Using pelagic fish movement data to estimate, predict and model CPUEs

José Miguel Ponciano: josemi@ufl.edu and Felipe Carvalho: fcorreia@ufl.edu

University of Florida, Biology Department

A novel approach to predict fish abundance using thermocline depth and animal movement data

- How can we use movement data to better understand and predict CPUE fluctuations of pelagic species?

A novel approach to predict fish abundance using thermocline depth and animal movement data

- How can we use movement data to better understand and predict CPUE fluctuations of pelagic species?
- The general context is the integration of satellite tag information into population dynamics modeling using simple stochastic processes, ecological and evolutionary principles.

Hierarchical models

- Reliable understanding of complex ecological data depends on the formulation of proper statistical models of the underlying processes.

Hierarchical models

- Reliable understanding of complex ecological data depends on the formulation of proper statistical models of the underlying processes.
- Hierarchical statistical models have proved useful towards achieving such goal.

Hierarchical models

- Reliable understanding of complex ecological data depends on the formulation of proper statistical models of the underlying processes.
- Hierarchical statistical models have proved useful towards achieving such goal.
- These models incorporate variability in parameters that otherwise is treated as fixed and

Hierarchical models

- Reliable understanding of complex ecological data depends on the formulation of proper statistical models of the underlying processes.
- Hierarchical statistical models have proved useful towards achieving such goal.
- These models incorporate variability in parameters that otherwise is treated as fixed and
- they incorporate multiple layers of uncertainty.

Hierarchical models in Ecology

$$\begin{aligned}\mathbf{Y} &\sim f(\mathbf{y}|\mathbf{X},\phi)\\ \mathbf{X} &\sim g(\mathbf{x}|\theta)\end{aligned}$$

it is known that the likelihood is

$$L(\theta,\phi)=\int f(\mathbf{y}|\mathbf{X},\phi)g(\mathbf{x}|\theta)d\mathbf{X}.$$

A few examples include:

- Stochastic population models with added observation error (De Valpine and Hastings 2002, Clark and Bjornstad 2004, Newman et al. 2006, Dennis et al 2006)
- Stochastic models of species abundance distributions (Etienne and Olff 2005)
- Capture-recapture models with uncertain capture probabilities (George and Robert 1992)

Non-linear, non-Gaussian SSM

$$\begin{aligned}\mathbf{Y} &\sim f(\mathbf{y}|\mathbf{X},\phi)\\ \mathbf{X} &\sim g(\mathbf{x}|\theta)\end{aligned}$$

it is known that the likelihood is

$$L(\theta,\phi) = \int f(\mathbf{y}|\mathbf{X},\phi)g(\mathbf{x}|\theta)d\mathbf{X}.$$

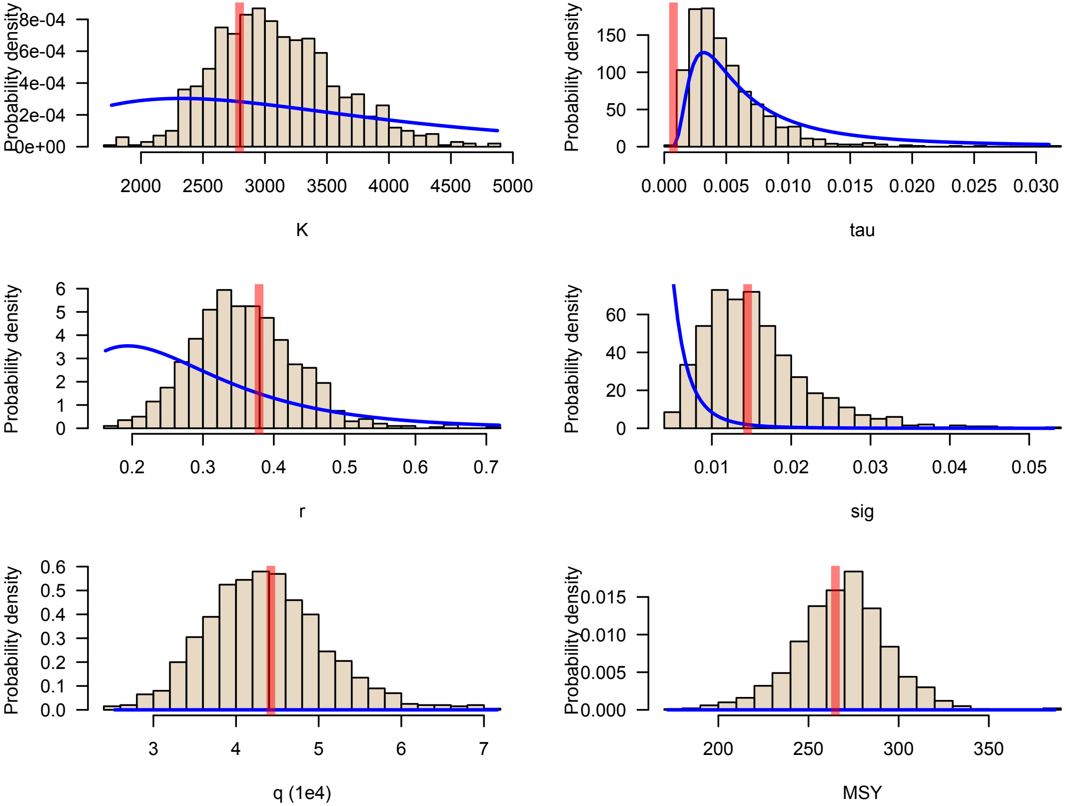
- Maximum likelihood was known to be very difficult for these models.
- Bayesian solutions to the study of hierarchical population models were much easier to implement until recently.
- However, it can be very difficult to specify non-informative priors to do “objective bayesian statistics” for hierarchical models (Nancy Reid, 2008):
  - Bayesian hierarchical Poisson models, (Gelman et al 2007)
  - Heinrich 2005, Proceedings of Phystat05 ($\mathrm{Poisson}\,(\epsilon s + b)$, s of interest, additional Poisson measurements of b and ϵ)
  - Bayesian probit regression (Jones 2008, Siddhartha and Chib 1984)

A persistent problem: non-Identifiability

Non-identifiability of a parameter: the value of the likelihood function evaluated at the data is unchanged when two different sets of parameters are used (Rannala 2002, Lele et al 2010). But need to consider 3 cases

- Model Non-Identifiability (MNI): model written in such a way that two or more parameters are non-separable.
  - Can inadvertently be introduced while formulating a Hidden Markov model or a state space model (McCullogh and Searle 2001, Yang and Rannala 2006)
- Sampling Non-Identifiability (SNI): by pure chance, sample contains 0 information about the parameter of interest.
  - In phylogenetics: all sampled loci contain no mutations in a section of the topology where a real branching event is present.
- Weak Estimability (WE): Data simply does not contain enough information to estimate the parameters of interest and the profile likelihood of certain parameters lacks a strong curvature.

An example in Fisheries, Meyer and Millar 1999



Are the model parameters identifiable?

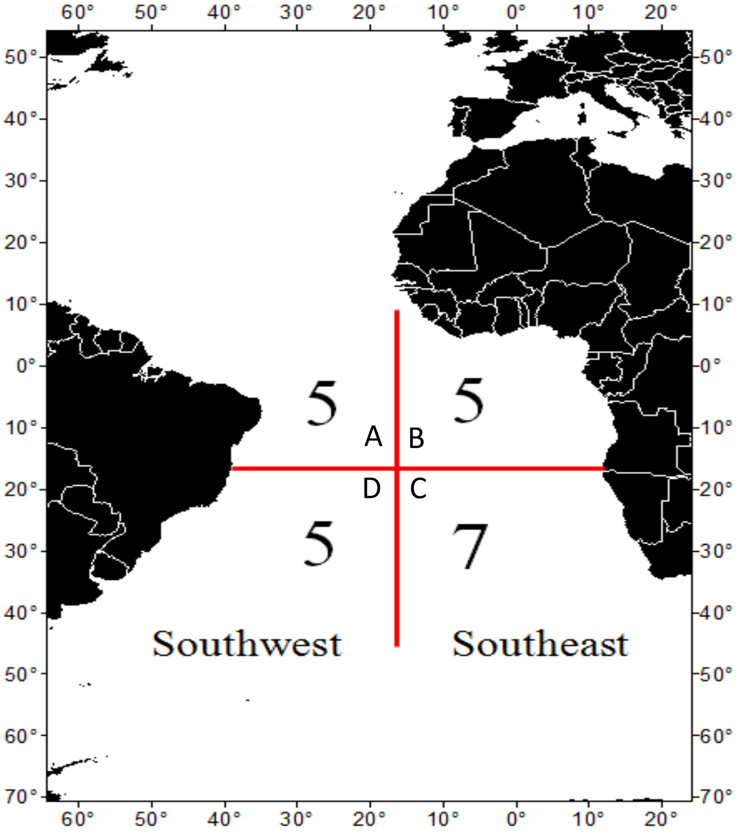
- Identifiability of model parameters is indeed very important!
- Accumulating evidence keeps telling us about the need to integrate different, independent levels of information to carry reliable statistical inference
- MCMC won't do miracles for us!!

- Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, and D. F. Staples. (2006) Estimating density dependence, process noise, and observation error. Ecological Monographs
- Lele S., Nadeem K., Schmuland B.(2010). Estimability and likelihood inference for generalized linear mixed models using data cloning. JASA
- Dennis, B., Ponciano, J.M. and M.L. Taper (2010). Replicated sampling increases efficiency in monitoring biological populations. Ecology.
- Lebreton, J.-D. and O. Gimenez (2012). Detecting and estimating density-dependence in wildlife populations. Journal of Wildlife Management.
- Ponciano, J.M., Burleigh, G., Braun, E. and M. L. Taper (2012) Assessing Parameter Identifiability in Phylogenetic Models Using Data Cloning. Systematic Biology.

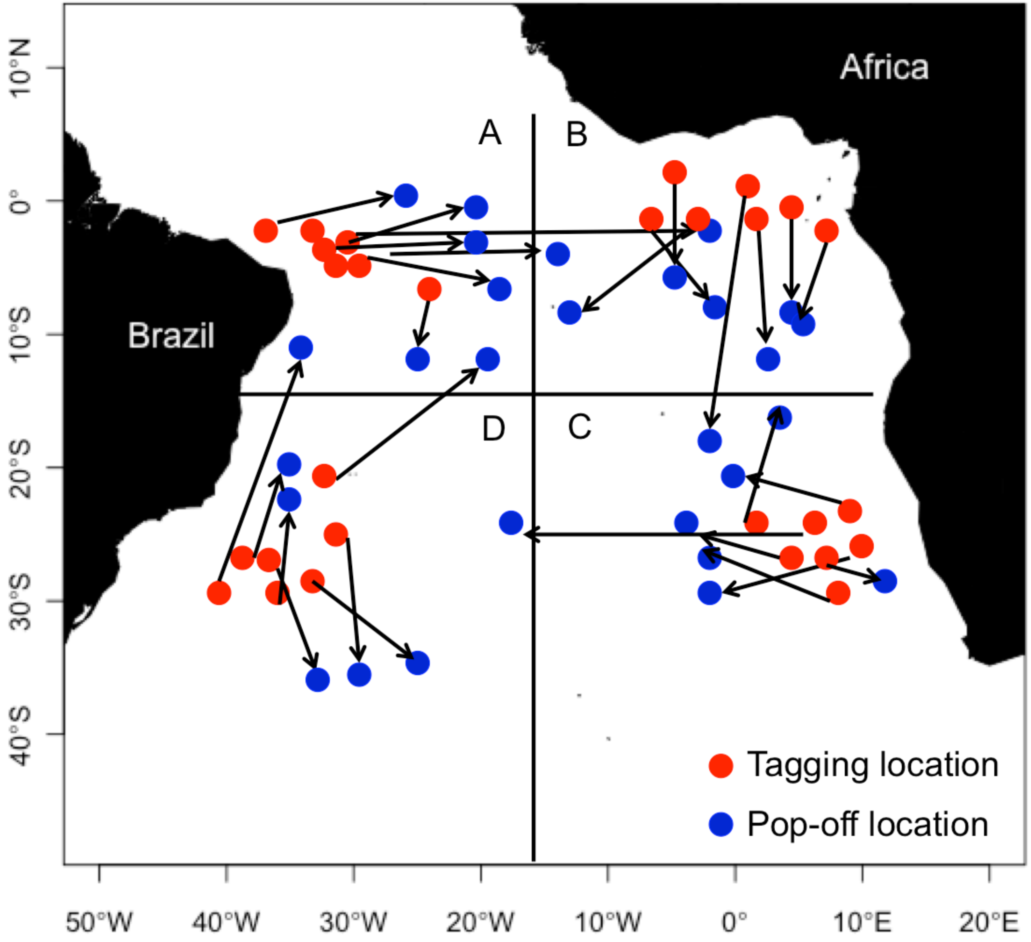
This talk: it’s not about arguing about a statistical problem. It’s about offering potential solutions

