IOTC longline tuna CPUE: Collaborative analyses

Simon Hoyle, Takayuki Matsumoto, Keisuke Satoh, Yu-Min Yeh, Yin Chang, Doo Nam Kim, Sung Il Lee, Hiroaki Okamoto, Toshihide Kitakado.

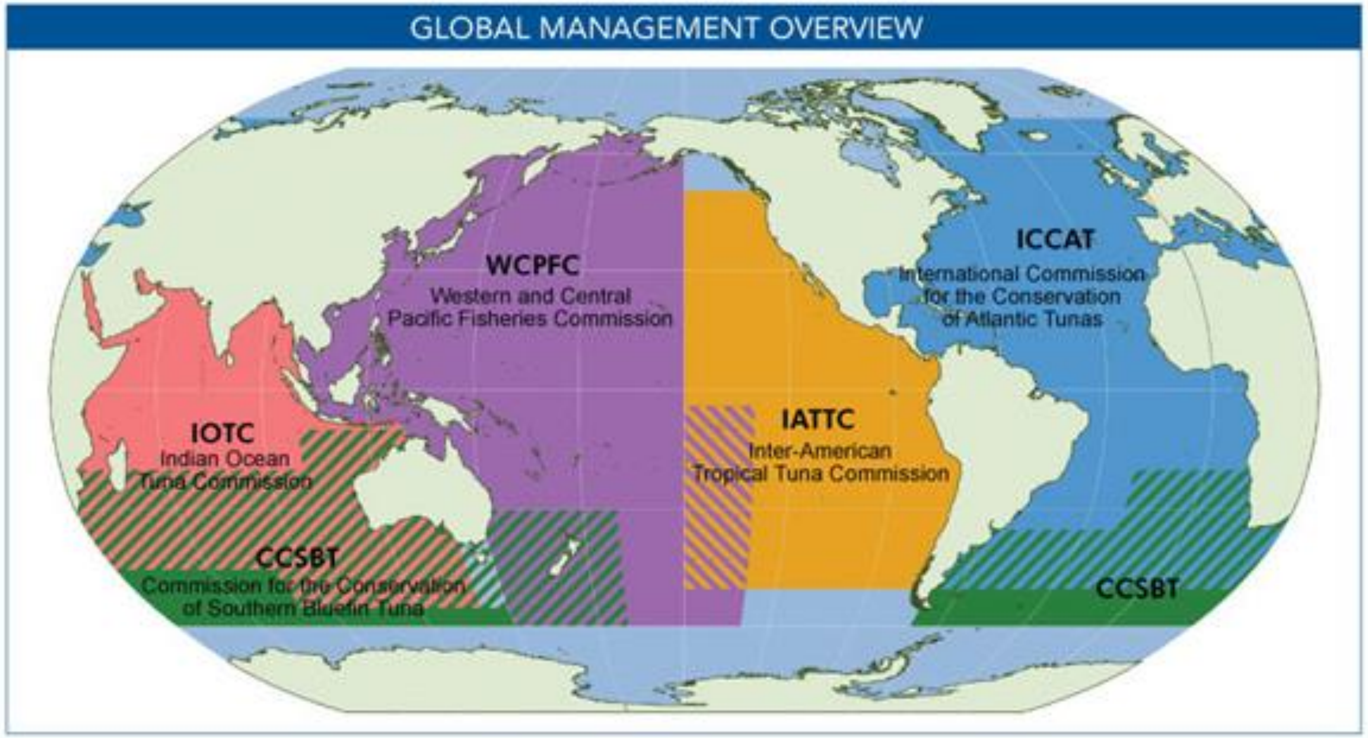
Indian Ocean Tuna Commission
Commission des Thons de l'Océan Indien
iotc ctoi

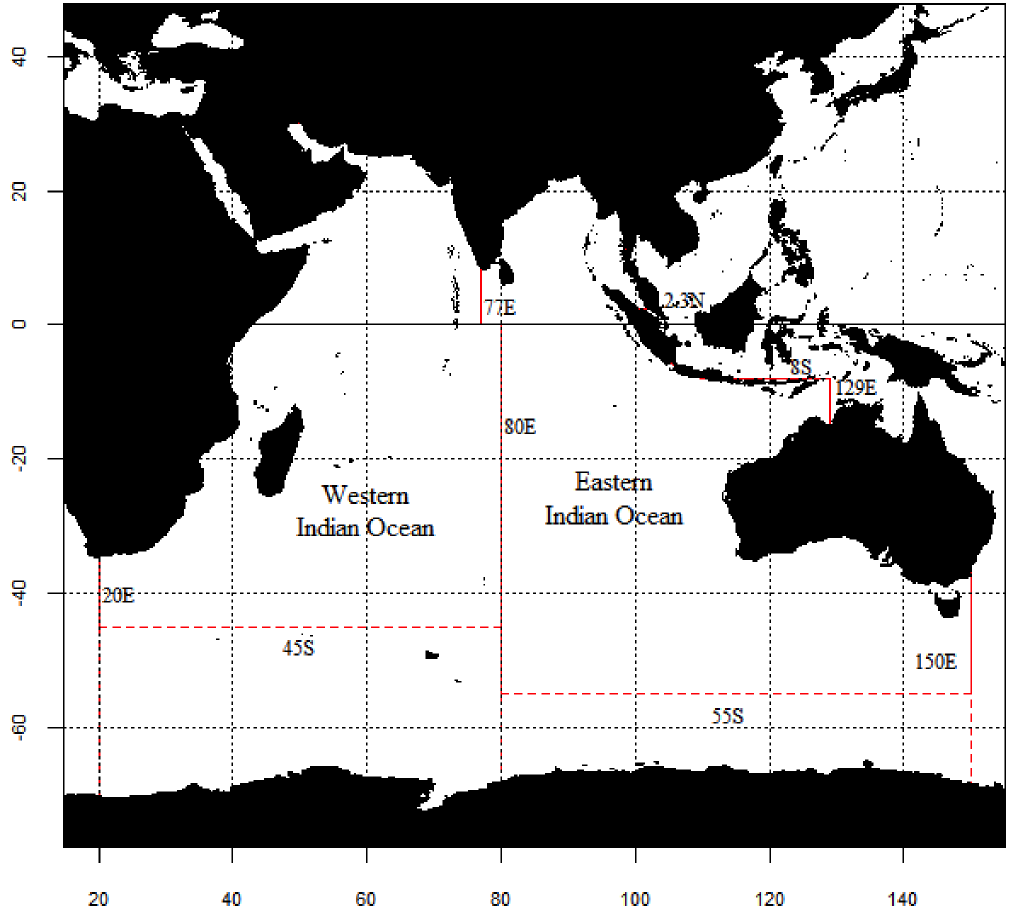
INTERNATIONAL SEAFOOD SUSTAINABILITY FOUNDATION

Introduction

- IOTC and the motivation & history of this work
- Targeting analyses & clustering
- Standardization methods
  - Basics
  - Modelling spatial effects
  - Adjusting for fleet movement
    - Area weighting
    - Spatial infilling
- Other issues
- Regional scaling

IOTC area


GLOBAL MANAGEMENT OVERVIEW
WCPFC
Western and Central Pacific Fisheries Commission
ICCAT
International Commission for the Conservation of Atlantic Tunas
IATTC
Inter-American Tropical Tuna Commission
IOTC
Indian Ocean Tuna Commission
CCSBT
Commission for the Conservation of Southern Bluefin Tuna
CCSBT


77E
2.3N
8S
129E
80E
Western Indian Ocean
Eastern Indian Ocean
20E
45S
150E
55S
40
20
0
-20
-40
-60
20
40
60
80
100
120
140

Background

- Longline CPUE indices of abundance
  - Most important factor driving stock assessment outcomes.
  - Pelagic species: bigeye, yellowfin, albacore, billfish, sharks
- ‘Triage’ required
  - Multiple problems to address
  - Limited analysis time
  - Focus on the issues considered most likely to affect biomass trends

Motivation

- Joint analysis
  - Japanese and Taiwanese bigeye tuna CPUE show different trends in some periods, which needed to be resolved.
  - Sparse data provided poor indices in some areas and years if using just one fleet
    - Japan fishery contracting spatially due to piracy & competition, low effort recently
    - Taiwanese fishery started later, sparse data in some periods, reliability concerns
    - Korea smaller dataset, can help fill gaps & identify issues in other datasets
    - Seychelles (added in 2017) can help fill spatial gaps in recent years
- Methodological issues
  - Target change through time was significant and affected indices.
  - Agreed, standard, and updated methods were needed for issues such as fleet turnover, environmental covariates, and spatial effects.

