Whales, Politics and Statisticians

Céline Cunen (joint work with Nils Lid Hjort amd Lars Walløe)

University of Oslo

25/05/2018

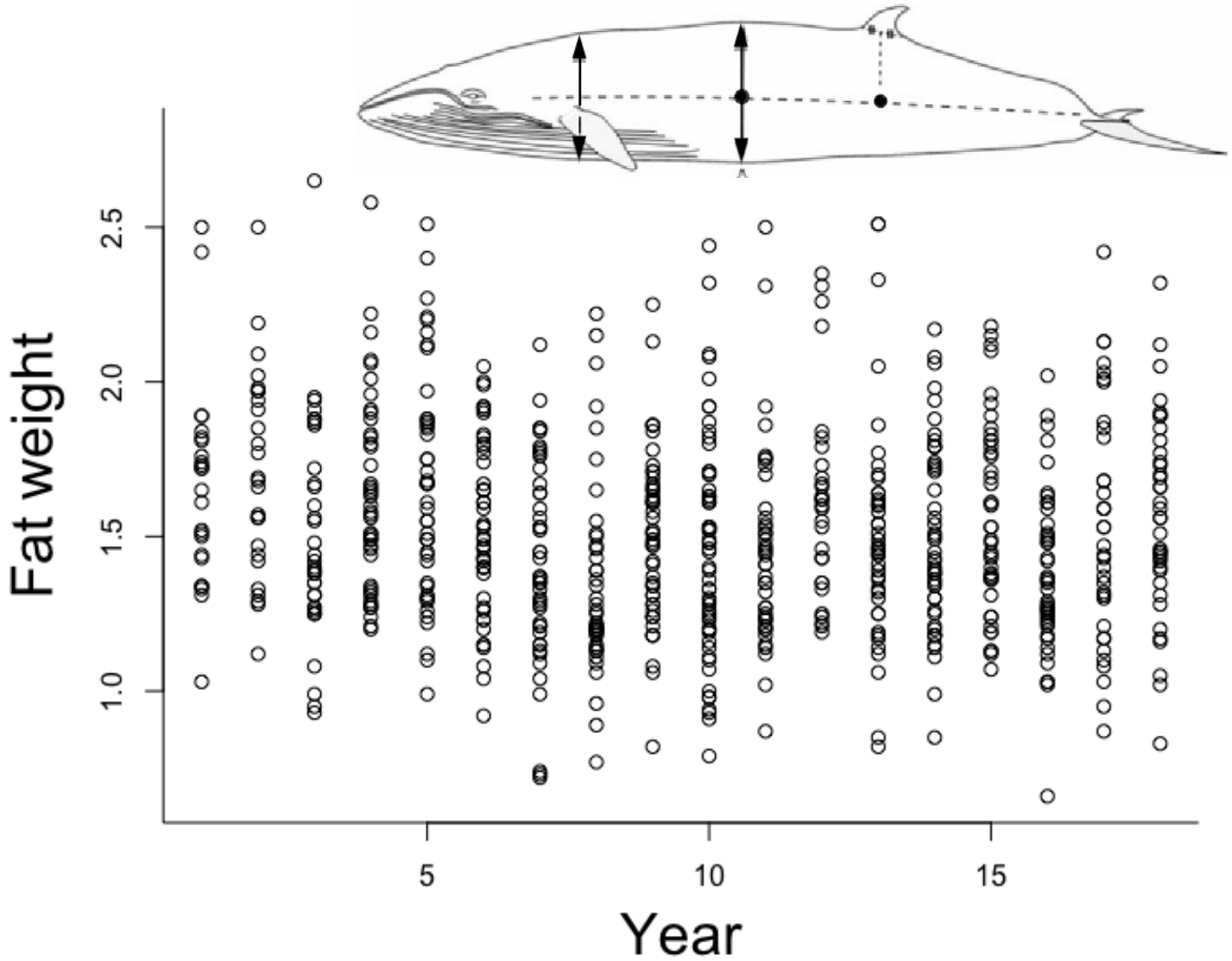
Antarctic Minke Whales

"Small" (8.5 meters long) $\rightarrow$ abundant ($\sim$ 515000), hunted by Japan

Photo: Australian Antarctic Division

The problem: are the whales getting thinner?