- Blue Shark Satellite data: A large project involving many institutions -including MADE- across different countries
  - Charlene Da Silva, responsible for part of the movement data from South Africa
  - Fabio Hazin and Paulo Travassos, Recife, Brazil: responsible for the research design, implementation and funds allocation for the tags.
  - Mariana Travassos, data analysis and logistics of the project in Brazil
- Sharks tagged with a satellite device: 10 males, 18 females, mature and immature, up to 180 days of tracking, across 4 areas of the Atlantic Ocean

Study area


5
5
A
B
D
C
5
7
Southwest
Southeast
60°
50°
40°
30°
20°
10°
0°
10°
20°
50°
40°
30°
20°
10°
0°
10°
20°
30°
40°
50°
60°
70°

Tag and Pop-off locations: Brazil and South-Africa



Subset of data analyzed in this study

From

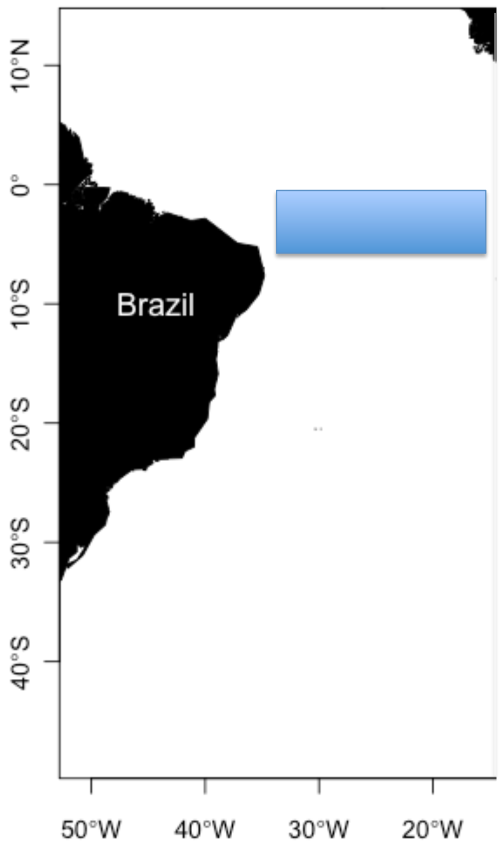
- 19 quadrants of 111 by 111 nautical miles.
- 4 female, tagged blue shark.

For each quadrant, we have available time series of length 16 time units t of

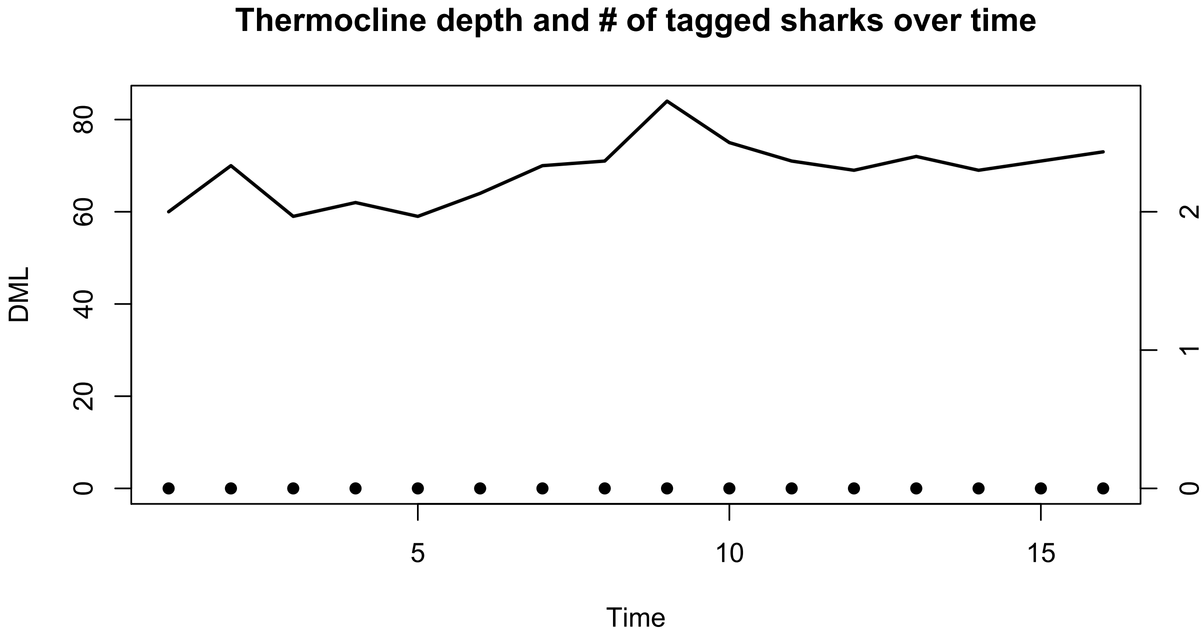
- The number of tagged sharks present in each quadrant at time t.
- DML values during time $(t, t+1)$
- CPUEs in the same time interval.

Finally, each time interval consisted of 3 days (for a total of 48 days). Interval chosen matches the minimum interval size for which presence/absence, DMLs and CPUE values are all simultaneously available.

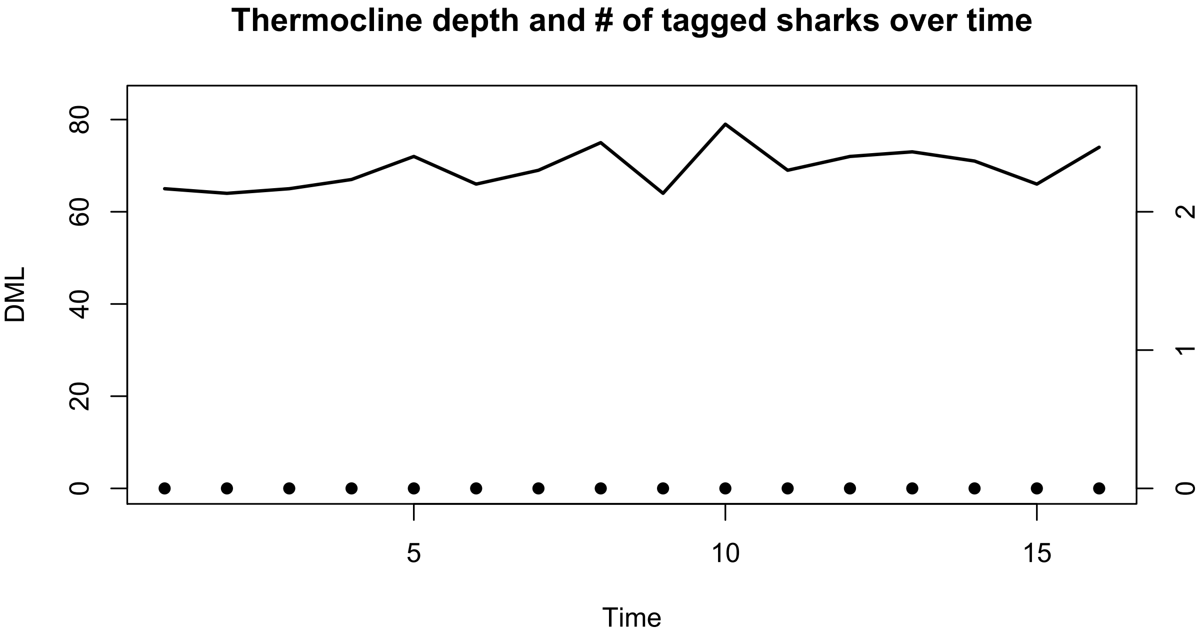
Location of subset of data



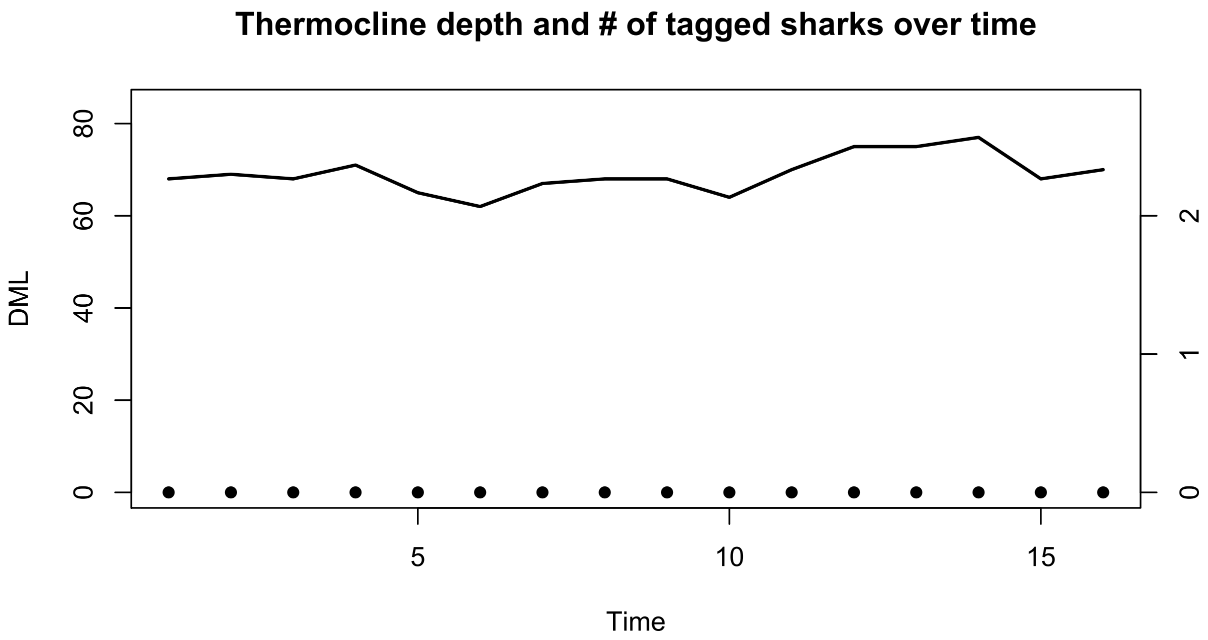
Time series of thermocline data and number of tagged sharks present at time t, quadrant 1