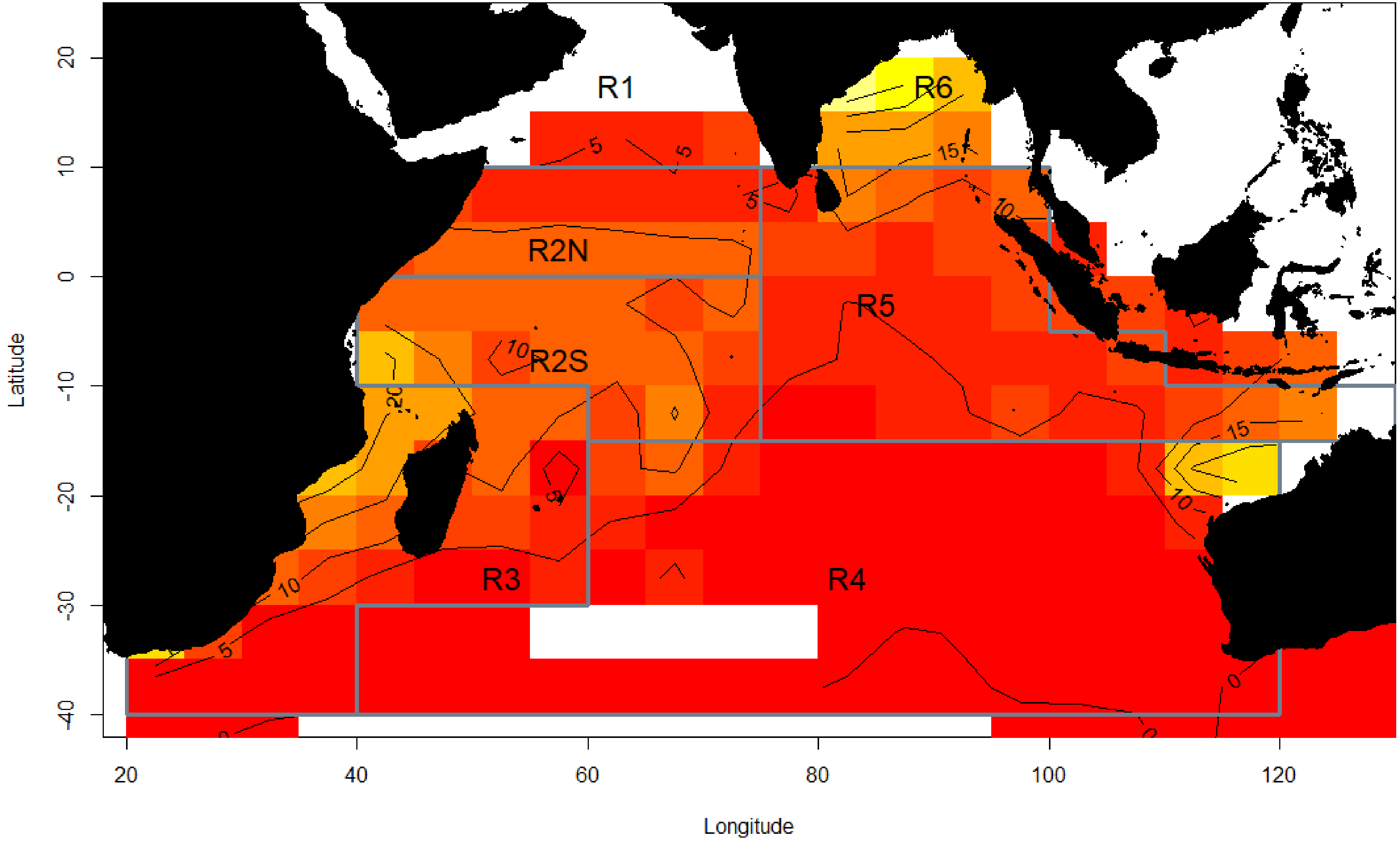
Progress

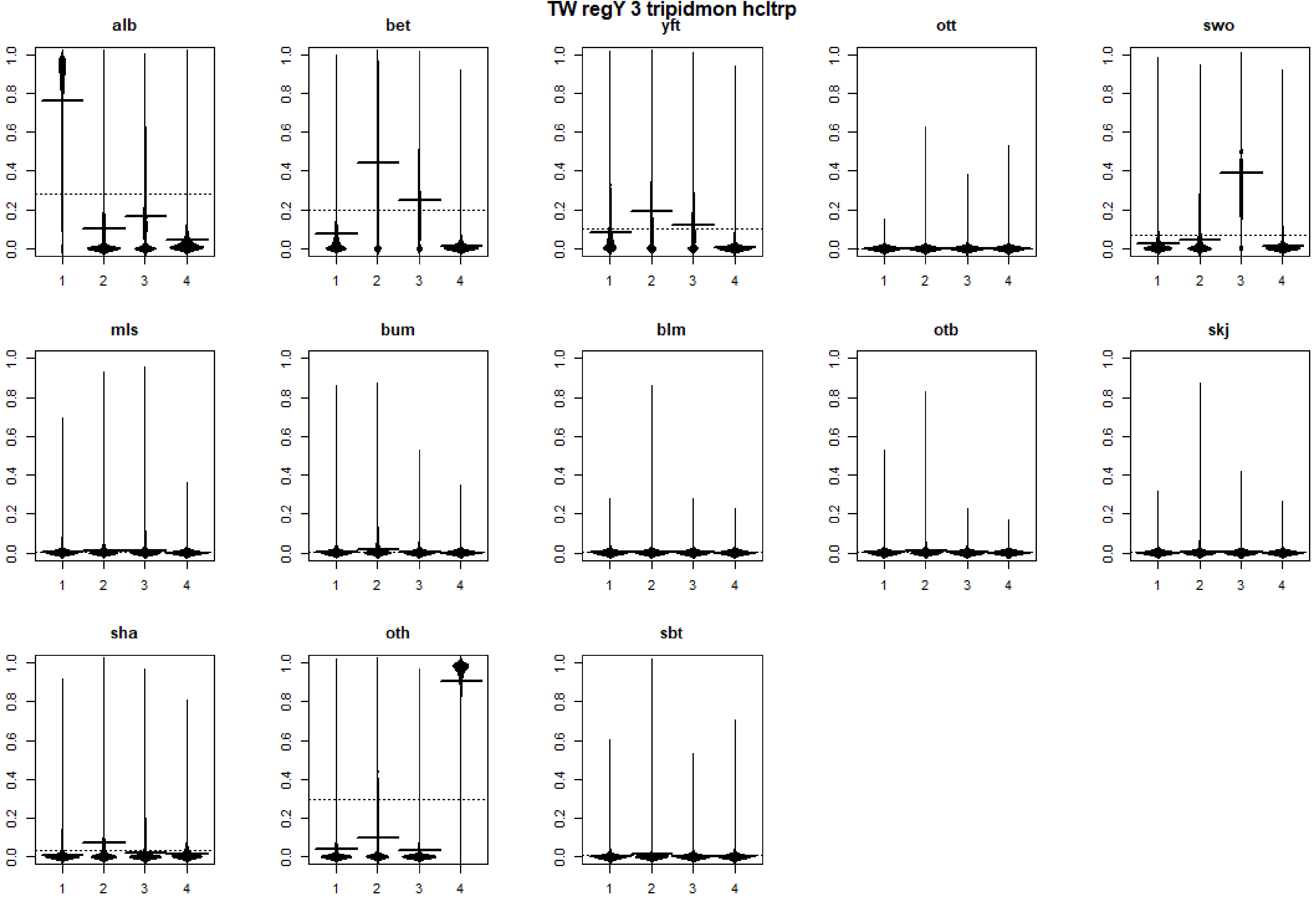
- History
  - 2015: Project started with YFT & BET, first access to operational data from JP, KR, TW
  - 2016: ALB added, and indices first used in assessments
  - 2017: Seychelles data added, explored relationships with size data, time-area interactions
- Logistics
  - Limited access to operational data, 2-3 weeks per year
  - In-country meetings for data exploration and preparation
    - Last year provided training for national scientists to prepare & cluster data, and develop national indices using standard approaches
  - Joint meetings for joint analysis, training, and discussion
  - We have to provide indices

