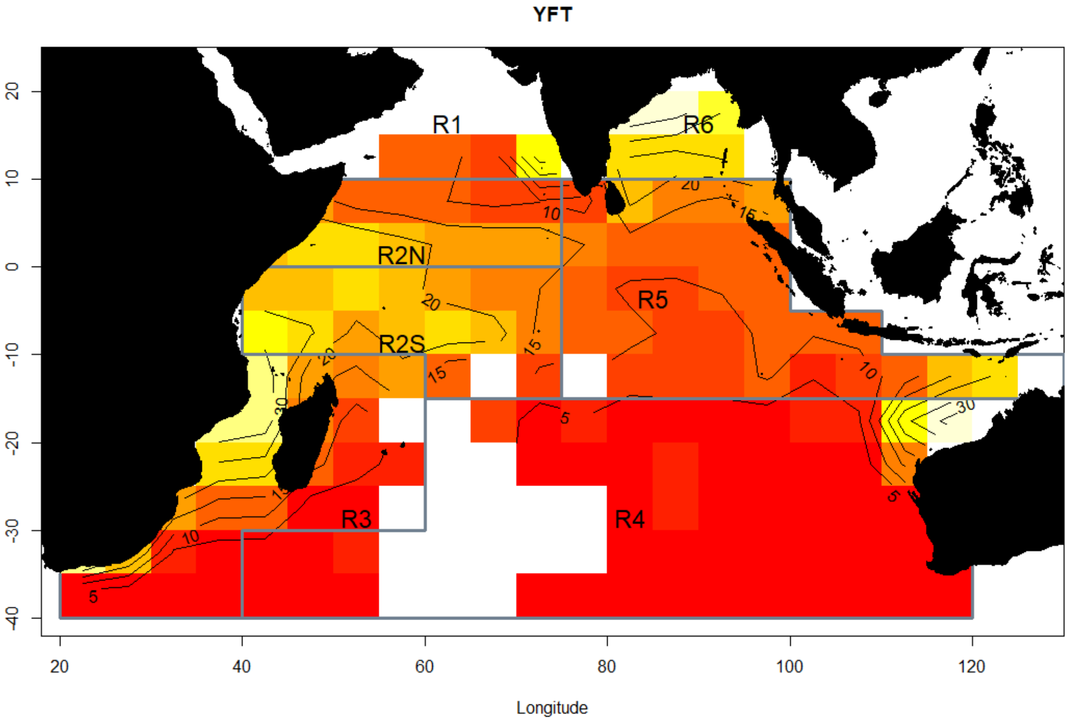
Results : Model selection 1979-1994

Species	Method	Variable dropped	Df	AIC	δAIC
Bigeye	5	-	-	26720	0
		year-qtr	59		376
		cell	119		6085
		fleet	1		147
	6	-	-	26547	0
		year	14		95
		cell	115		3432
		fleet	1		120
		reg.qtr	15		450
Yellowfin	5	-	-	31308	0
		year-qtr	59		681
		cell	146		10882
		fleet	1		103
	6	-	-	31051	0
		year	14		464
		cell	140		5329
		fleet	1		110
		reg.qtr	21		469