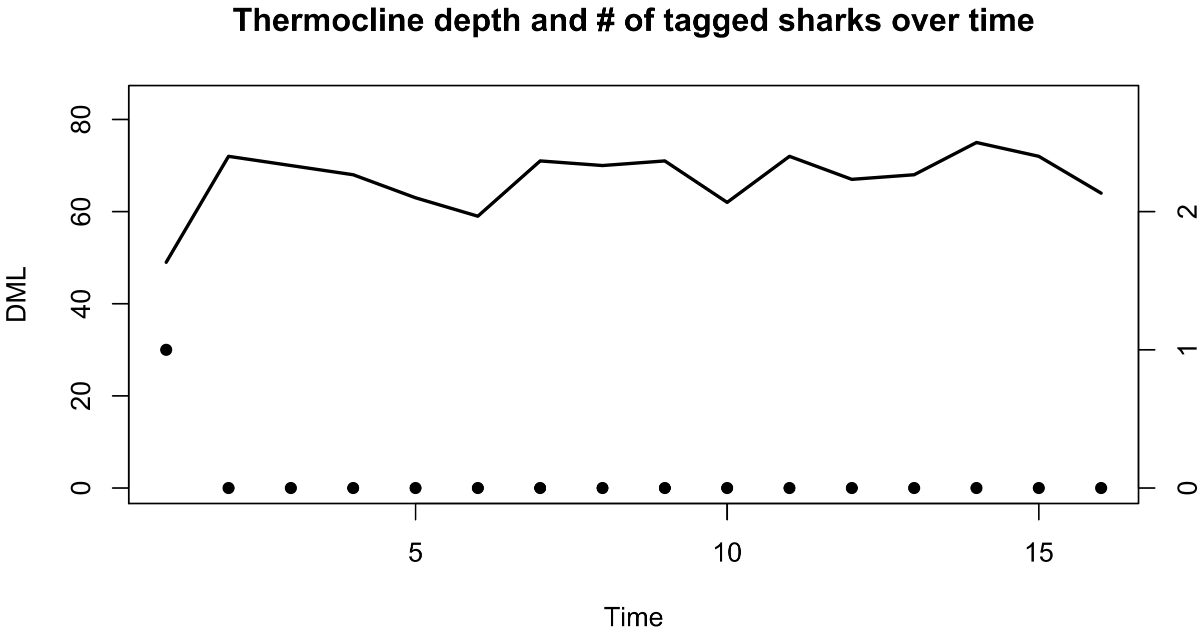
Time series of thermocline data and number of tagged sharks present at time t, quadrant 2



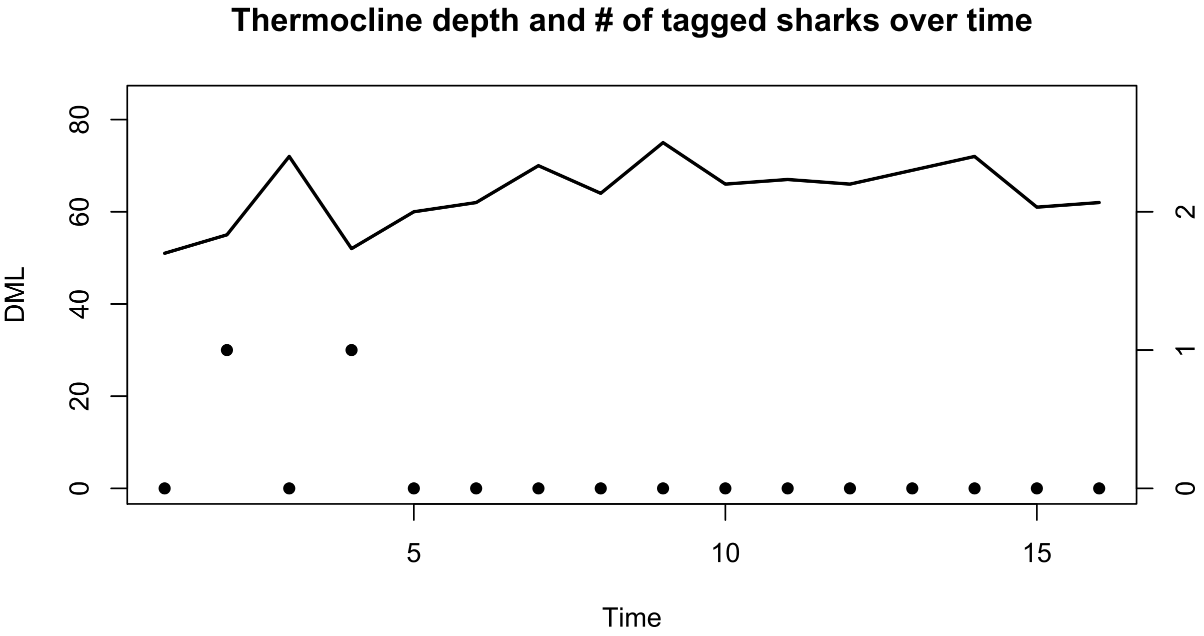
Time series of thermocline data and number of tagged sharks present at time t, quadrant 3



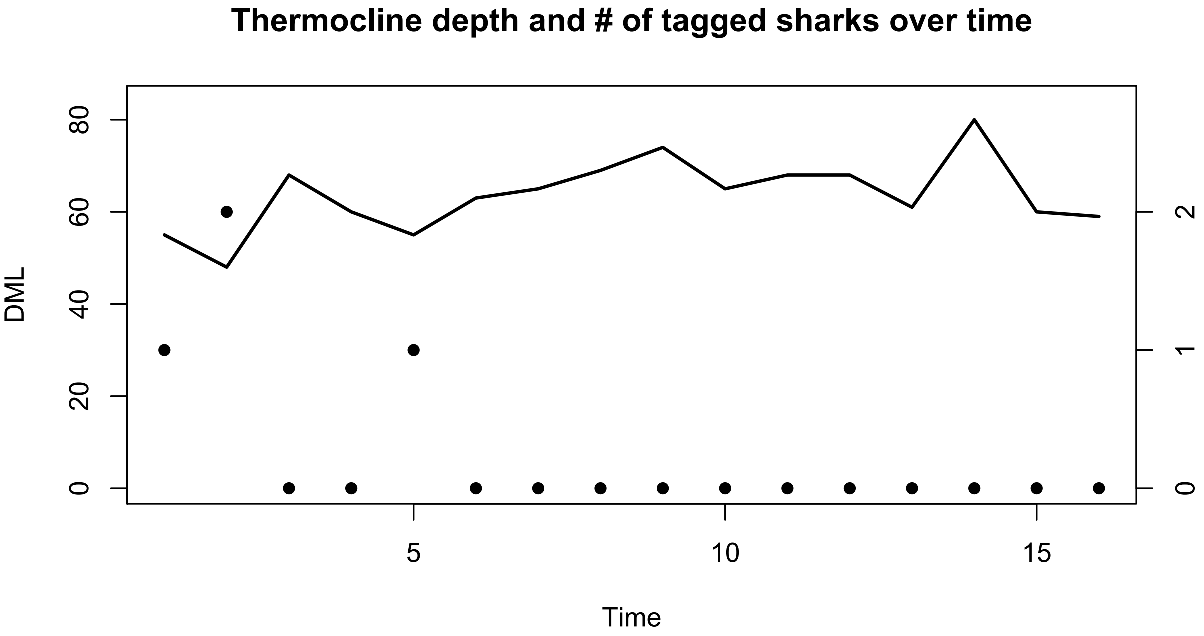
Time series of thermocline data and number of tagged sharks present at time t, quadrant 4



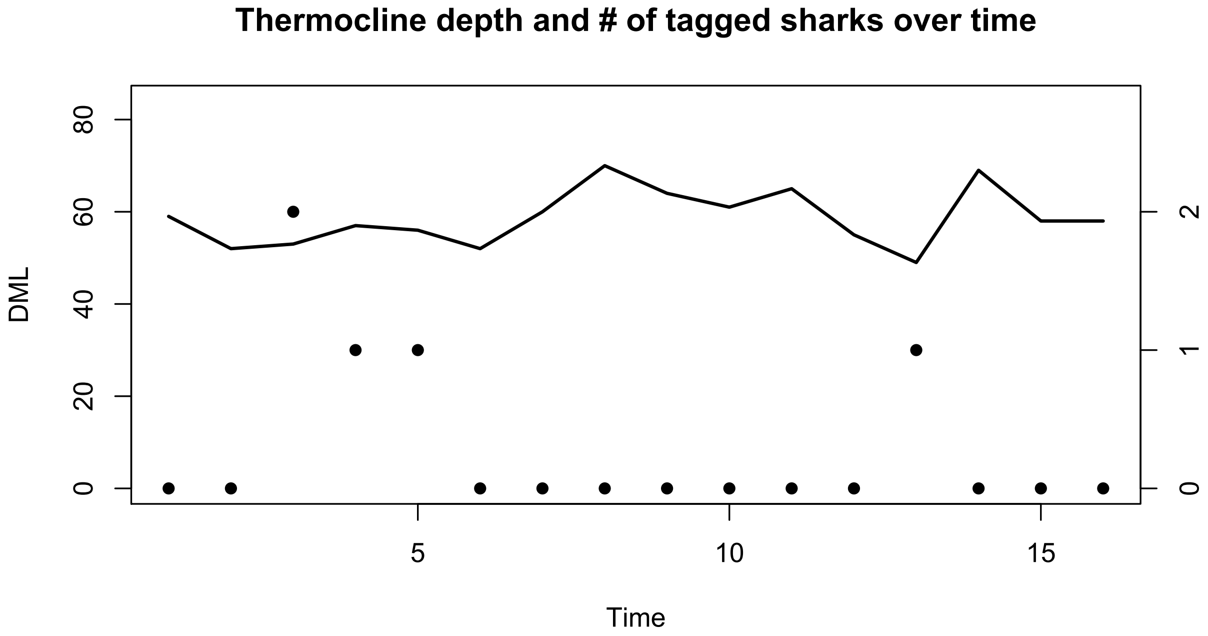
Time series of thermocline data and number of tagged sharks present at time t, quadrant 5



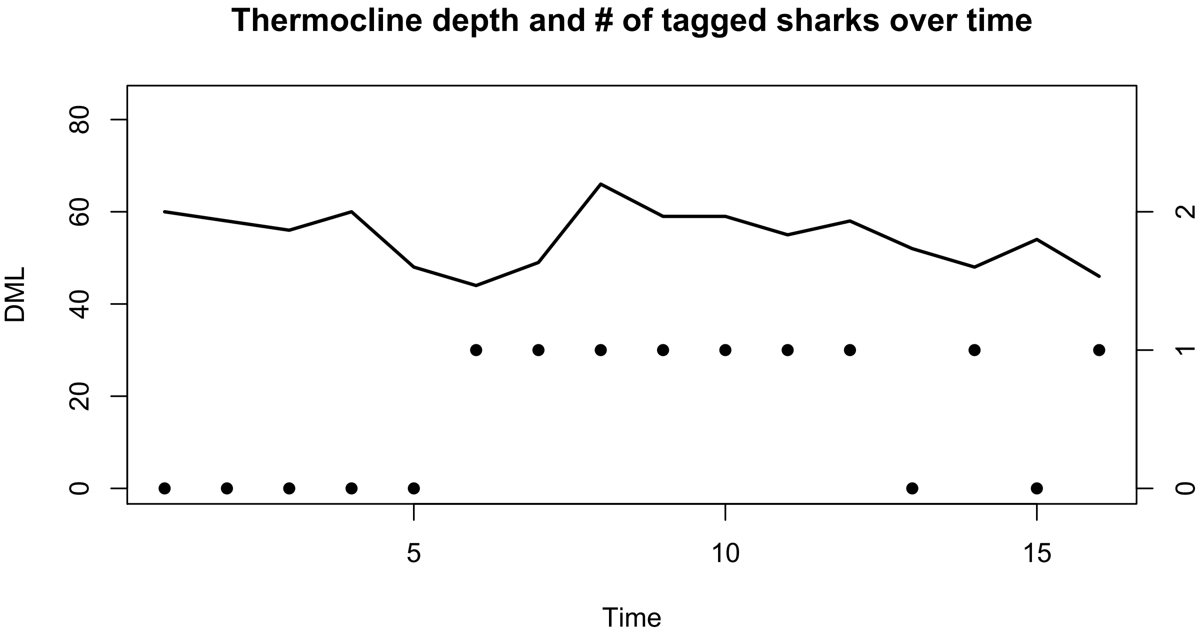
Time series of thermocline data and number of tagged sharks present at time t, quadrant 6