Analysis process

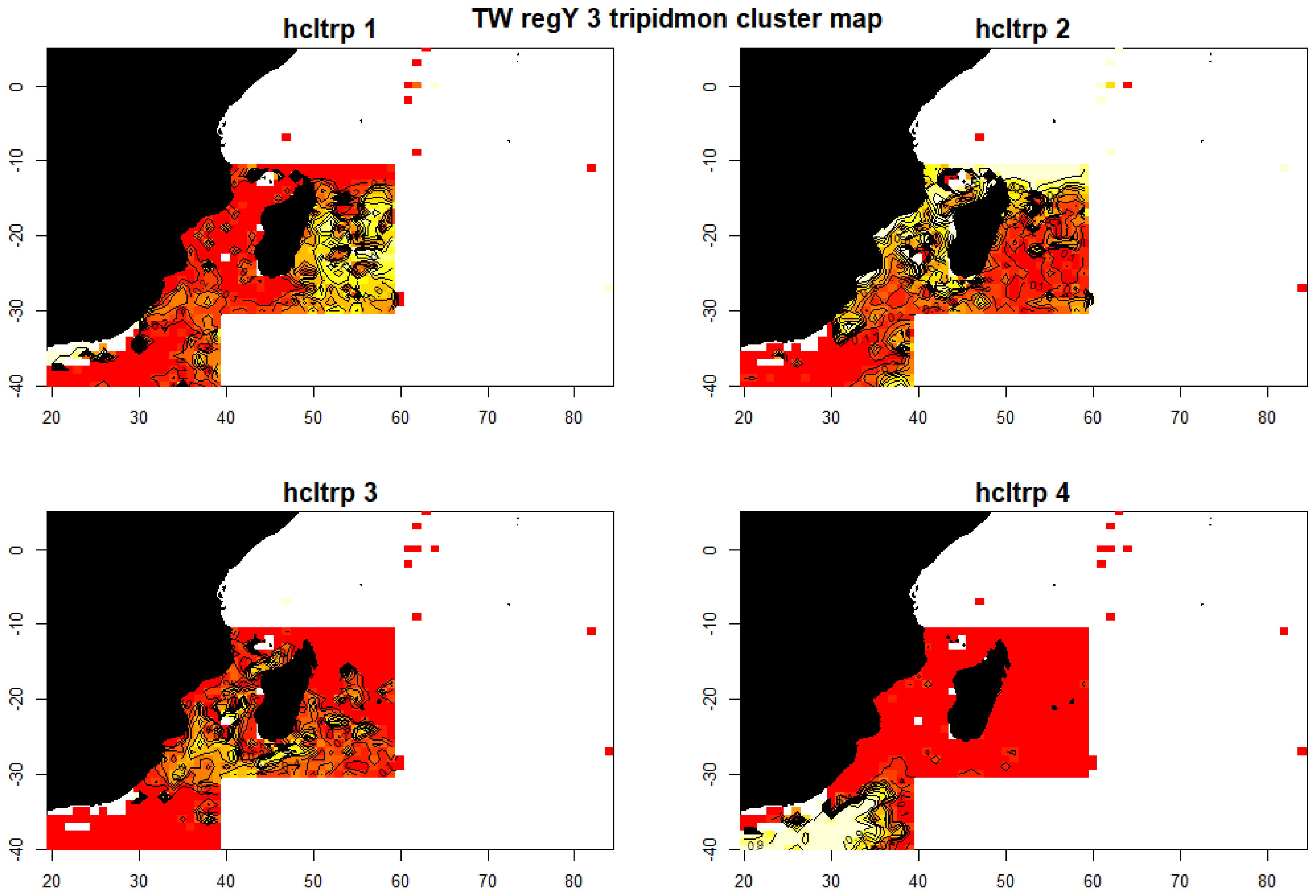
1. Load, clean data
2. Explore data
   - Plot, document everything
3. Targeting analyses
   - Clustering by species composition to identify fisheries
4. Standardization

Assessment regions




TW regY 3 tripidmon hcltrp
alb
bet
yft
ott
swo
mls
bum
blm
otb
skj
sha
oth
sbt
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4

TW YFT R3

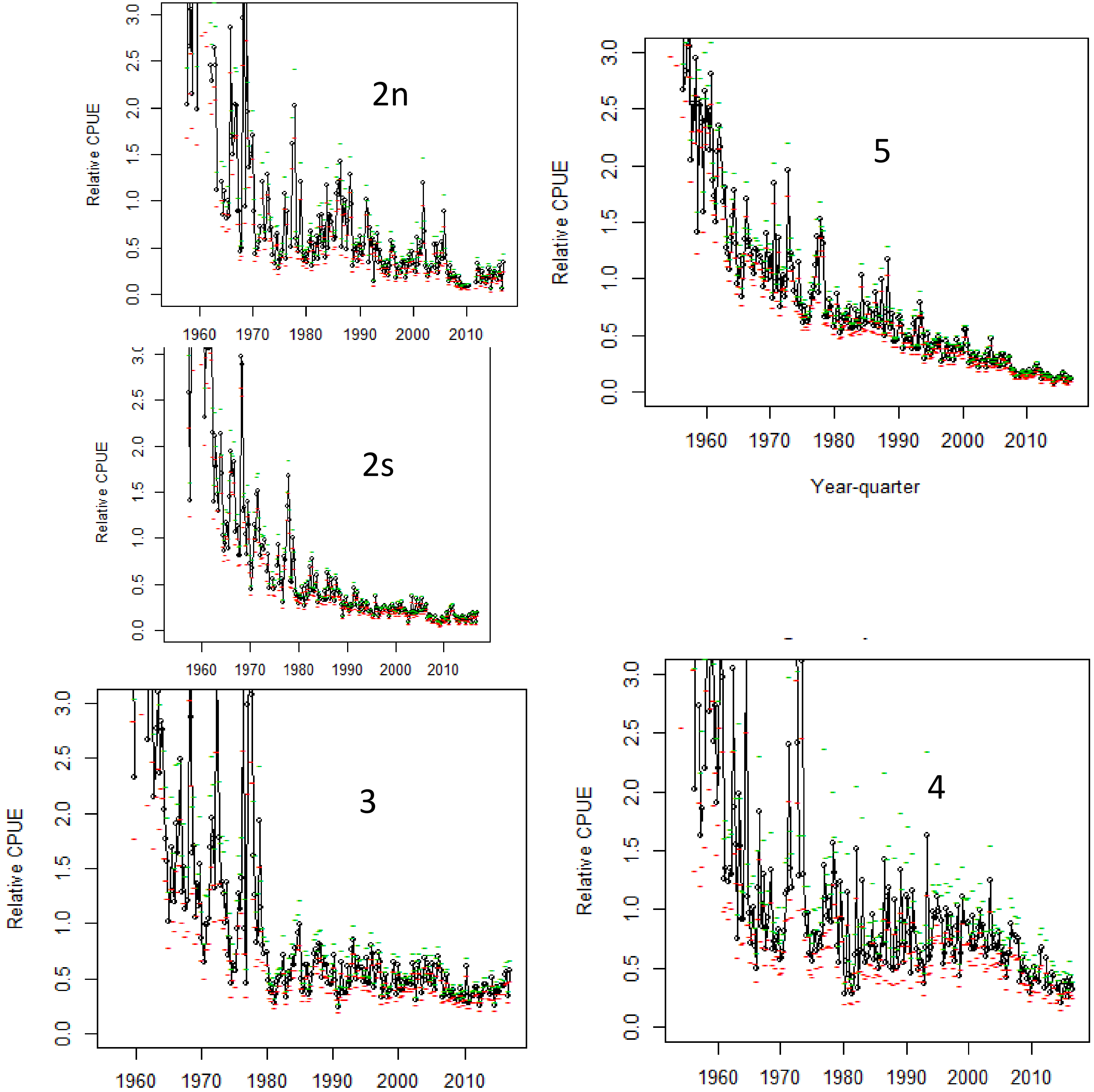

TW regY 3 tripidmon cluster map
hcltrp 1
hcltrp 2
hcltrp 3
hcltrp 4
0
-10
-20
-30
-40
20
30
40
50
60
70
80

CPUE standardization methods, basics 1

- Generalized linear models in R, modelling CPUE at the set level
- Data omits clusters catching very few of target species
- Delta lognormal
  - $(CPUE = 0) \sim yrqtr + vessel + latlong5 + (cluster\ or\ HBF) + \epsilon$
  - $\log(CPUE) \sim yrqtr + vessel + latlong5 + (cluster\ or\ HBF) + \epsilon$, for nonzero sets
- *yrqtr*, *vessel*, *latlong5*, and *cluster* are categorical variables
- Hooks between floats *(HBF)* parameter is a cubic spline