5 (correlated) measurements of body condition: blubber thickness, girth, and **fat weight** (742 observations from 1987 to 2005).



The adjustment game

Measured variables

- **Sex**: females have more fat*
- **Fetus length**: females with longer fetuses have more fat ($\leftrightarrow$?)
- **Body length**: longer whales have more fat*
- **Total body weight**: heavier whales have more fat*, but...
- (Age)
- Date (in each season)
- Diatom load
- Spatial variables:
  - Latitude
  - Longitude
  - Ice
  - Areas

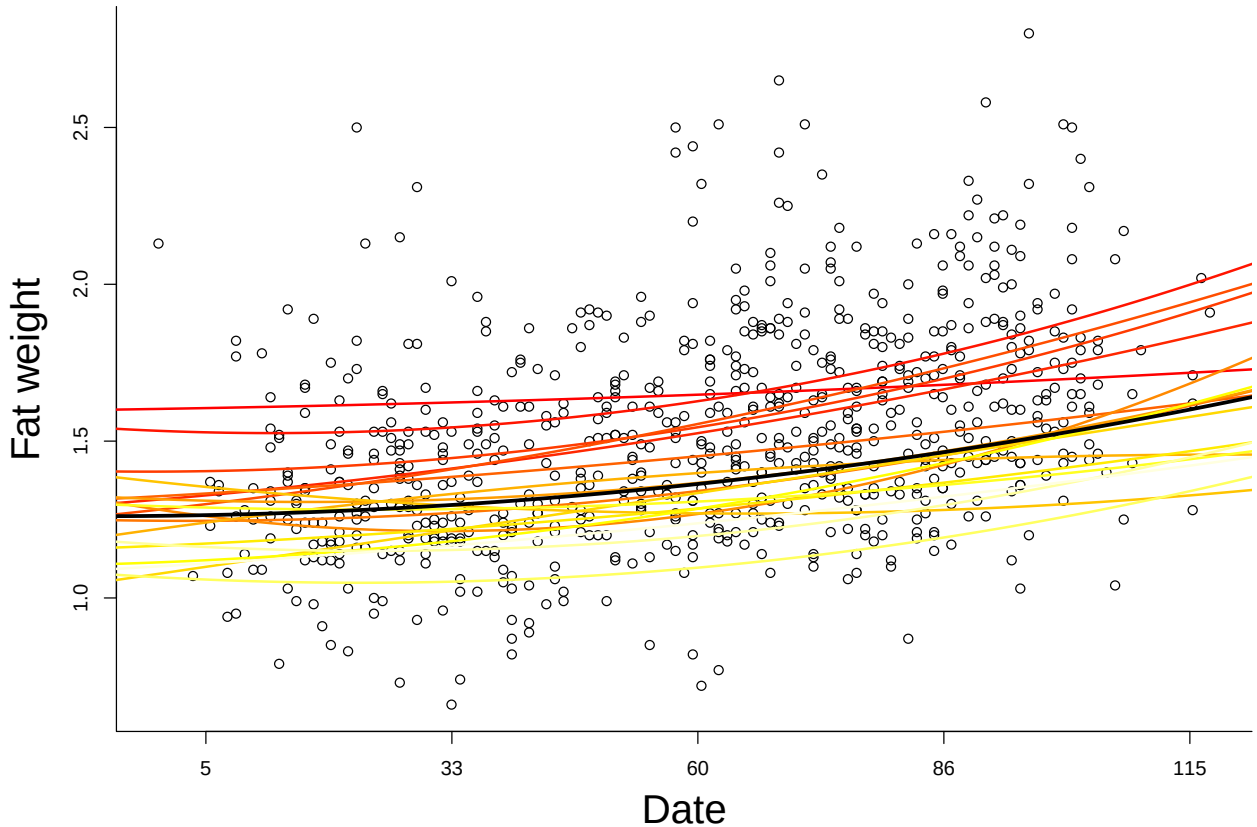
* all other things being equal

The adjustment game

Measured variables

- Sex
- Fetus length
- Body length
- Total body weight
- (Age)
- **Date**
- **Diatom load**
- Spatial variables:
  - Latitude
  - Longitude
  - Ice
  - Areas