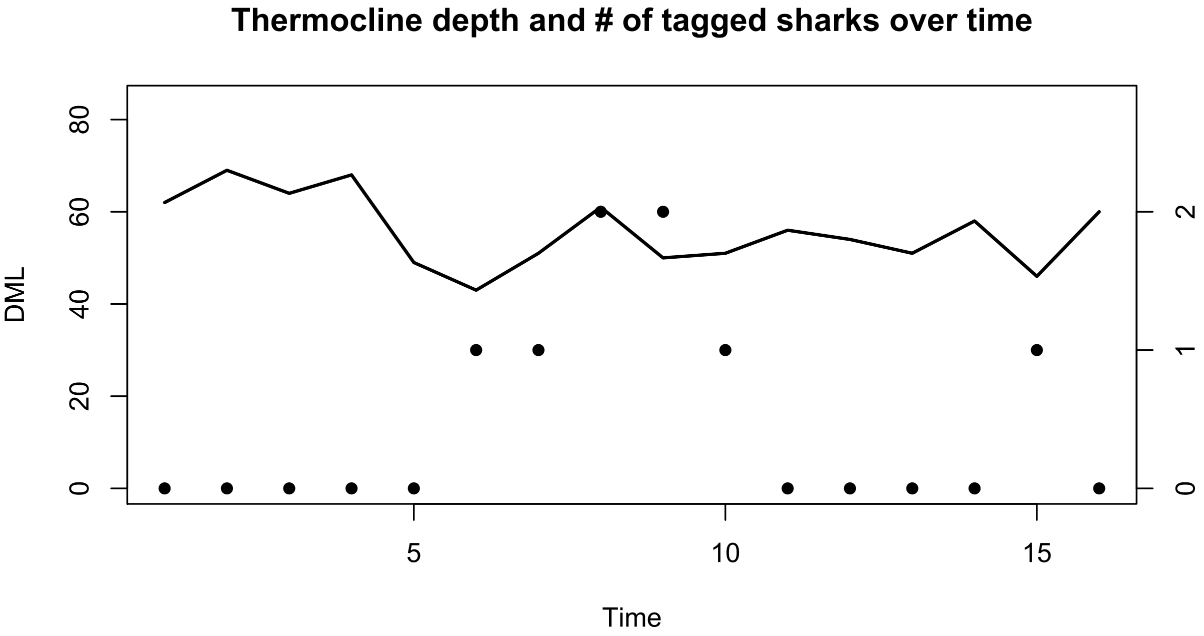
Time series of thermocline data and number of tagged sharks present at time t, quadrant 7



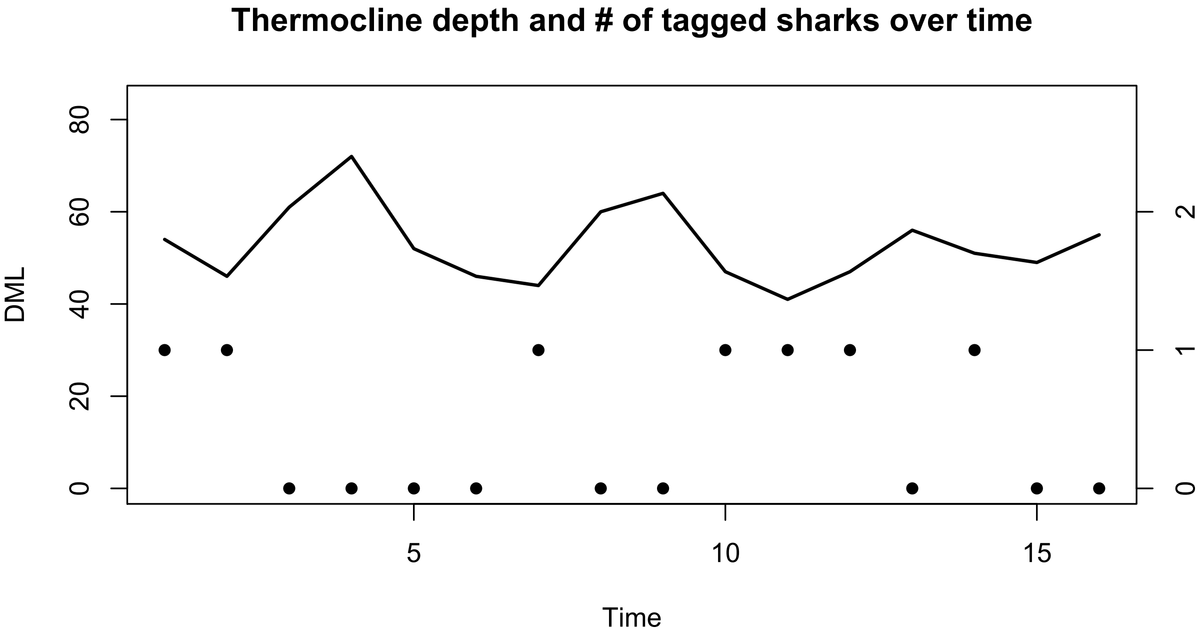
Time series of thermocline data and number of tagged sharks present at time t, quadrant 8



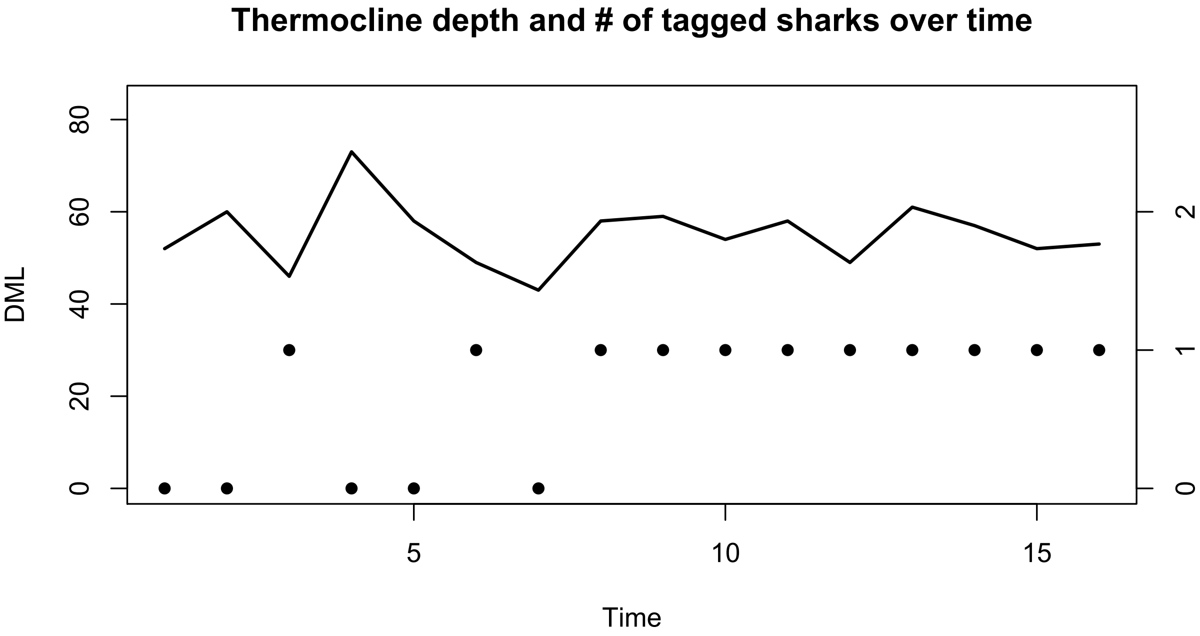
Time series of thermocline data and number of tagged sharks present at time t, quadrant 9



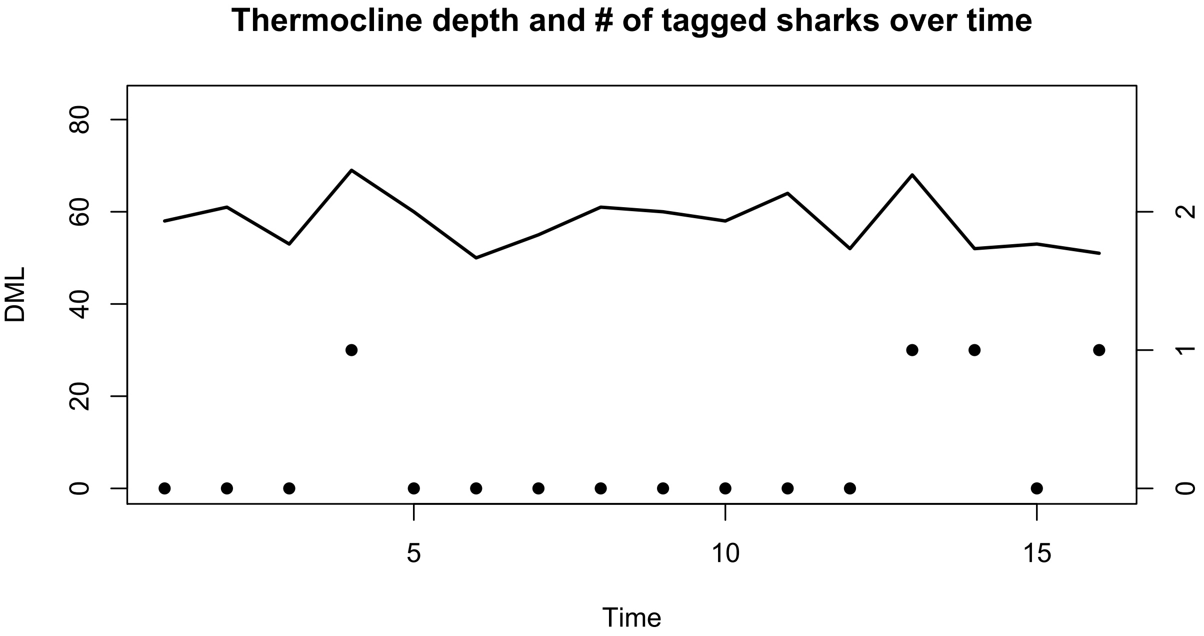
Time series of thermocline data and number of tagged sharks present at time t, quadrant 10



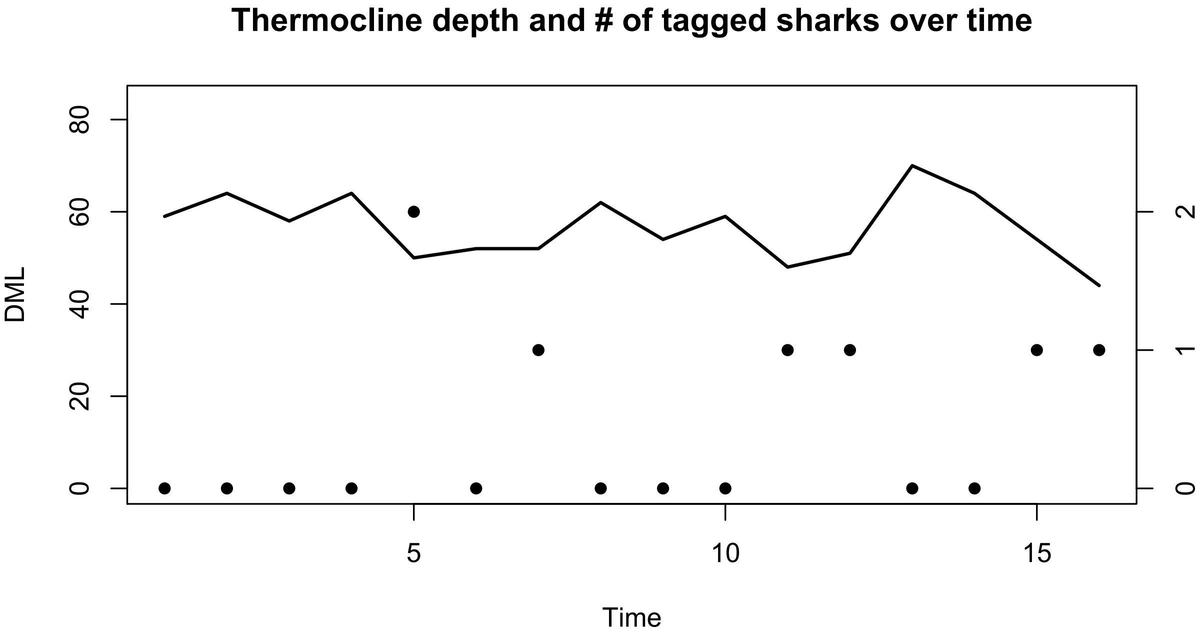
Time series of thermocline data and number of tagged sharks present at time t, quadrant 11