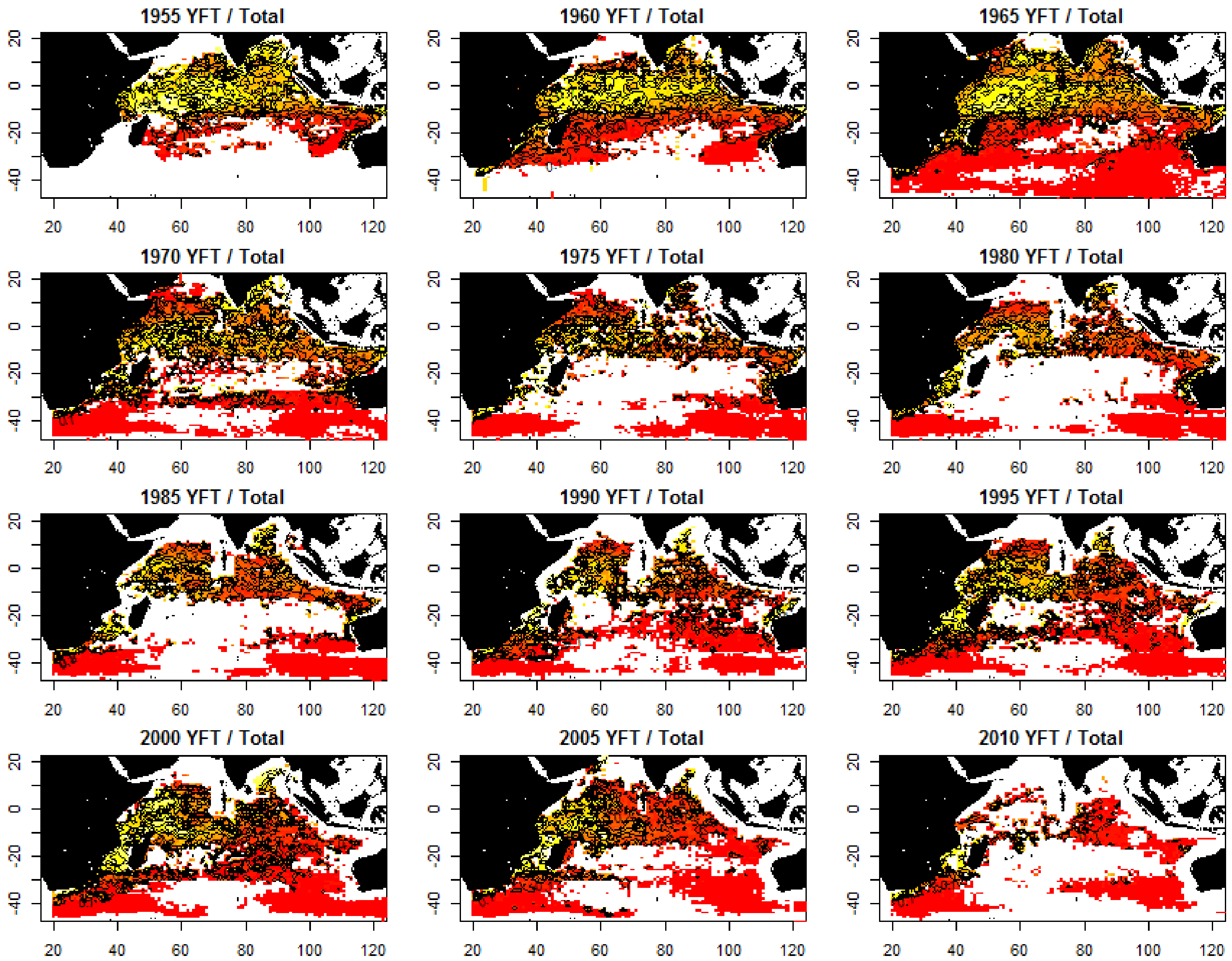
Approach to modelling spatial effects

- Assessment regions are modelled independently, consistent with their treatment in the assessment.
- Within a region, 5° cells are modelled as independent categorical variables.



Changes in distribution & coverage

- In a 65 year dataset (since 1952), effort concentration moves around
  - Causes
    - Initial expansion into new areas
    - Area closures due to EEZs, Somali piracy
    - Markets changing target preferences, e.g. sashimi market raising value of BET/YFT vs ALB
  - Effects
    - Areas without effort
    - Changing statistical weights among areas, biasing the indices



Within region: What to assume about areas without effort?

1. Time area interactions, spatial infilling (one size does not fit all)
    1. During expansion, unfished areas have high biomass & higher CPUE
        - Unfished areas never fished, so assume ~ initial CPUE in those areas
        - But catchability probably higher in the initial phase
    2. Later, when index fleet leaves an area, assumptions depend on …
        - Do other fleets remain (e.g. exclusion from EEZ, outcompeted by other fleets)?
        - Is there less fishing effort (piracy)? Biomass may trend up.
2. Within a region, model is CPUE ~ time + area, which avoids the need for infilling
    - Problematic to the extent that fish distributions change
3. Combined approach – explored last year
    - Time x area model (latlong5 + lat5 * qtr + lat5 * year)
    - Fill time-area 'holes' with estimates from time + area model

Biases due to changing effort distribution

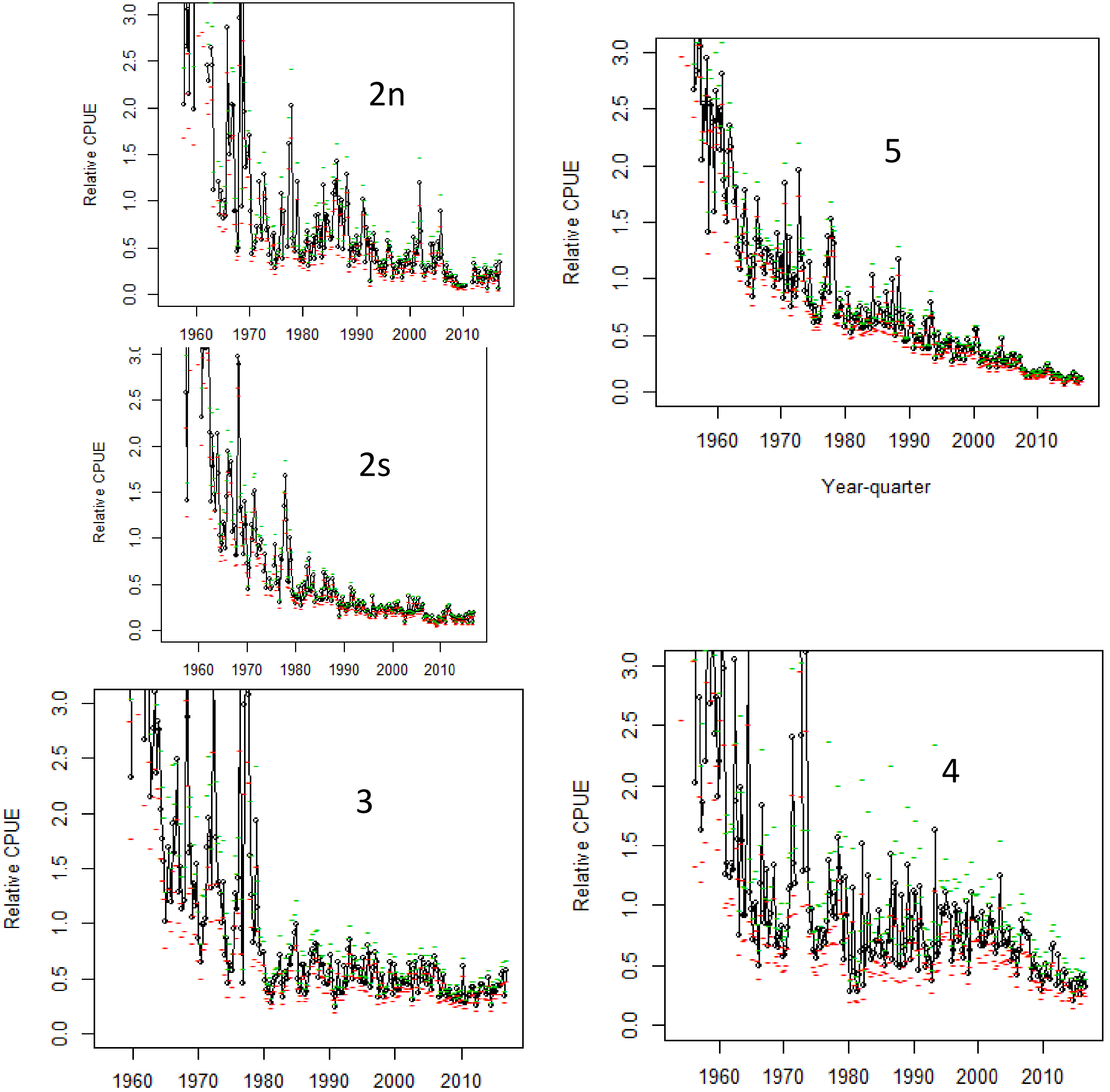
- Shifting effort introduces bias. We do the following:
  1. Remove 5° cells with fewer than N1 sets across all years
  2. Randomly select N2 sets from each yq*cell stratum (applied when total # sets in dataset > limit ~ 60000)
  3. Adjust statistical weights to give each yq stratum the same influence (Punsly 1987, Campbell 2004)

  For set *j* in area *i* and year-qtr *t*, $w_{ijt} = \frac{log(h_{ijt}+1)}{\sum_{j=1}^{n} \log(h_{ijt}+1)}$