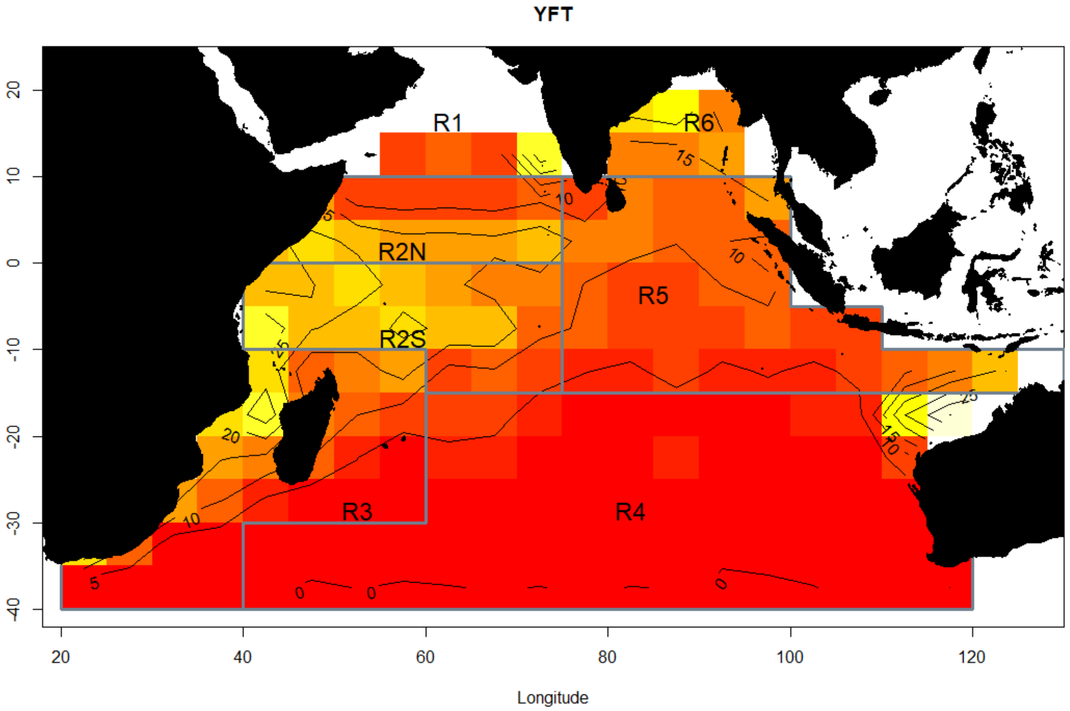
- For both BET and YFT
  - Method 6 (reg.qtr interaction) is preferred over 1 to 5.
  - All variables are statistically significant
- Method 8 selected
  - Most cells are estimated using method 6
  - Missing cells are filled using method 7 (spatial smoother)