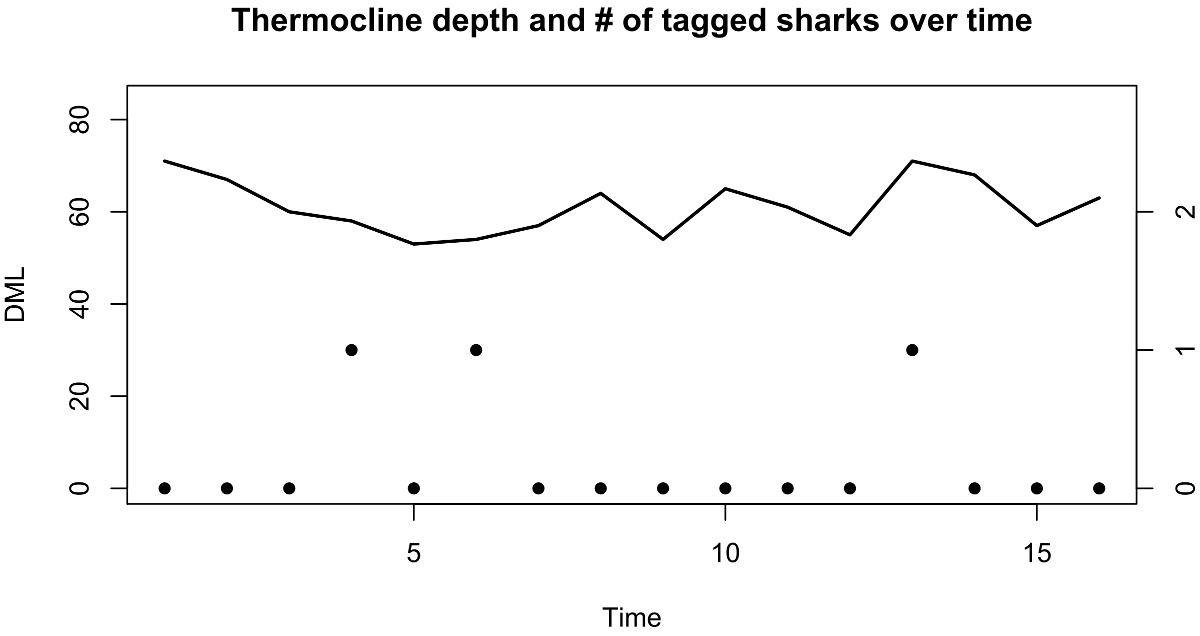
Time series of thermocline data and number of tagged sharks present at time t, quadrant 12



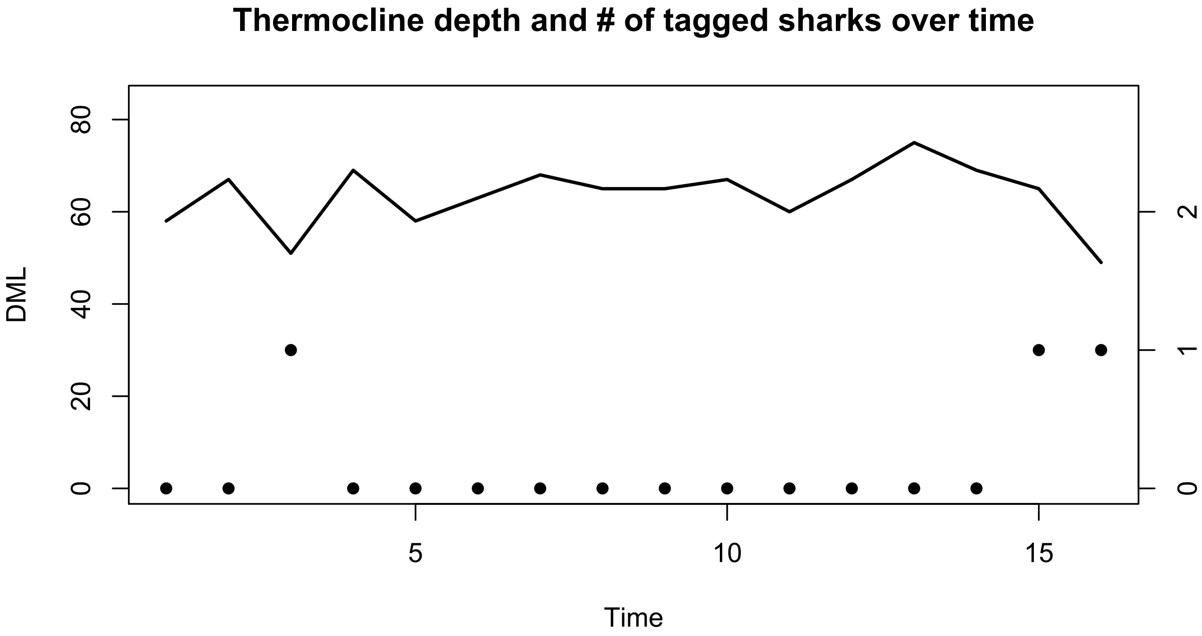
Time series of thermocline data and number of tagged sharks present at time t, quadrant 13



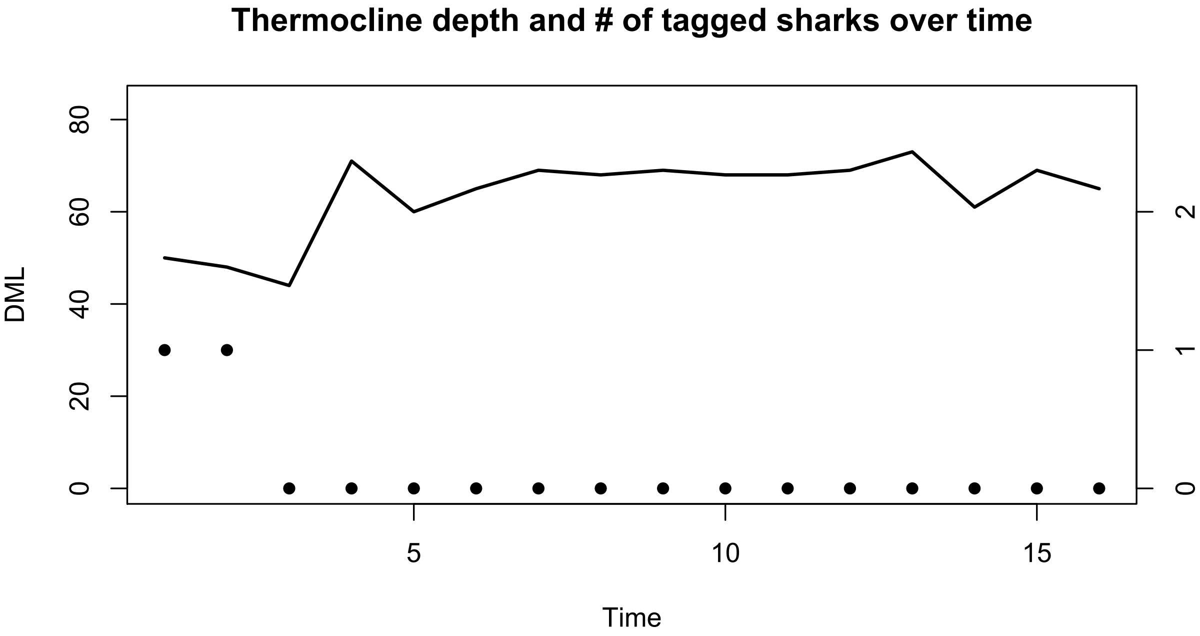
Time series of thermocline data and number of tagged sharks present at time t, quadrant 14



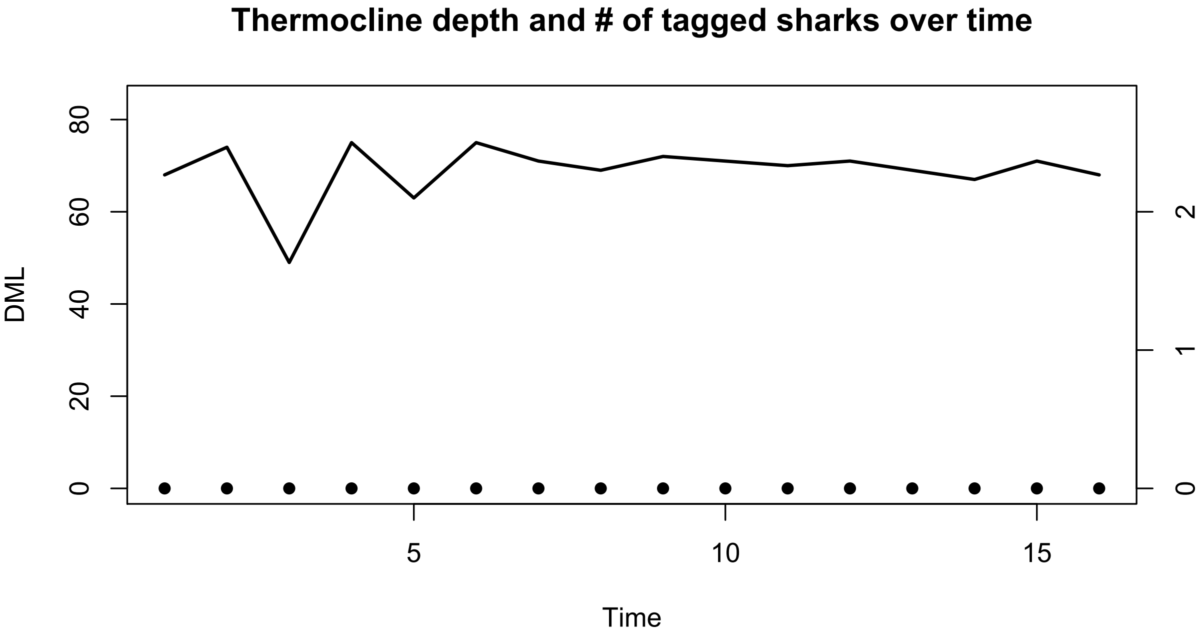
Time series of thermocline data and number of tagged sharks present at time t, quadrant 15