CPUE standardisation – some details

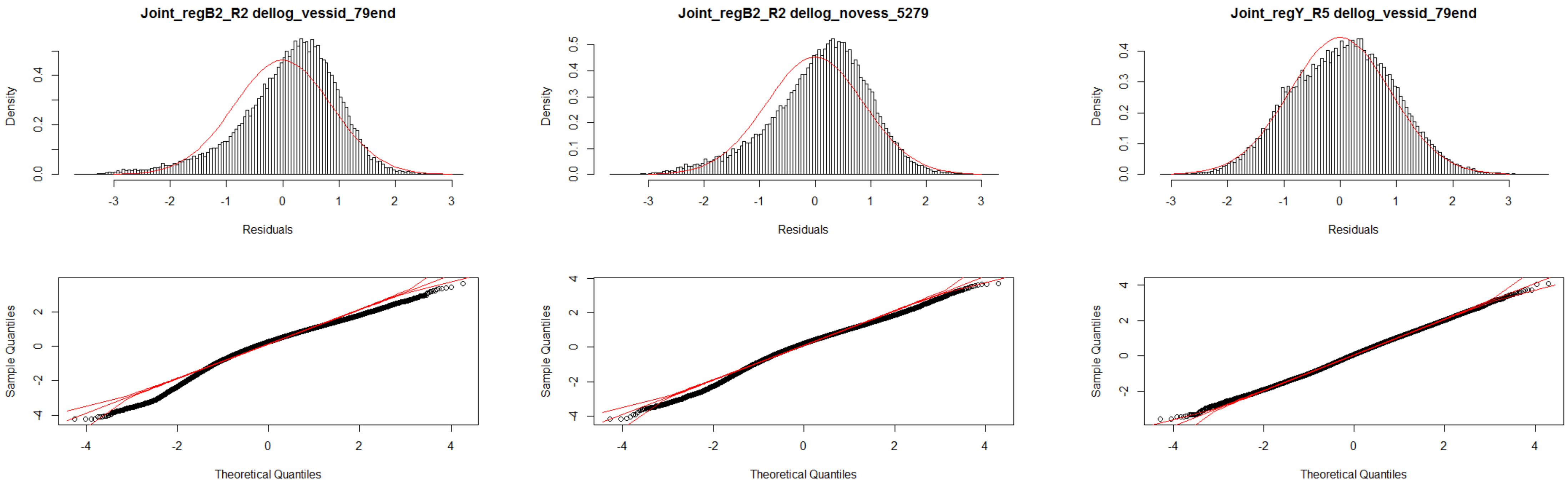
- Problems with large datasets and multiple strata
  - Very long runtimes
  - Large memory use (> 16GB)
  - Hard to debug and fix problems
- Solutions
  - Reduce number of strata
    - Remove vessels fishing < N_1 qtrs
    - Remove cells, yr-qtrs, & vessels with < N_2 sets
  - Subsample data at random
    - Randomly sample (without replacement) N_3 sets from each year-qtr x cell stratum
    - Tested with WCPO data, indices stable with ~ 15 sets per stratum (Hoyle and Okamoto 2011)
    - Limited benefit from extra precision – important sources of uncertainty are elsewhere

YFT indices

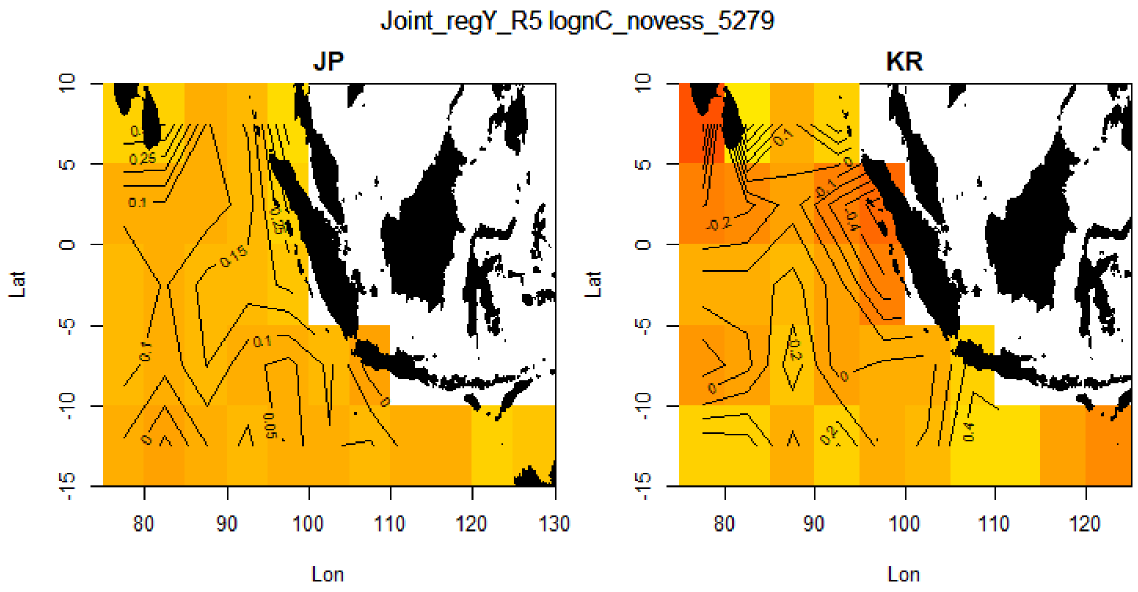


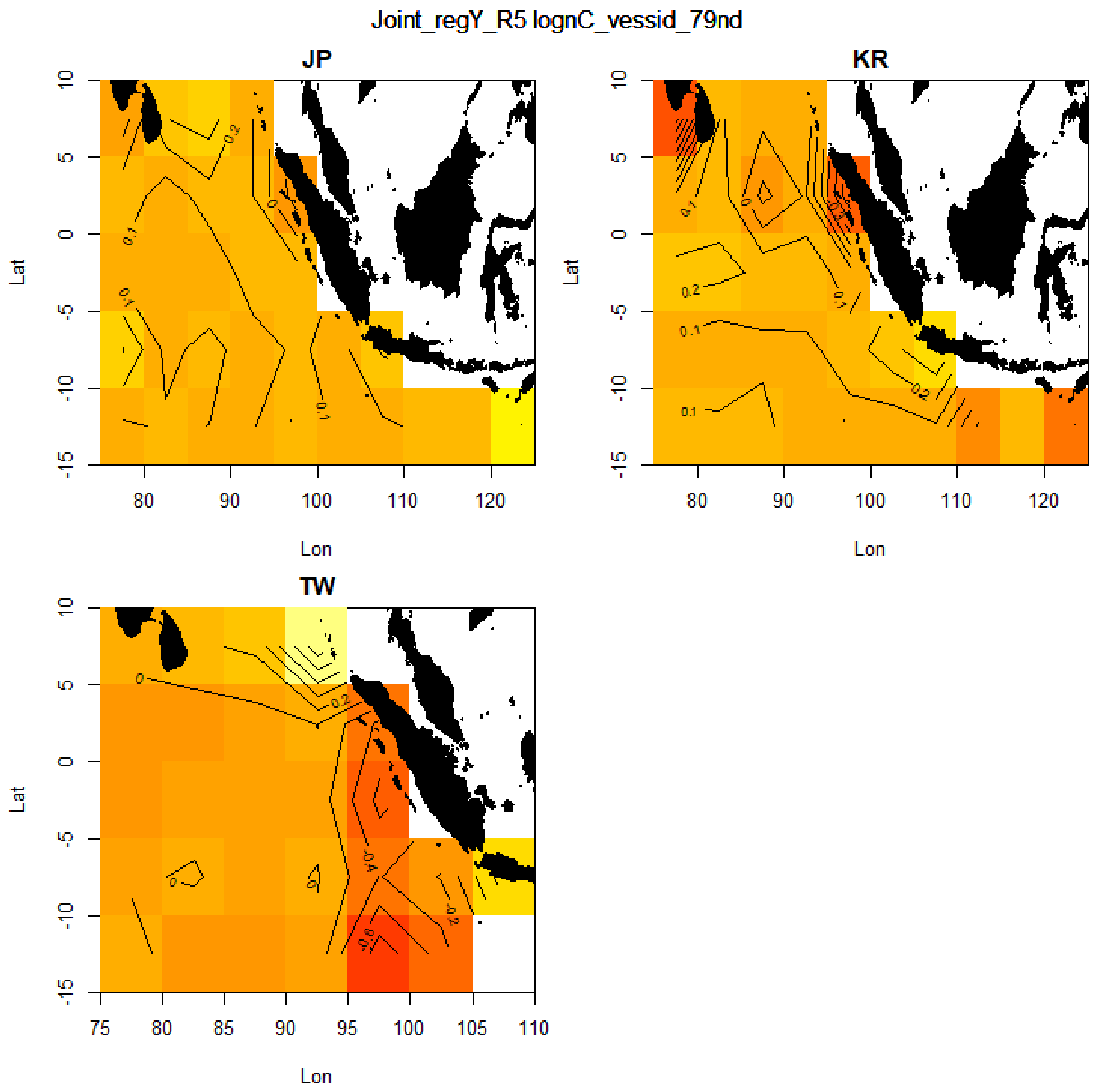
2n
2s
3
4
5
Relative CPUE
Year-quarter
0.0
0.5
1.0
1.5
2.0
2.5
3.0
1960
1970
1980
1990
2000
2010