Heat maps for YFT, 1979-1994 period

Means method

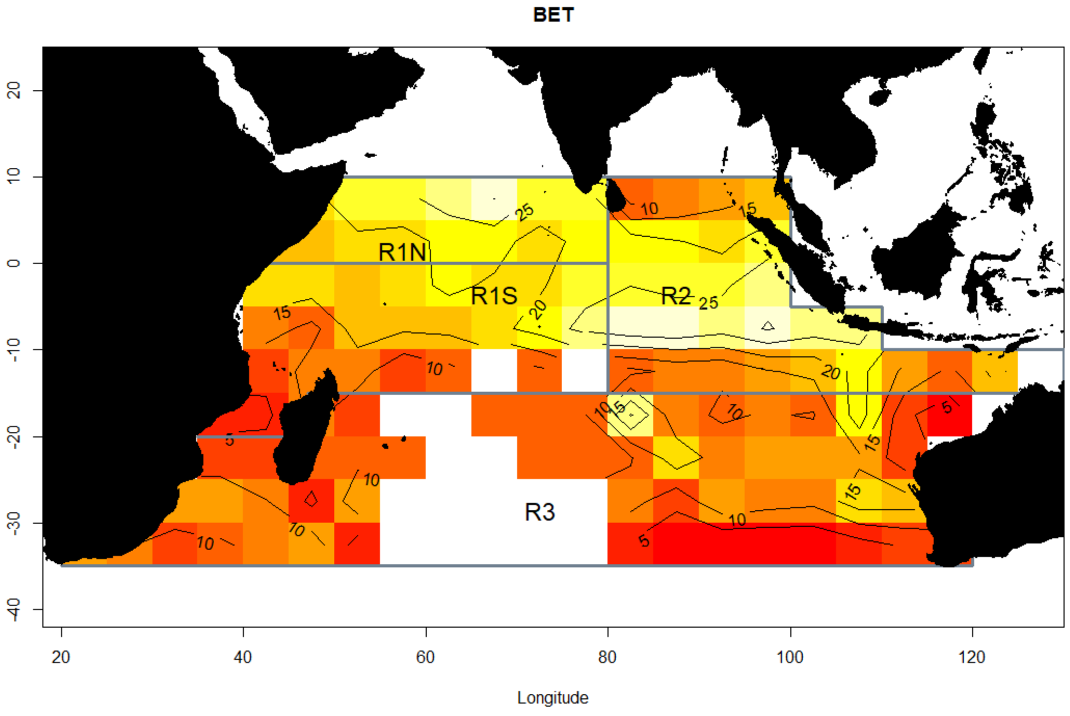


Method 8

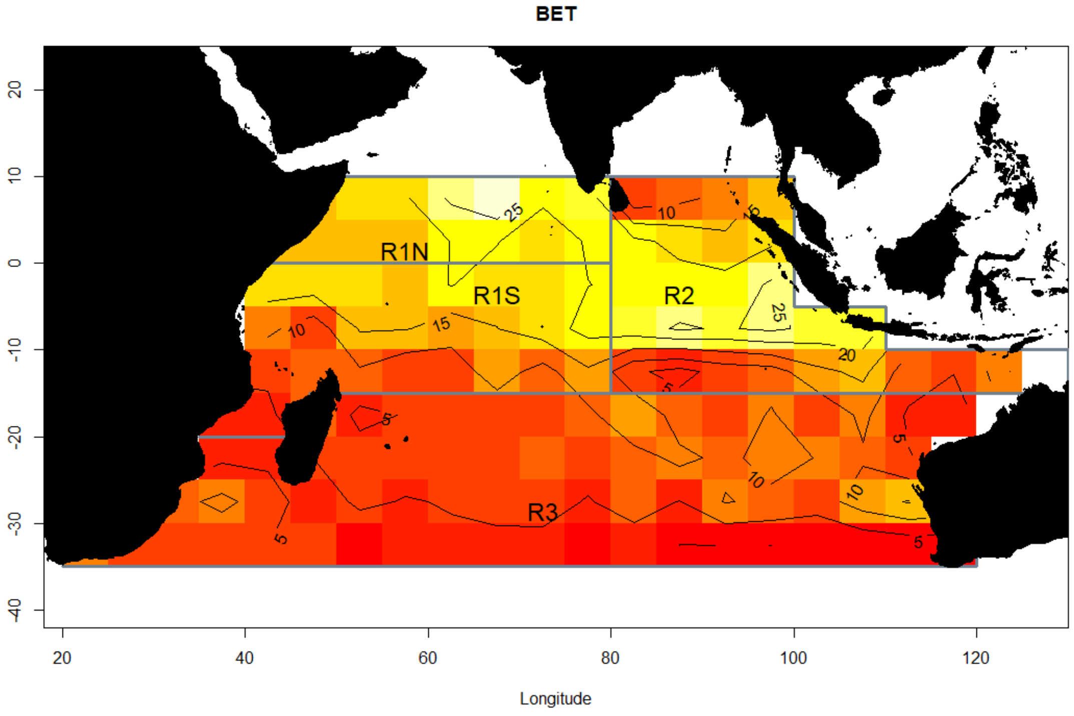


Heat maps for BET, 1979-1994 period

Means method

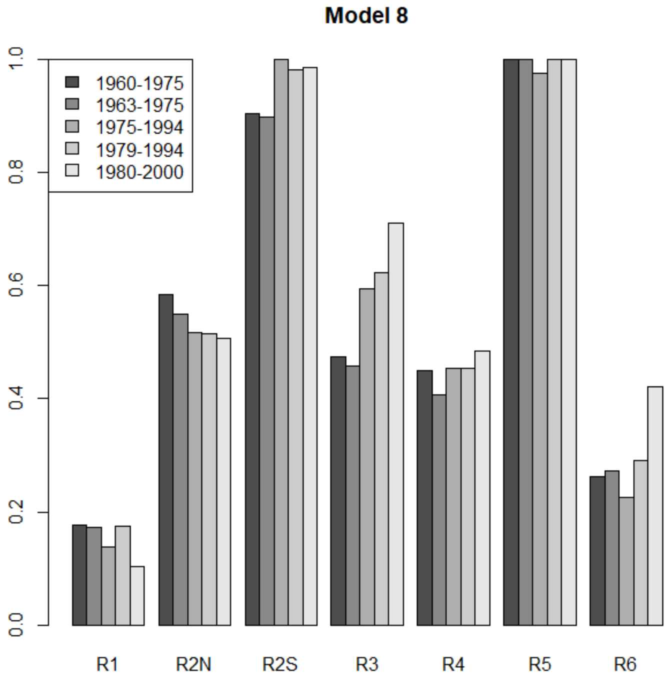


Method 8

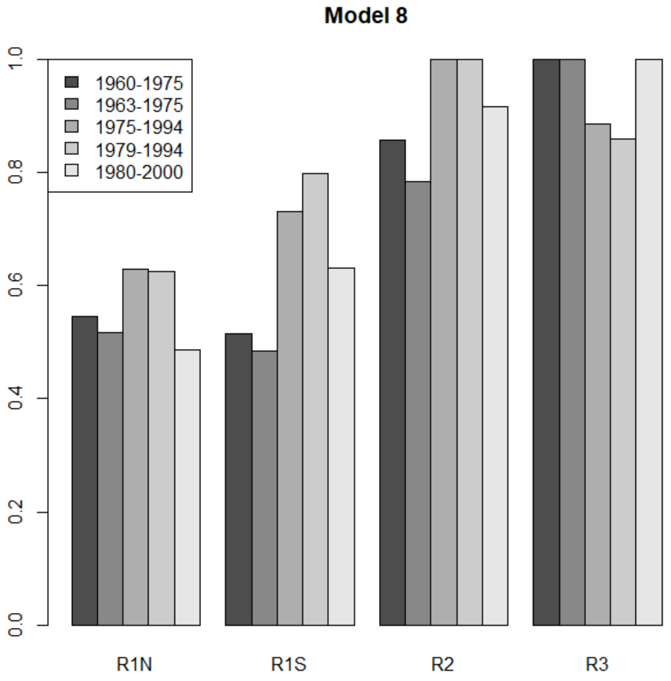


Regional scaling parameters compared by period, adjusted to same baseline

YFT

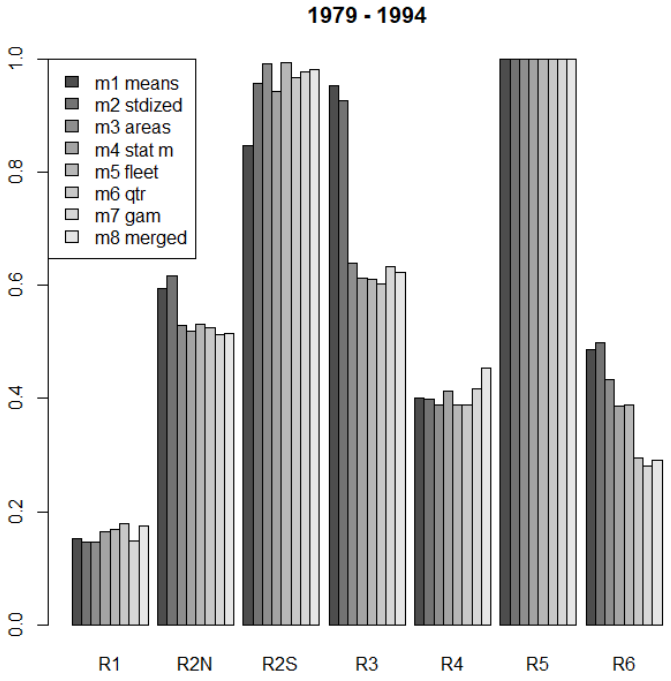


BET

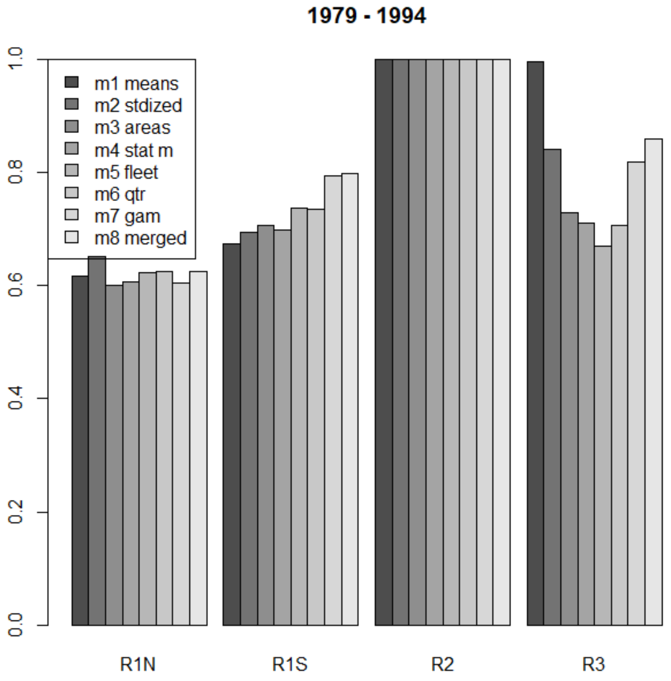


Regional scaling parameters compared by method

YFT



BET



Conclusions

- A primary objective was for the approach to be documented and peer-reviewed.
- Both the choice of period and the method have moderate to large effects on results.
  - The 1979-1994 period best meets the period requirements
  - Allowing for ocean areas is inarguably better, and the most influential change
  - The method 6 model fits the data best
  - Method 8 is recommended because the gaps are filled
- Future work
  - Analyses using operational data are needed, to adjust for target change.
    - Code development required, and approaches to deal with memory issues due to large datasets.
    - Could integrate into the normal CPUE standardization process.
  - Explore information from other fleets, to estimate significance of biomass outside the standard regions.