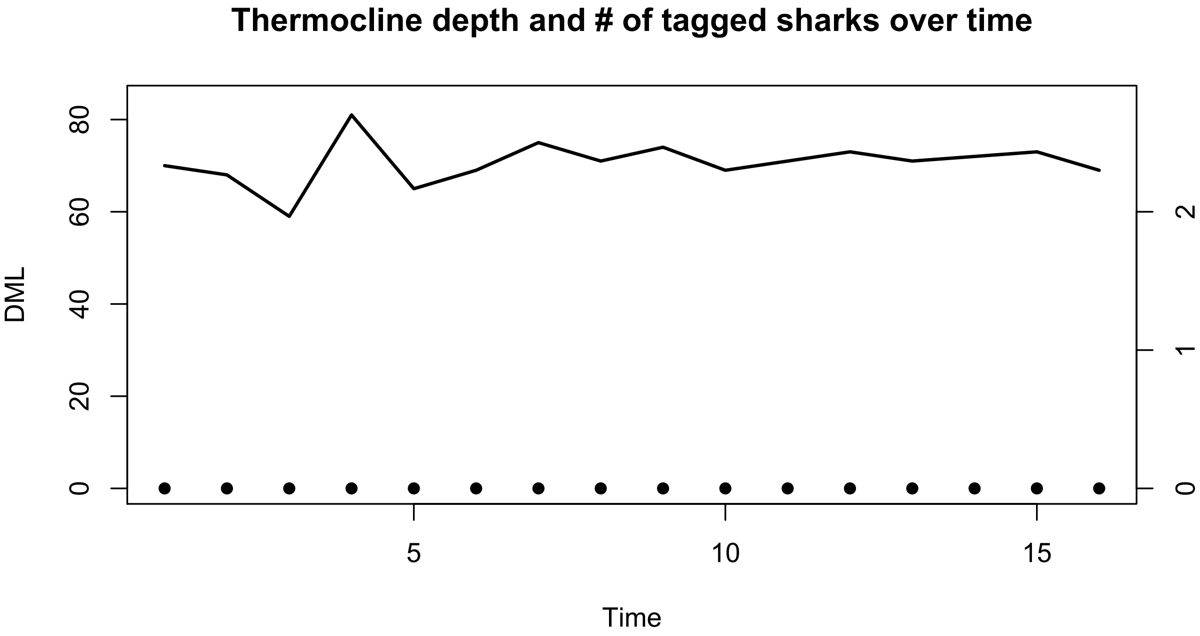
Time series of thermocline data and number of tagged sharks present at time t, quadrant 16



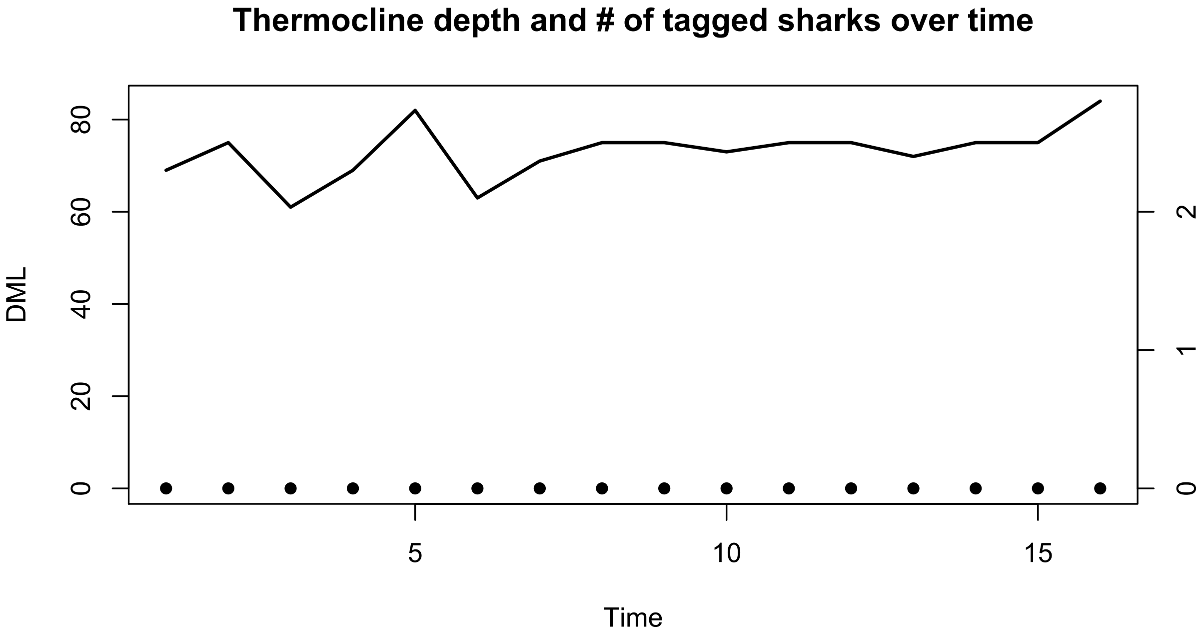
Time series of thermocline data and number of tagged sharks present at time t, quadrant 17



Time series of thermocline data and number of tagged sharks present at time t, quadrant 18



Time series of thermocline data and number of tagged sharks present at time t, quadrant 19



A simple probabilistic setting: Discrete, self-decomposable processes:

Within a single quadrant, the time series of arrivals and departures of individual, tagged sharks can be modeled according to:

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} \quad \text{where}$$

A simple probabilistic setting: Discrete, self-decomposable processes:

Within a single quadrant, the time series of arrivals and departures of individual, tagged sharks can be modeled according to:

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} \quad \text{where}$$

- $\alpha = Pr($an individual shark stays from time $t-1$ to time $t)$, $\alpha \star X_{t-1} = \sum_{i=1}^{X_{t-1}} B_i(\alpha)$ and

$$B_i(\alpha) = \begin{cases} 1 & \text{w. prob.} \quad \alpha \\ 0 & \text{w. prob.} \quad 1-\alpha \end{cases}$$

A simple probabilistic setting: Discrete, self-decomposable processes:

Within a single quadrant, the time series of arrivals and departures of individual, tagged sharks can be modeled according to:

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} \quad \text{where}$$

- $\alpha = Pr(\text{an individual shark stays from time } t-1 \text{ to time } t)$, $\alpha \star X_{t-1} = \sum_{i=1}^{X_{t-1}} B_i(\alpha)$ and

$$B_i(\alpha) = \begin{cases} 1 & \text{w. prob.} \quad \alpha \\ 0 & \text{w. prob.} \quad 1-\alpha \end{cases}$$

- If I specify a probabilistic model for X_{t-1} (say $X_{t-1} \sim \mathrm{Pois}(\theta)$), then I can find the distribution of $\alpha \star X_{t-1}$ ($\mathrm{Pois}(\alpha\theta)$ in this case).

A simple probabilistic setting: Discrete, self-decomposable processes:

Within a single quadrant, the time series of arrivals and departures of individual, tagged sharks can be modeled according to:

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} \quad \text{where}$$

- $\alpha = Pr($an individual shark stays from time $t-1$ to time $t)$, $\alpha \star X_{t-1} = \sum_{i=1}^{X_{t-1}} B_i(\alpha)$ and

$$B_i(\alpha) = \begin{cases} 1 & \text{w. prob.} \quad \alpha \\ 0 & \text{w. prob.} \quad 1-\alpha \end{cases}$$

- If I specify a probabilistic model for X_{t-1} (say $X_{t-1} \sim \mathrm{Pois}(\theta)$), then I can find the distribution of $\alpha \star X_{t-1}$ ($\mathrm{Pois}(\alpha\theta)$ in this case).
- And if I specify a model for the arrivals (say $\epsilon_t \sim \mathrm{Pois}((1-\alpha)\theta)$), then I can find the exact distribution of the process ($X_t \sim \mathrm{Pois}(\theta)$ in this case) and
- the *conditional transition distribution of the process*, $P(X_t = x_t | X_{t-1} = x_{t-1})$

Connecting the observed time series of sharks presence/absence with the model

For a given time series of observations $x_0, x_1, \ldots, x_q$, maximizing the joint probability of the observations, given the model and data at hand *i.e* -the likelihood function-

$$P(X_0 = x_0, X_1 = x_1 \ldots, X_q = x_q) = P(X_0 = x_0) \prod_{i=1}^{q} P(X_i = x_i | X_{i-1} = x_{i-1})$$

Connecting the observed time series of sharks presence/absence with the model

For a given time series of observations $x_0, x_1, \ldots, x_q$, maximizing the joint probability of the observations, given the model and data at hand *i.e* -the likelihood function-

$$P(X_0 = x_0, X_1 = x_1 \ldots, X_q = x_q) = P(X_0 = x_0) \prod_{i=1}^{q} P(X_i = x_i | X_{i-1} = x_{i-1})$$

yields the Maximum Likelihood estimates of the model parameters α and θ.

Connecting the observed time series of sharks presence/absence with the model

For a given time series of observations $x_0, x_1, \dots, x_q$, maximizing the joint probability of the observations, given the model and data at hand *i.e* -the likelihood function-

$$P(X_0 = x_0, X_1 = x_1 \dots, X_q = x_q) = P(X_0 = x_0) \prod_{i=1}^{q} P(X_i = x_i | X_{i-1} = x_{i-1})$$

yields the Maximum Likelihood estimates of the model parameters α and θ. Note that the simple Poisson models mentioned above

- assume homogeneity and independence across quadrants (no spatial dependence),

Connecting the observed time series of sharks presence/absence with the model

For a given time series of observations $x_0, x_1, \ldots, x_q$, maximizing the joint probability of the observations, given the model and data at hand *i.e* -the likelihood function-

$$P(X_0 = x_0, X_1 = x_1 \ldots, X_q = x_q) = P(X_0 = x_0) \prod_{i=1}^{q} P(X_i = x_i | X_{i-1} = x_{i-1})$$

yields the Maximum Likelihood estimates of the model parameters α and θ. Note that the simple Poisson models mentioned above

- assume homogeneity and independence across quadrants (no spatial dependence),
- cannot account for all these 0's I showed you before,

Connecting the observed time series of sharks presence/absence with the model

For a given time series of observations $x_0, x_1, \ldots, x_q$, maximizing the joint probability of the observations, given the model and data at hand *i.e* -the likelihood function-

$$P(X_0 = x_0, X_1 = x_1 \ldots, X_q = x_q) = P(X_0 = x_0) \prod_{i=1}^{q} P(X_i = x_i | X_{i-1} = x_{i-1})$$

yields the Maximum Likelihood estimates of the model parameters α and θ. Note that the simple Poisson models mentioned above

- assume homogeneity and independence across quadrants (no spatial dependence),
- cannot account for all these 0's I showed you before,
- do not model the probability of staying or arriving to a particular location as a function of the environmental conditions (DML's, for instance).

Connecting the observed time series of sharks presence/absence with the model

For a given time series of observations $x_0, x_1, \ldots, x_q$, maximizing the joint probability of the observations, given the model and data at hand *i.e* -the likelihood function-

$$P(X_0 = x_0, X_1 = x_1 \ldots, X_q = x_q) = P(X_0 = x_0) \prod_{i=1}^{q} P(X_i = x_i | X_{i-1} = x_{i-1})$$

yields the Maximum Likelihood estimates of the model parameters α and θ. Note that the simple Poisson models mentioned above

- assume homogeneity and independence across quadrants (no spatial dependence),
- cannot account for all these 0's I showed you before,
- do not model the probability of staying or arriving to a particular location as a function of the environmental conditions (DML's, for instance).
- does not connect movement to CPUEs.

Accounting for heterogeneity across quadrants

General idea: using the derivation of the Negative Binomial as a conditional Poisson distribution where the rate is itself a random variable (gamma distributed) (Corbert, Fisher and Williams, 1943), and averaged over all possible rates.

Accounting for heterogeneity across quadrants

General idea: using the derivation of the Negative Binomial as a conditional Poisson distribution where the rate is itself a random variable (gamma distributed) (Corbert, Fisher and Williams, 1943), and averaged over all possible rates.

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} .$$

Accounting for heterogeneity across quadrants

General idea: using the derivation of the Negative Binomial as a conditional Poisson distribution where the rate is itself a random variable (gamma distributed) (Corbert, Fisher and Williams, 1943), and averaged over all possible rates.