Date (in each year)

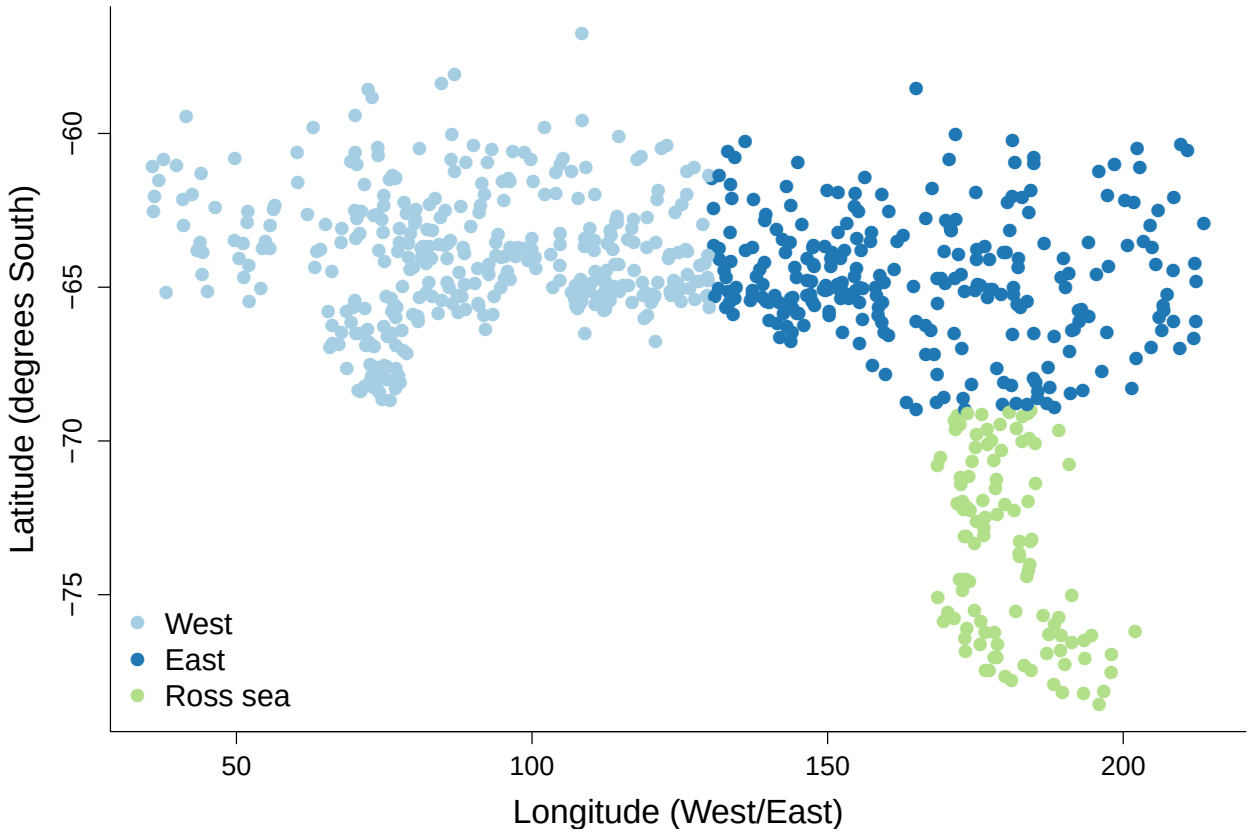


December 1st = day 1

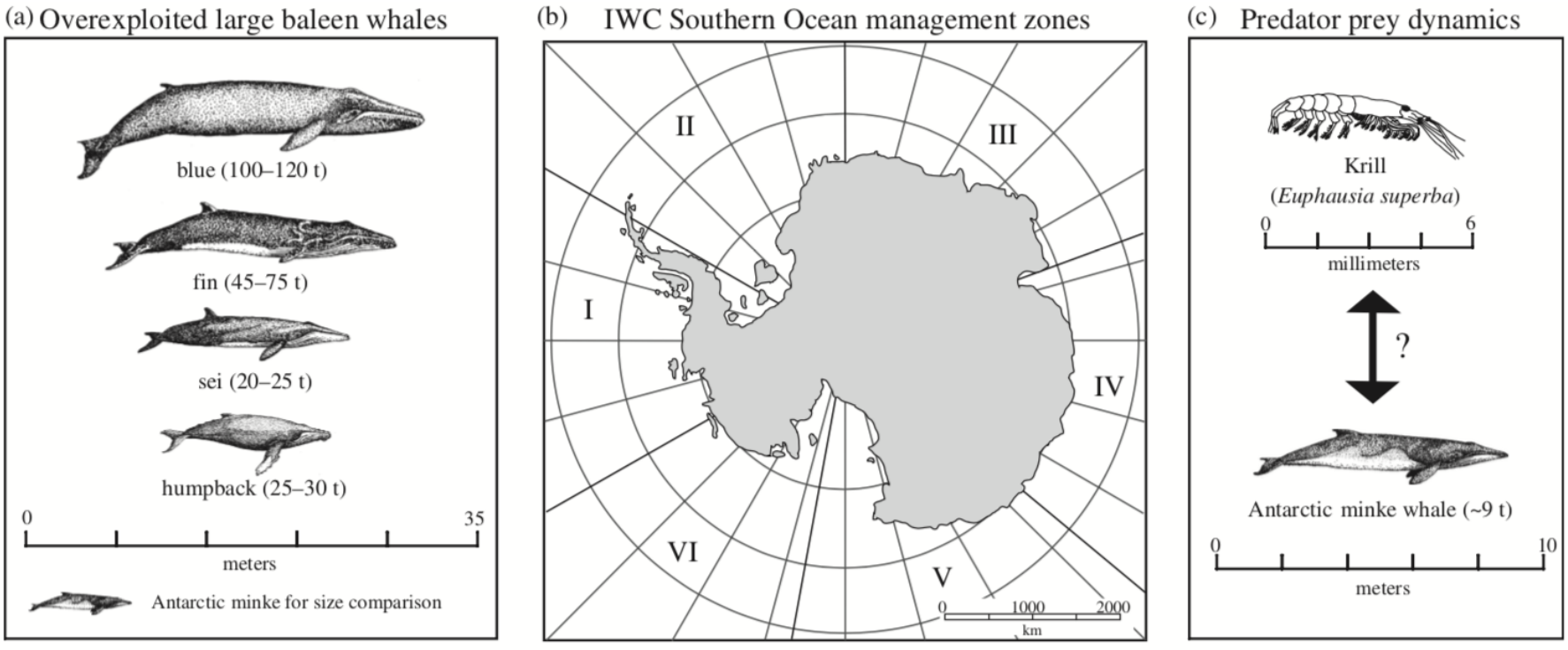
The adjustment game

Measured variables

- Sex
- Fetus length
- Body length
- Total body weight
- (Age)
- Date
- Diatom load
- **Spatial variables:**
  - Latitude
  - Longitude
  - Ice
  - Areas



The biological context: the Krill Surplus Hypothesis



Ruegg et al. Are Antarctic minke whales unusually abundant because of 20th century whaling?. *Molecular Ecology* (2010).

Before: fewer large whales $\rightarrow$ more krill for minke whales $\rightarrow$ more and fatter minke whales?

Now: more large whales $\rightarrow$ less krill for minke whales $\rightarrow$ less and thinner minke whales?

The meta-problem

The data result from the JARPA - Japanese Whale Research Program under Special Permit in the Antarctic, and have been presented and discussed in the Scientific Committee of the International Whaling Commission (IWC).

The specific minke whale problem above has been discussed for more than ten years:

- 2007: first paper presented and discussed
- from 2011: new papers and discussions every year
- **2014: the Committee concluded that the decline in body condition had been sufficiently well documented**, but Australian scientists were not present at that meeting, so the discussions have continued since

The meta-problem

What are the positions?

- **Japanese (and Norwegian) scientists**: there has been a decline in body condition.
- **Australian scientists**: we cannot draw any conclusion, too much uncertainty.

What do the discussions concern?