Distribution diagnostics



Spatial patterns in residuals – east tropical yellowfin





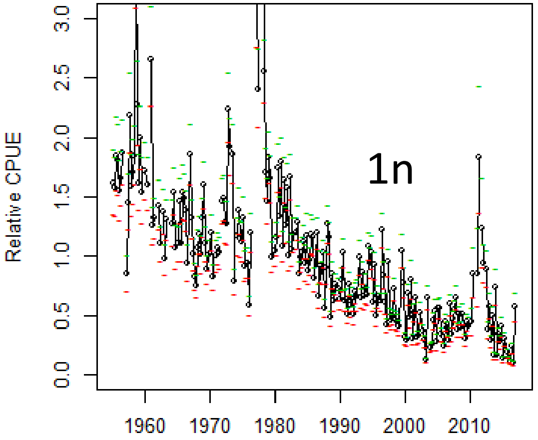
Residual concerns

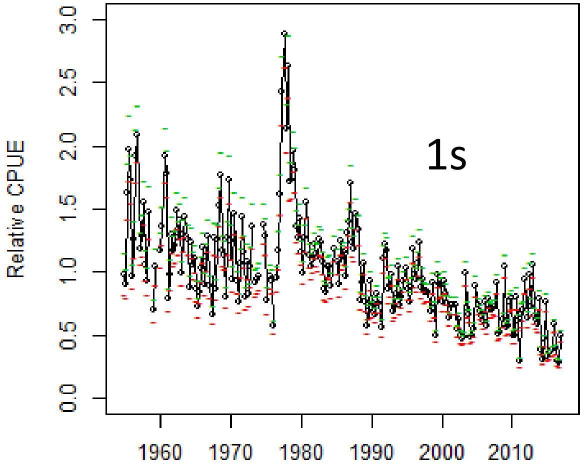
- Potential for differences between (& within) fleets
  - Factors not available for analysis
  - Different bait, gear configurations, reporting behaviour
  - Time series patterns in individual vessel behaviour
- Model issues
  - Assuming no interactions, e.g. between:
    - Targeting behaviour and vessel catchability
    - Season and spatial effects
- Possible future options
  - Random effects on e.g. vessel by target, to permit exploration of interactions
  - mgcv: as before, but add te(lat, lon, yr) + te(lat, lon, qtr)
  - VAST