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} \quad .$$

$$\underbrace{X_t}_{\text{distributed ?}} = \underbrace{\alpha \star X_{t-1}}_{\text{NegBinom}} + \underbrace{\epsilon_t,}_{\text{distributed ?}}$$

Accounting for heterogeneity across quadrants

General idea: using the derivation of the Negative Binomial as a conditional Poisson distribution where the rate is itself a random variable (gamma distributed) (Corbert, Fisher and Williams, 1943), and averaged over all possible rates.

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} .$$

$$\underbrace{X_t}_{\text{distributed ?}} = \underbrace{\alpha \star X_{t-1}}_{\text{NegBinom}(\gamma, p=\frac{\gamma}{\gamma+m})} + \underbrace{\epsilon_t,}_{\text{distributed ?}} .$$

It turns out that X_t is a stationary Markov Chain if we assume that the total size of each arrival results from a random no. of arrival waves between $(t-1, t)$ and a random no. of animals coming at each arrival wave,

Accounting for heterogeneity across quadrants

General idea: using the derivation of the Negative Binomial as a conditional Poisson distribution where the rate is itself a random variable (gamma distributed) (Corbert, Fisher and Williams, 1943), and averaged over all possible rates.

$$\underbrace{X_t}_{\text{no. of sharks at time } t} = \underbrace{\alpha \star X_{t-1}}_{\text{no. of sharks that stayed from } t-1 \text{ to } t} + \underbrace{\epsilon_t,}_{\text{no. of sharks that arrived}} .$$

$$\underbrace{X_t}_{\text{distributed ?}} = \underbrace{\alpha \star X_{t-1}}_{\text{NegBinom}(\gamma, p=\frac{\gamma}{\gamma+m})} + \underbrace{\epsilon_t,}_{\text{distributed ?}} .$$

It turns out that X_t is a stationary Markov Chain if we assume that the total size of each arrival results from a random no. of arrival waves between $(t-1, t)$ and a random no. of animals coming at each arrival wave, *i.e.*, if

$$\epsilon_t = \sum_{i=1}^{N} \alpha^{U_i} \star Y_i = \sum_{i=1}^{N} \sum_{j=1}^{Y_i} B_j(\alpha^{U_i}), \text{ where}$$

$N \sim \mathrm{Poi}(-\gamma \ln \alpha)$, $Y_i \sim \mathrm{Geo}(1, \beta)$, $\beta = \gamma/m$, and $U \sim \mathrm{Unif}(0, 1)$.

Accounting for the extra zeros using a process-based model

The self-decomposable process can be embedded within a zero inflated process by assuming that there are two ways of observing 0 tagged sharks at a given point in time and space:

Accounting for the extra zeros using a process-based model

The self-decomposable process can be embedded within a zero inflated process by assuming that there are two ways of observing 0 tagged sharks at a given point in time and space:

1. with probability π the quadrant is absolutely unsuitable at time t so no sharks will arrive and stay, or

Accounting for the extra zeros using a process-based model

The self-decomposable process can be embedded within a zero inflated process by assuming that there are two ways of observing 0 tagged sharks at a given point in time and space:

1. with probability π the quadrant is absolutely unsuitable at time t so no sharks will arrive and stay, or
2. the quadrant is suitable with probability $1 - \pi$ but the sharks that were in it just left, *i.e.*

Accounting for the extra zeros using a process-based model

The self-decomposable process can be embedded within a zero inflated process by assuming that there are two ways of observing 0 tagged sharks at a given point in time and space:

1. with probability π the quadrant is absolutely unsuitable at time t so no sharks will arrive and stay, or
2. the quadrant is suitable with probability $1-\pi$ but the sharks that were in it just left, *i.e.*

Idea: Model π as a function of the value of the thermocline depth w. To do that, we take two approaches:

Accounting for the extra zeros using a process-based model

The self-decomposable process can be embedded within a zero inflated process by assuming that there are two ways of observing 0 tagged sharks at a given point in time and space:

1. with probability π the quadrant is absolutely unsuitable at time t so no sharks will arrive and stay, or
2. the quadrant is suitable with probability $1 - \pi$ but the sharks that were in it just left, *i.e.*

Idea: Model π as a function of the value of the thermocline depth w. To do that, we take two approaches:

1. Use the traditional -yet phenomenological- function $\pi(w) = 1/(1 + \exp^{-f(w)})$, where $f(w)$ is a linear function of the covariate w,

Accounting for the extra zeros using a process-based model

The self-decomposable process can be embedded within a zero inflated process by assuming that there are two ways of observing 0 tagged sharks at a given point in time and space:

1. with probability π the quadrant is absolutely unsuitable at time t so no sharks will arrive and stay, or
2. the quadrant is suitable with probability $1 - \pi$ but the sharks that were in it just left, *i.e.*

Idea: Model π as a function of the value of the thermocline depth w. To do that, we take two approaches:

1. Use the traditional -yet phenomenological- function $\pi(w) = 1/(1 + \exp^{-f(w)})$, where $f(w)$ is a linear function of the covariate w,
2. Use a novel approach about niche modeling from ecological and evolutionary ideas.

An example from modeling fitness components (reproduction and survival)

An example from modeling fitness components (reproduction and survival)

- How does the function of how good a species fares as a function of a covariate has to look like?

An example from modeling fitness components (reproduction and survival)

- How does the function of how good a species fares as a function of a covariate has to look like?
- How does the shape of such function can arise from simple ecological principles (*i.e.* compatible with Hutchinsons niche concept)?

Modeling fitness components (reproduction and survival)

- Define $\overline{\phi(w)} = \int \phi(w) h_1(\phi(w)) d\phi(w)$ as the expected survival probability of any given individual in a population, as a function of w.

Modeling fitness components (reproduction and survival)

- Define $\overline{\phi(w)} = \int \phi(w) h_1(\phi(w)) d\phi(w)$ as the expected survival probability of any given individual in a population, as a function of w.
- $h_1(\phi(w))$ is the pdf of the survival probabilities $\phi(w)$ as quantities that depend on w

Modeling fitness components (reproduction and survival)

- Define $\overline{\phi(w)} = \int \phi(w) h_1(\phi(w)) d\phi(w)$ as the expected survival probability of any given individual in a population, as a function of w.
- $h_1(\phi(w))$ is the pdf of the survival probabilities $\phi(w)$ as quantities that depend on w
- For a given large number m of individuals, $\overline{\phi(w)} \approx \frac{1}{m} \sum_{i=1}^{m} \phi_i(w)$.

Modeling fitness components (reproduction and survival)

- Define $\overline{\phi(w)} = \int \phi(w) h_1(\phi(w)) d\phi(w)$ as the expected survival probability of any given individual in a population, as a function of w.
- $h_1(\phi(w))$ is the pdf of the survival probabilities $\phi(w)$ as quantities that depend on w
- For a given large number m of individuals, $\overline{\phi(w)} \approx \frac{1}{m} \sum_{i=1}^{m} \phi_i(w)$.
- Assuming that the climatic condition w affects individuals' survival probability independently from each other, then

$$\left(1 - \frac{\sum_{i=1}^{m} \phi_i(w)}{m}\right)^{m} \approx e^{-\sum_{i=1}^{m} \phi_i(w)}$$

  is an estimate of the probability that m individuals do not survive under w.

Modeling fitness components (reproduction and survival)

- Define $\overline{\phi(w)} = \int \phi(w) h_1(\phi(w)) d\phi(w)$ as the expected survival probability of any given individual in a population, as a function of w.
- $h_1(\phi(w))$ is the pdf of the survival probabilities $\phi(w)$ as quantities that depend on w
- For a given large number m of individuals, $\overline{\phi(w)} \approx \frac{1}{m} \sum_{i=1}^{m} \phi_i(w)$.
- Assuming that the climatic condition w affects individuals' survival probability independently from each other, then

  $$\left(1 - \frac{\sum_{i=1}^{m} \phi_i(w)}{m}\right)^{m} \approx e^{-\sum_{i=1}^{m} \phi_i(w)}$$

  and an estimate of the probability that m individuals survive under w is then

  $$1 - e^{-\sum_{i=1}^{m} \phi_i(w)}$$

Modeling fitness components (reproduction and survival)

- Likewise, letting $\overline{\psi(w)} = \int \psi(w) h_1(\psi(w)) d\psi(w)$ be the average probability of not reproducing given the value of w and that survival has occurred, we get that $e^{-\sum_{i=1}^{m} \psi_i(w)}$ can be taken as an estimate of the probability that m individuals actually reproduce given the value of w.

Modeling fitness components (reproduction and survival)

- Likewise, letting $\overline{\psi(w)} = \int \psi(w)h_1(\psi(w))d\psi(w)$ be the average probability of not reproducing given the value of w and that survival has occurred, we get that $e^{-\sum_{i=1}^{m} \psi_i(w)}$ can be taken as an estimate of the probability that m individuals actually reproduce given the value of w.
- Without loss of generality, we can write

$$\sum_{i=1}^{m} \phi_i(w) = \delta(w) = \delta_0 + \delta_1 w \text{ and } \gamma(w) = \gamma_0 + \gamma_1 w \approx \sum_{i=1}^{m} \psi_i(w)$$

.

Modeling fitness components (reproduction and survival)

- Likewise, letting $\overline{\psi(w)} = \int \psi(w)h_1(\psi(w))d\psi(w)$ be the average probability of not reproducing given the value of w and that survival has occurred, we get that $e^{-\sum_{i=1}^{m}\psi_i(w)}$ can be taken as an estimate of the probability that m individuals actually reproduce given the value of w.
- Without loss of generality, we can write

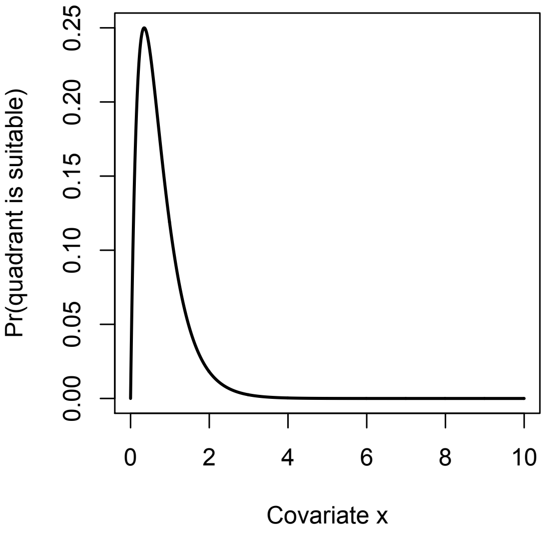
$$\sum_{i=1}^{m}\phi_i(w) = \delta(w) = \delta_0 + \delta_1 w \text{ and } \gamma(w) = \gamma_0 + \gamma_1 w \approx \sum_{i=1}^{m}\psi_i(w)$$

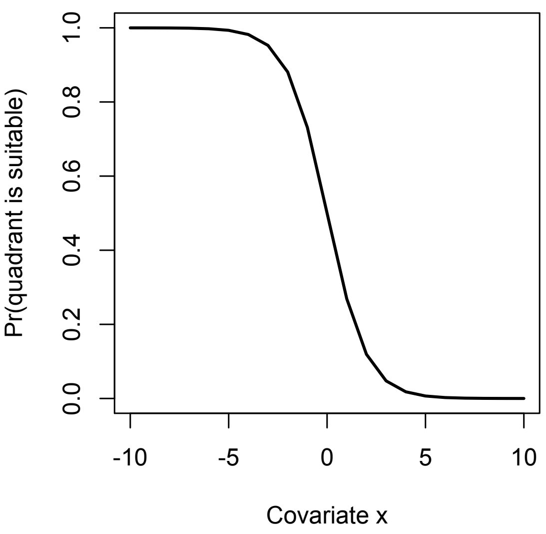
.