- Some biology: the right/best response variable, the covariates to include in the model,...
- A lot of statistics: the sampling process, potential non-independences, inference in linear mixed effect (LME) models, goodness-of-fit in LME models, AIC vs. BIC as model selection criteria, post-selection issues,...

Our analysis

Our task (according to the Scientific Committee):

- use linear mixed effect models to account for (potential) non-independences between the observations;
- try to make the estimated yearly trend "more precise" $\longrightarrow$ FIC?!

We decided on doing two things:

1. develop a large, biologically plausible model (the wide model) and investigate the yearly trend there;
2. try to simplify the wide model using FIC (to find a more precise estimate of the year trend).

Linear Mixed Effect models

Data with n natural groups, $i = 1, \ldots, n$, with m_i observations in each group. The general LME model takes the form

$$\underset{m_i \times 1}{y_i} = \underset{m_i \times p}{X_i}\,\beta + \underset{m_i \times k}{Z_i}\,b_i + \epsilon_i, \quad \text{with } \epsilon_i \sim \mathrm{N}_{m_i}(0, \sigma^2 I).$$

with

- Z_i design matrix corresponding to the random effects;
- b_i the group-specific vector of random effects, with $b_i \sim \mathrm{N}_k(0, \sigma^2 D)$.

Equivalently,

$$y_i \sim \mathrm{N}_{m_i}\big(X_i\beta, \sigma^2(I + Z_i D Z_i^{\mathrm{t}})\big),$$

and independence across groups.

Wide model

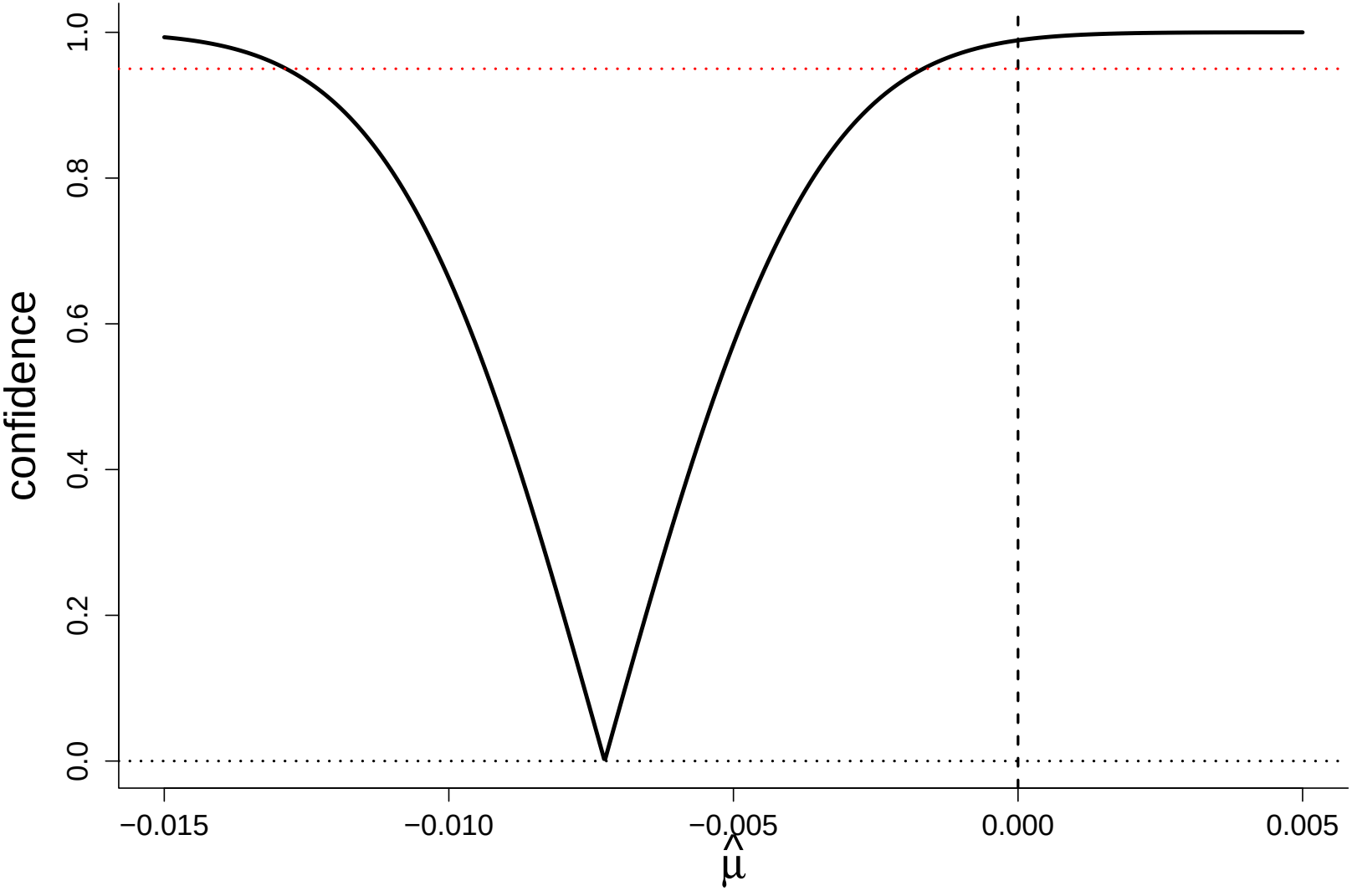
After reading and thinking:

$$
\begin{aligned}
Y \sim & \mathit{Year} + \mathit{Year}^2 + \mathit{BodyLength} + \mathit{Sex} + \mathit{Diatom} + \mathit{Date} + \mathit{Date}^2 \\
& + \mathit{Latitude} + \mathit{Sex} * \mathit{FetusLength} + \mathit{Sex} * \mathit{Diatom} + \mathit{Diatom} * \mathit{Date} \\
& + \mathit{Diatom} * \mathit{Date}^2 + \mathit{Latitude} * \mathit{Date} + \mathit{Latitude} * \mathit{Date}^2 \\
& + \mathit{Region} + \mathit{Year} * \mathit{Region} + \mathit{Year}^2 * \mathit{Region} + \mathit{Latitude} * \mathit{Region} \\
& + \mathit{Sex} * \mathit{Region} + \mathit{Diatom} * \mathit{Region} + (1 + \mathit{Date} + \mathit{Date}^2 | \mathit{Year}),
\end{aligned}
$$

28 fixed effect parameters, 35 parameters in total.

Focus parameter: $\mu = \beta_{\text{year}} + 2\beta_{\text{year2}}\bar{x}_{\text{year}}$.

Wide model result



Building a FIC for LMEs

- Large models can give large variances for the estimates;
- FIC offers a principled way of choosing a simpler/smaller model with a specific focus in mind.

In candidate model M we have a linear year effect, so $\mu_M = \beta_{\text{year}}$ with an associated estimate $\hat{\beta}_{M,\text{year}}$. FIC ranks models according to their estimated mean-squared error (MSE) $\text{FIC}(M) = \widehat{\text{mse}}_M$,

$$\text{mse}_M = \text{E}\,(\hat{\beta}_{M,\text{year}} - \beta_{\text{year}})^2 = \text{Var}\,\hat{\beta}_{M,\text{year}} + (\text{E}\,\hat{\beta}_{M,\text{year}} - \beta_{\text{year}})^2.$$

Here, we compute both the variance and the bias **with respect to the wide model**.

FIC – type 3?

The wide model is "the largest" plausible parametric model.