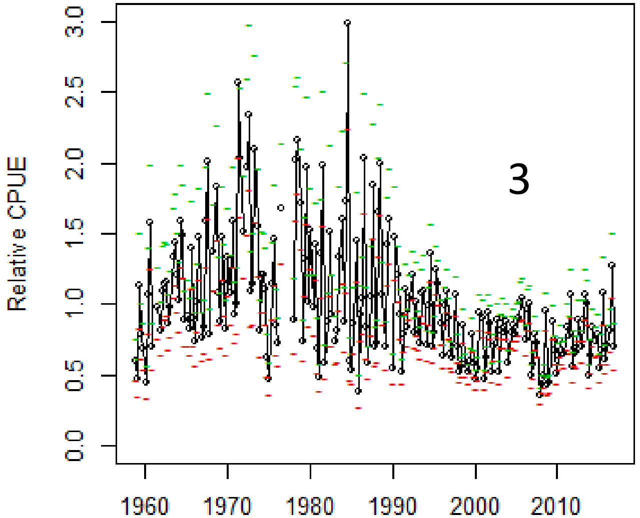
Other issues

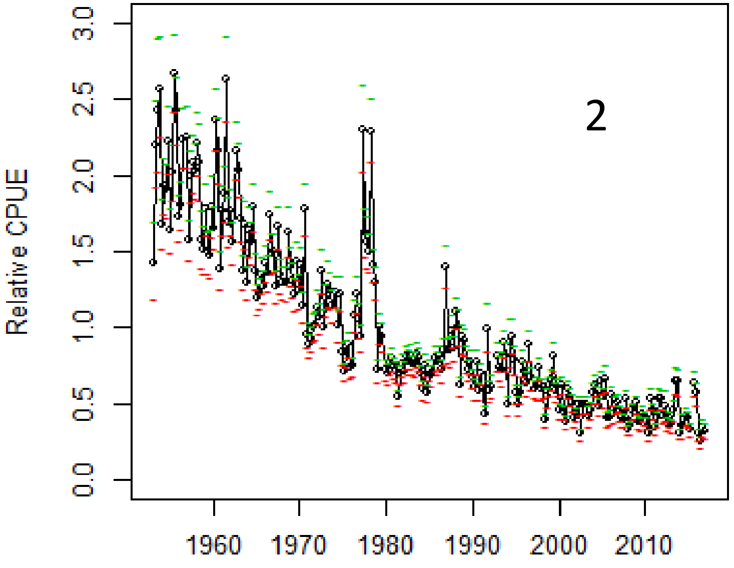
1: Spike in late 70s

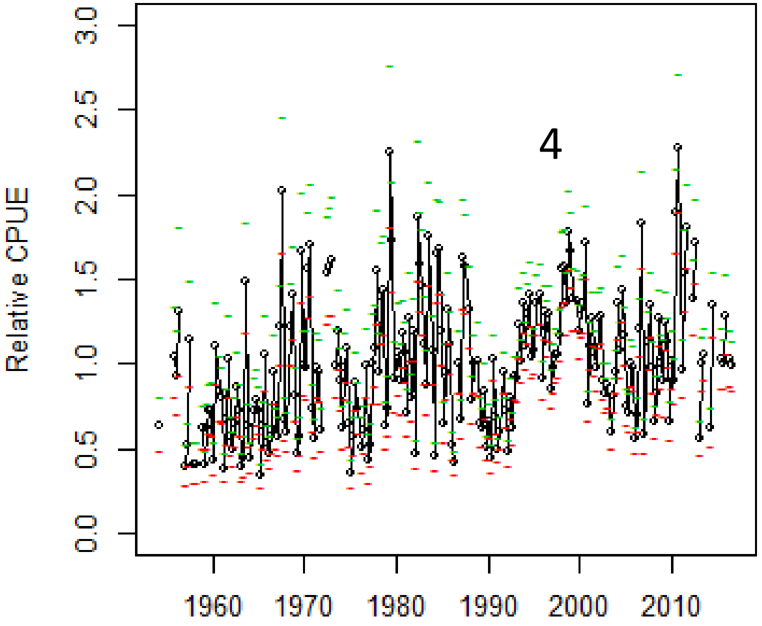
2. Post-piracy local spike ~ 2010



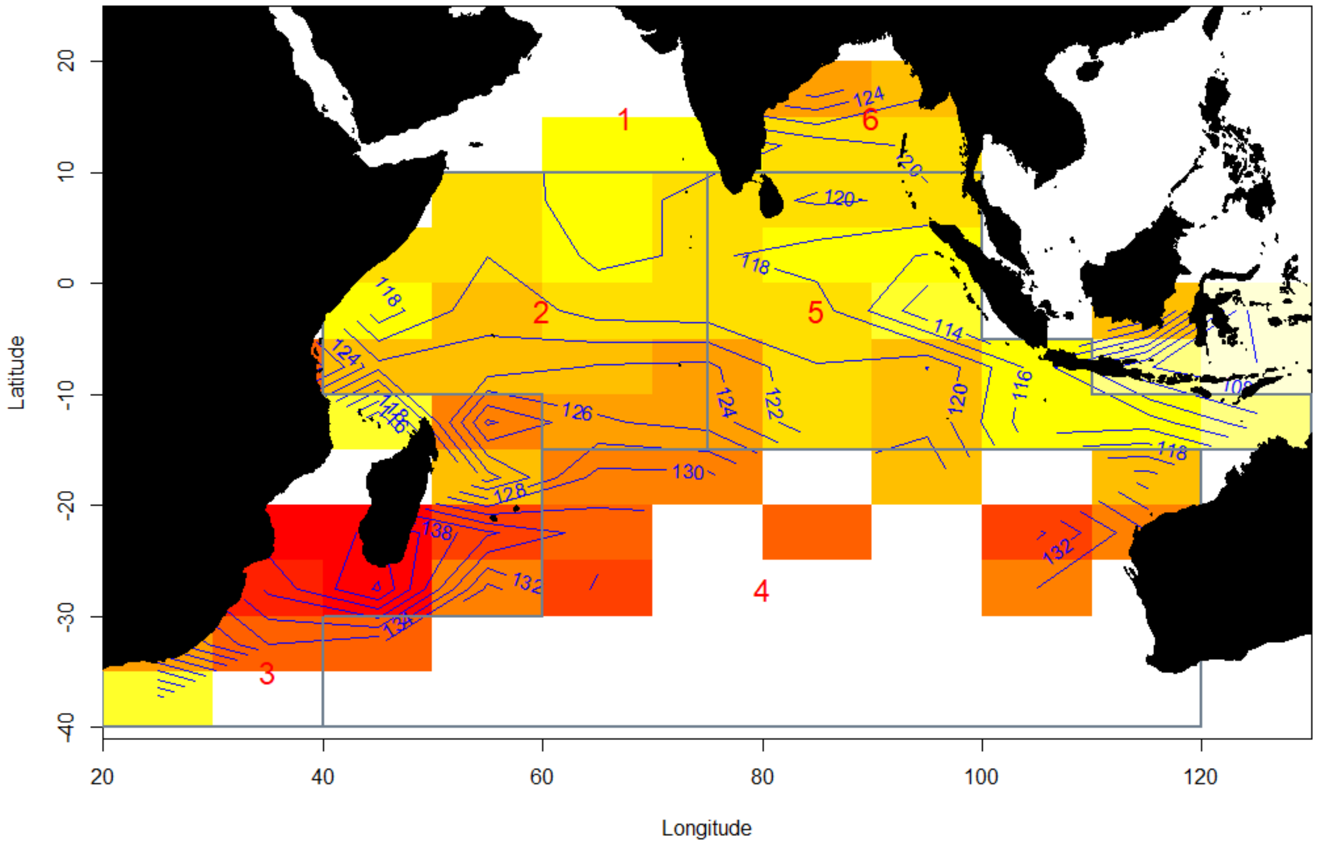


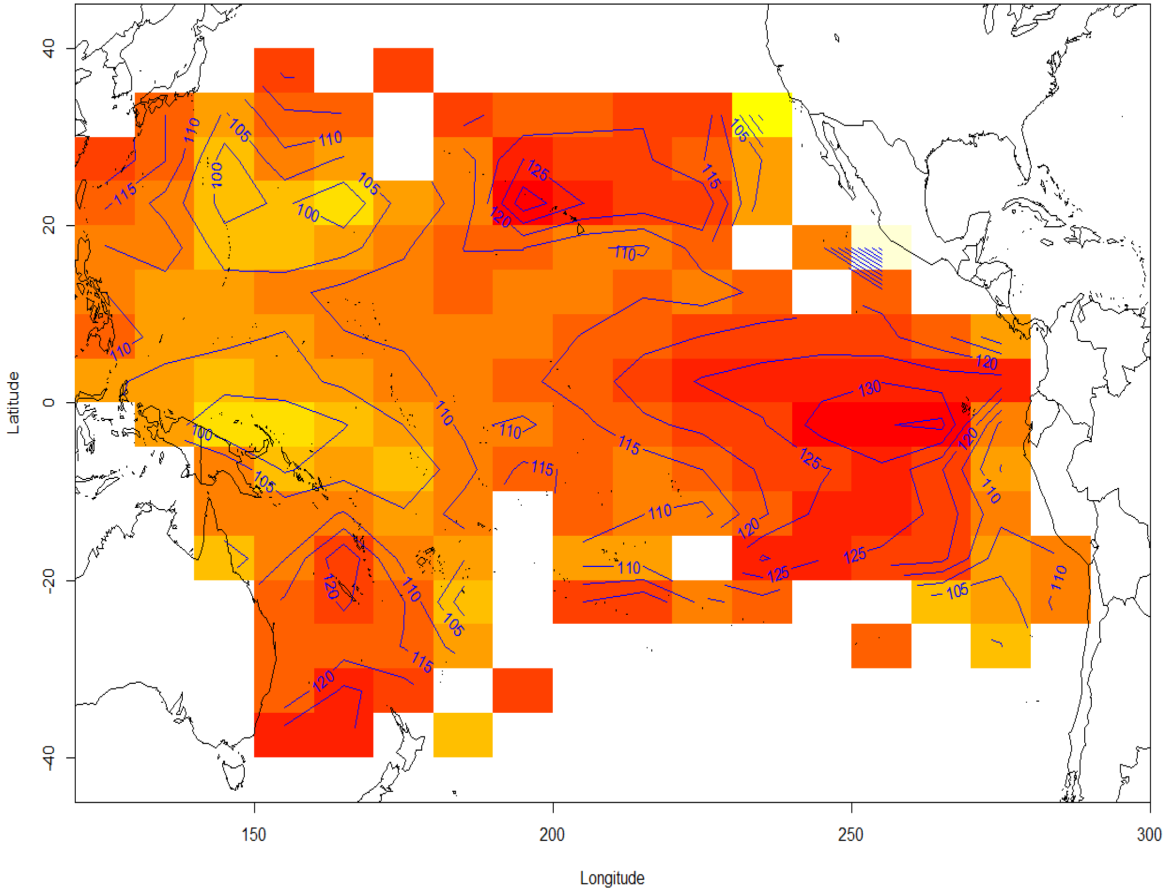




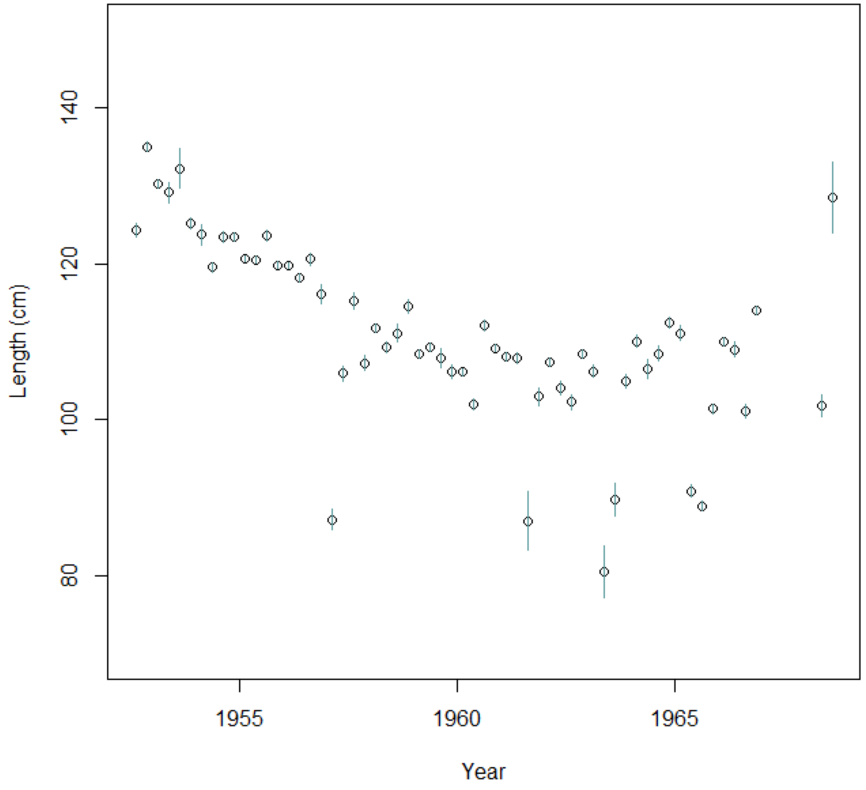


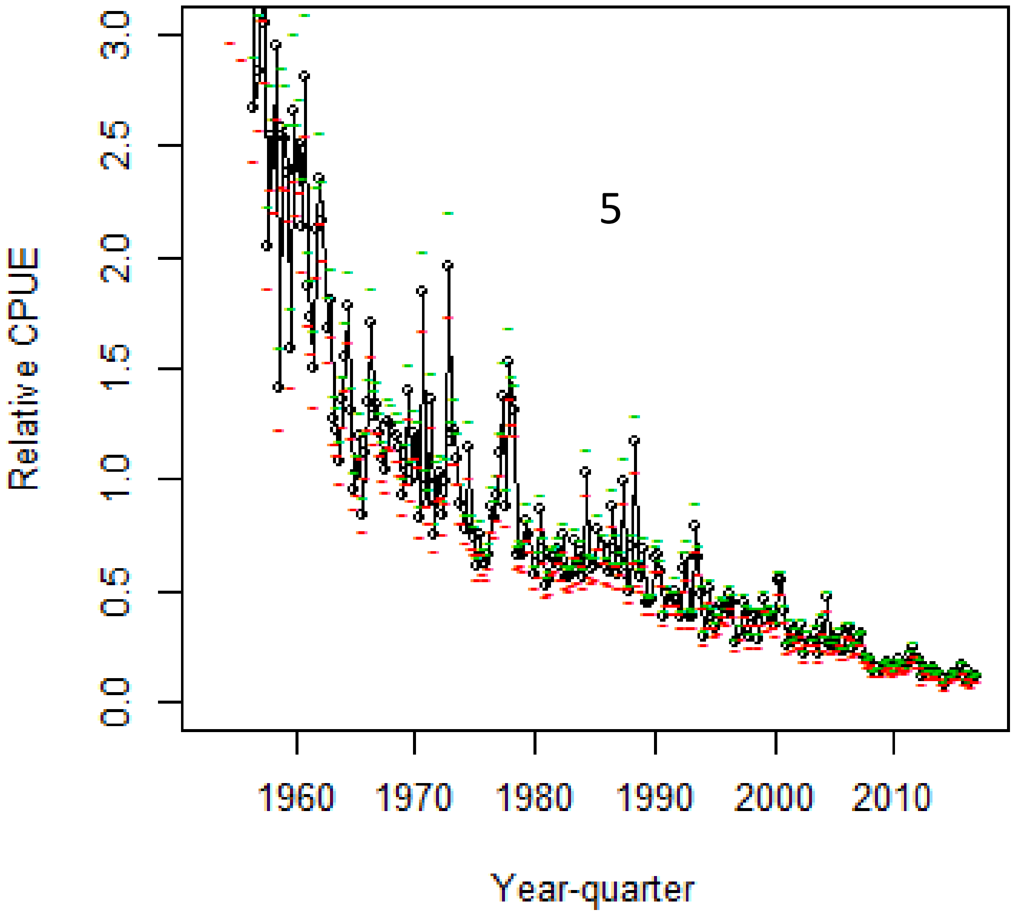
Other issues 3: size-area patterns





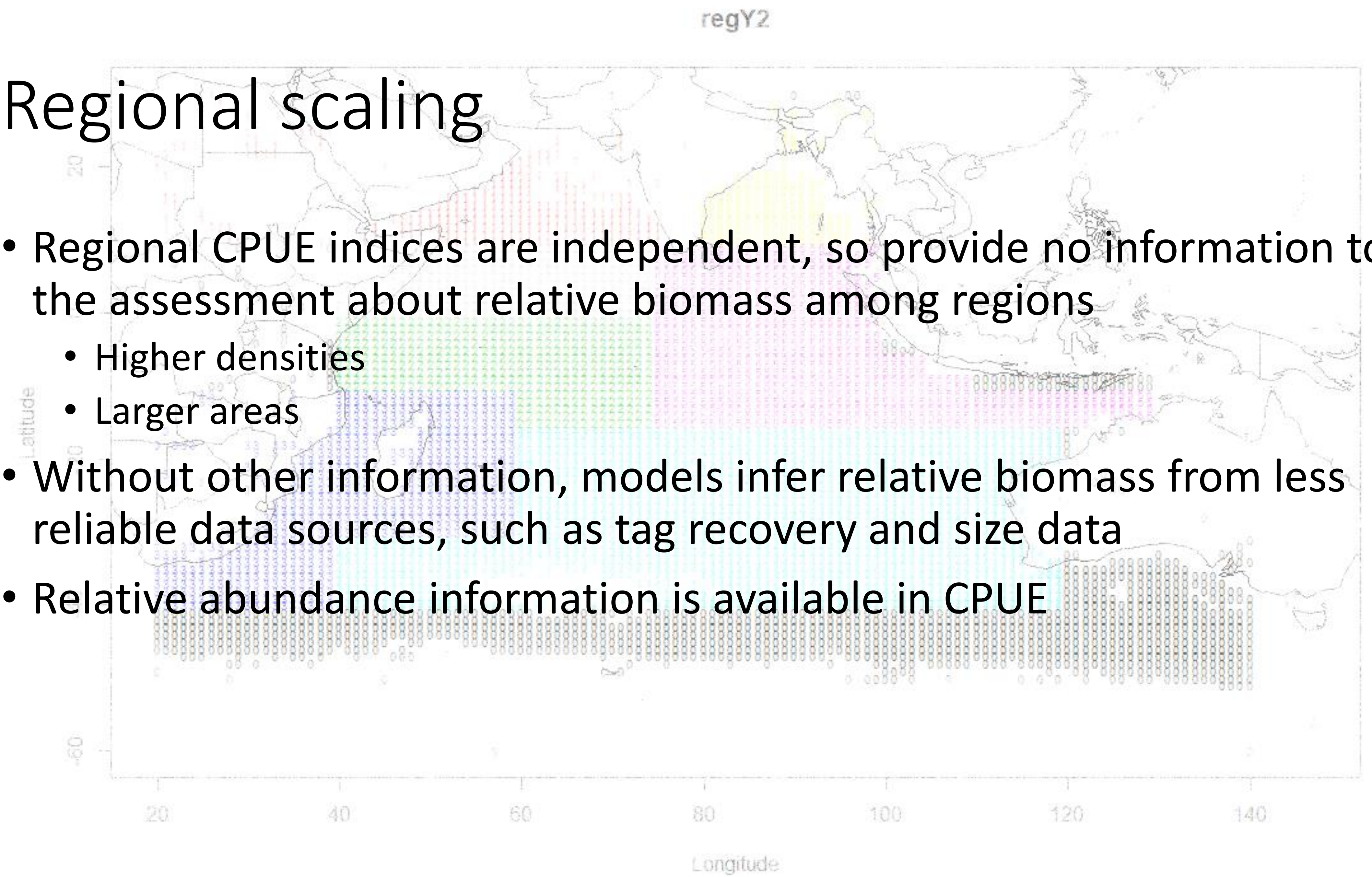
Other issues 4: probable catchability changes





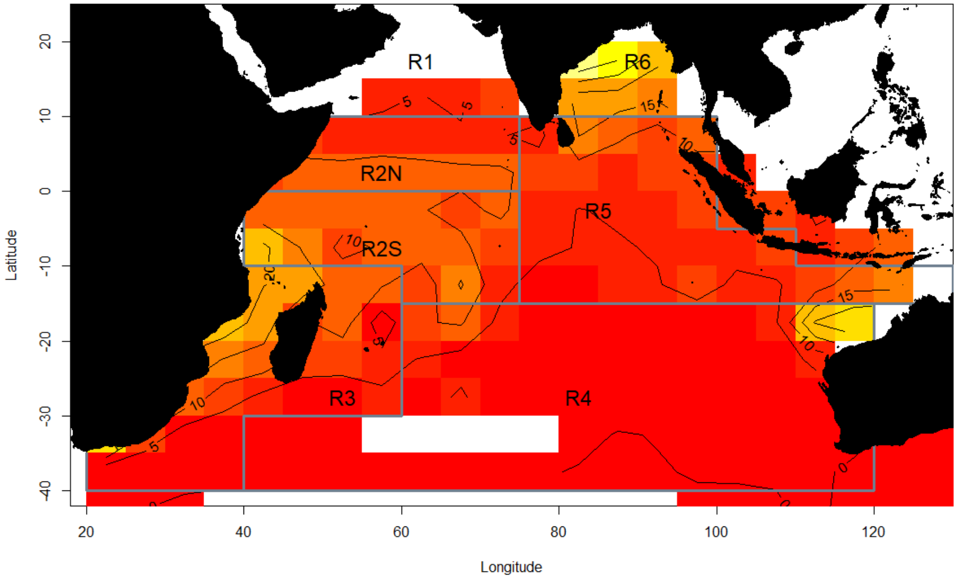
Regional scaling

- Regional CPUE indices are independent, so provide no information to the assessment about relative biomass among regions
  - Higher densities
  - Larger areas
- Without other information, models infer relative biomass from less reliable data sources, such as tag recovery and size data
- Relative abundance information is available in CPUE



Regional scaling: Adjusting for relative abundance among regions

- Use relative catch rates among regions as a proxy for density
- Abundance = sum(density x area)
- Area estimates are predicted 5° cell densities from a simple model, based on a shortish period with widespread fishing
- Currently using aggregated data
  - Need period with stable targeting
  - Simple model due to limited covariates
    - log(CPUE + constant) ~ cell + year-quarter



Conclusions

- Developed indices for the assessments
  - Plenty of room for more sophisticated spatio-temporal modelling
- Methods are significantly different from previous approaches used in IOTC assessments
  - Vessel effects, clustering esp. in temperate areas, use of operational data, area weighting, better data coverage due to multiple fleets, delta lognormal distribution
- Results comparable but some important differences
  - E.g. considerably higher YFT CPUE in recent years
- Residual analysis identified concerns
- Unanswered questions
  - Changes in late 70's, 2010, and the ability to target YFT vs BET
- Many opportunities to better understand fisheries, and improve the indices

Acknowledgments

- International Seafood Sustainability Foundation (ISSF) and the Indian Ocean Tuna Commission (IOTC) for funding the project.
- IOTC for facilitating the project, particularly Rishi Sharma, James Geehan, David Wilson, Dan Fu.
- Fishery Agencies of JP, TW, KR, and SY for granting access to their data.
- Taiwanese Overseas Fisheries Development Council, the Japanese National Research Institute of Far Seas Fisheries, and Shanghai Ocean University for providing their facilities and support.