Hence, the covariate-dependent maximum growth rate $a(w)$ could be written as

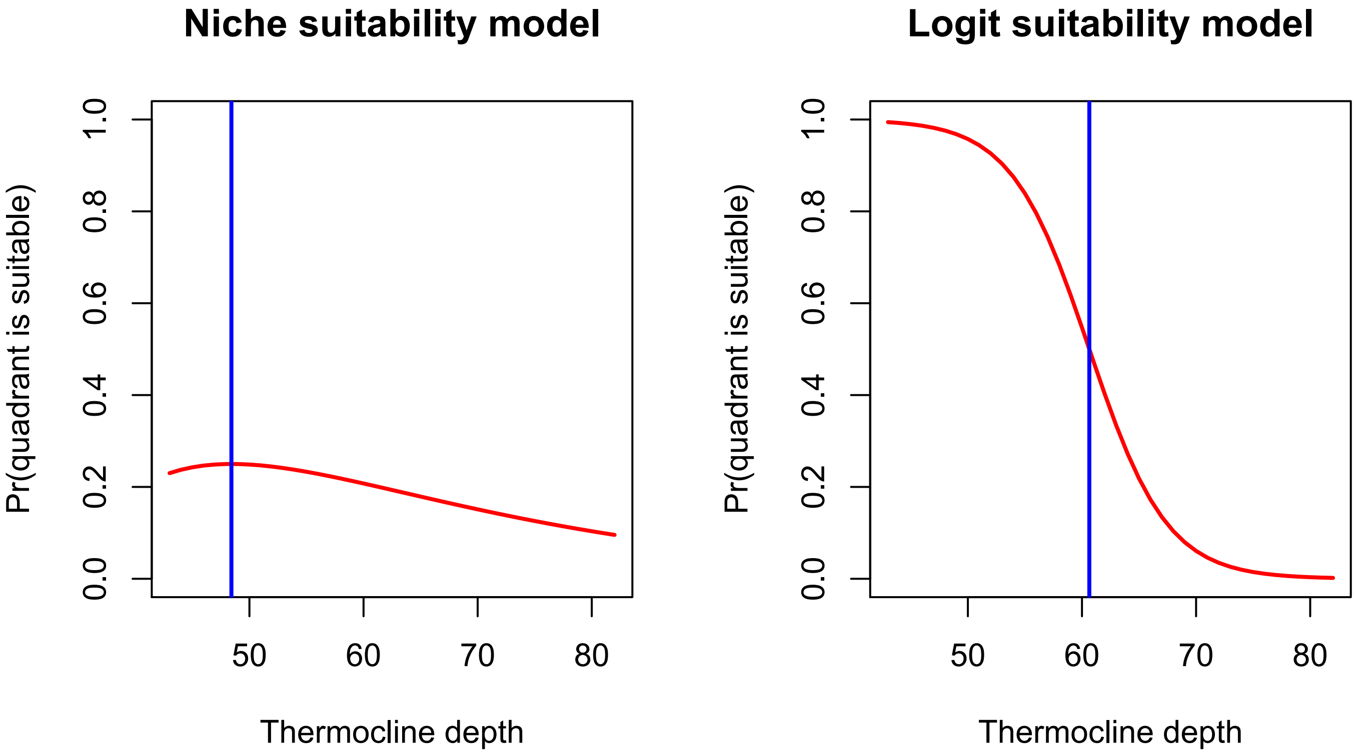
$$\begin{aligned} a(w) &= \text{constant} \times P(\text{surviving})P(\text{reproducing}|\text{survival has occurred}) \\ &= \lambda\left(1 - e^{-\delta(w)}\right)e^{-\gamma(w)}. \end{aligned}$$

The two hypotheses regarding quadrant suitability


Hypothesis 1: Niche model optimum


Hypothesis 2: logit transform model

Estimated suitabilities using only movement and DML data



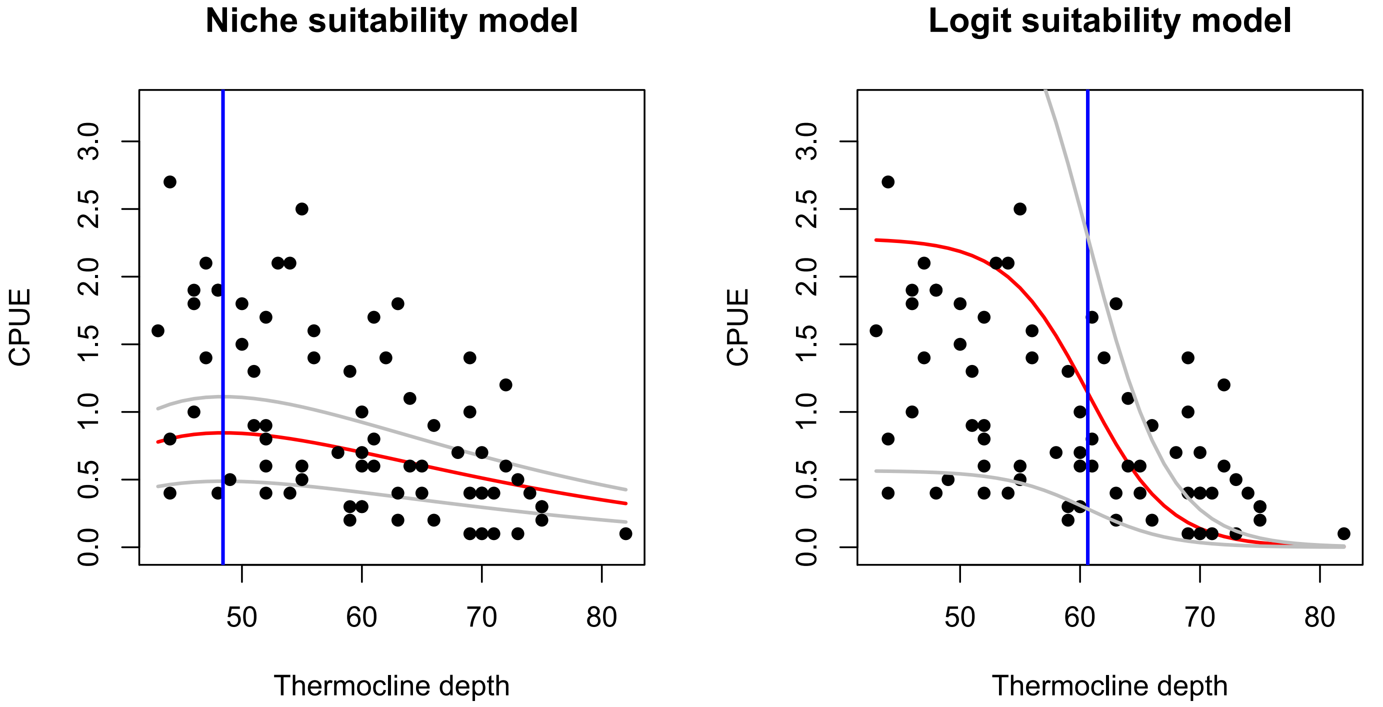
CPUES: a random sample from a distribution with mean proportional to suitability

Once we have an estimate (and CI's) of the suitability modeled as a function of the depth of the thermocline, we model the CPUEs as random samples from a delta-lognormal distribution whose mean is proportional to these suitabilities. *i.e.* we let the CPUEs Y_i

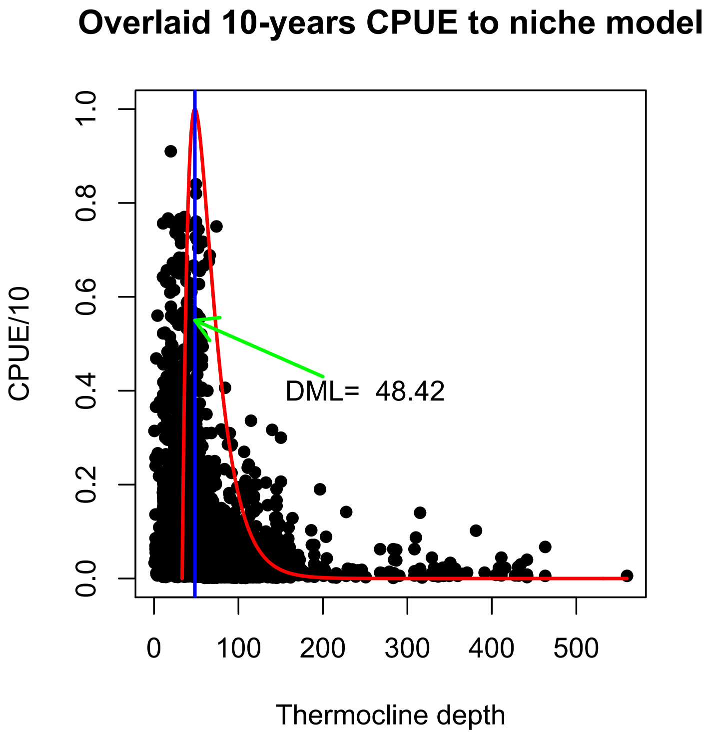
$$Y_i \sim \delta \mathrm{logNorm}(\delta, \mu = \ln(c.\pi(w_i)) - \frac{\tau^2}{2}, \tau^2), \qquad \text{so it follows that}$$

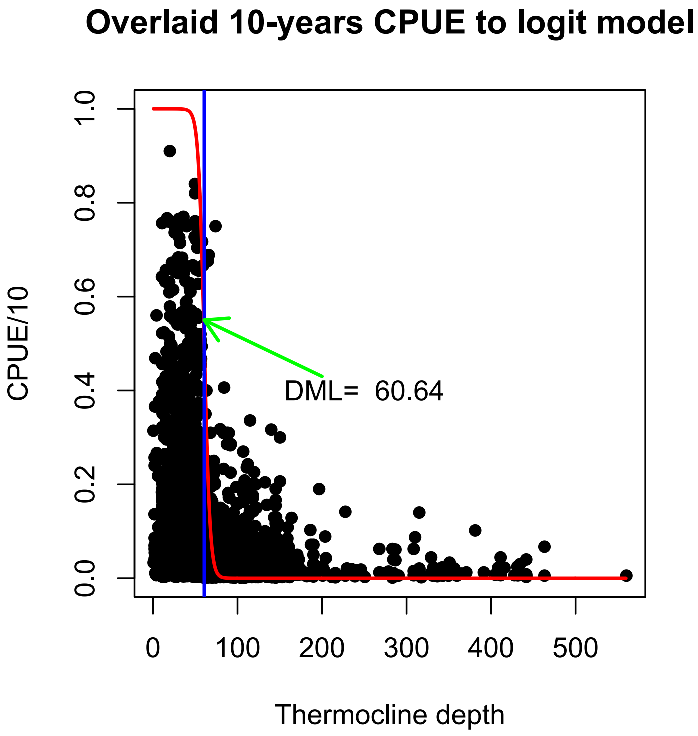
$$\mathrm{E}[Y_i] = \delta \exp\left\{\mu + \frac{\tau^2}{2}\right\}$$

Modeling the CPUEs with a δ-lognormal with mean proportional to the suitability



Re-scaled estimated suitabilities from movement data overlaid to 10-year span CPUEs


Overlaid 10-years CPUE to niche model


Overlaid 10-years CPUE to logit model

Acknowledgments

BRASIL
o país do pescado

agriculture,
forestry & fisheries
Department:
Agriculture, Forestry and Fisheries
REPUBLIC OF SOUTH AFRICA

MADE

UF UNIVERSITY of FLORIDA

ICCAT CICTA CICAA

EX SEMINE SEGES
UFRPE

NOAA
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
U.S. DEPARTMENT OF COMMERCE

PFRP Pelagic Fisheries Research Program
of the Joint Institute for Marine and Atmospheric Research (JIMAR)
IN THE SCHOOL OF OCEAN AND EARTH SCIENCE AND TECHNOLOGY AT THE UNIVERSITY OF HAWAI'I AT MĀNOA