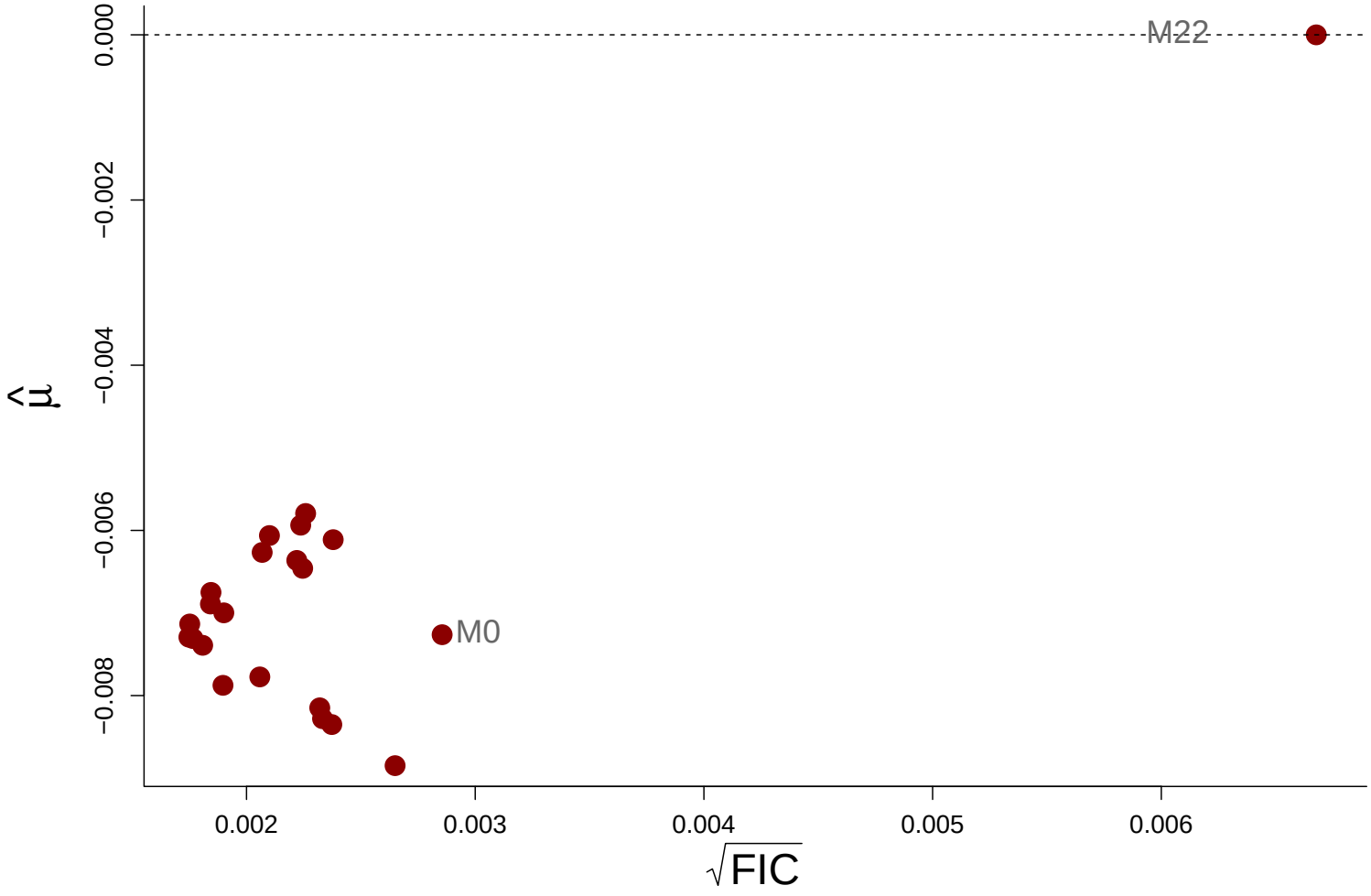
- no local misspecification;
- no non-parametric wide model.

Thus, in the limit, the wide model will always win*.

$$\begin{pmatrix} \sqrt{n}(\widehat{\mu} - \mu_{true}) \\ \sqrt{n}(\widehat{\mu}_M - \mu_{M,0}) \end{pmatrix} \approx \mathrm{N}_2 \left(0, \begin{pmatrix} \nu_{wide} & \nu_{M,c} \\ \nu_{M,c} & \nu_M \end{pmatrix} \right),$$

with $\nu_{wide} = c^{\mathrm{t}} J_n^{-1} c$, $\nu_{M,c} = c^{\mathrm{t}} J_n^{-1} C_{M,n} J_{M,n}^{-1} c_M$,
$\nu_M = c_M^{\mathrm{t}} J_{M,n}^{-1} K_{M,n} J_{M,n}^{-1} c_M$.
We have explicit formulas for all the necessary quantities.

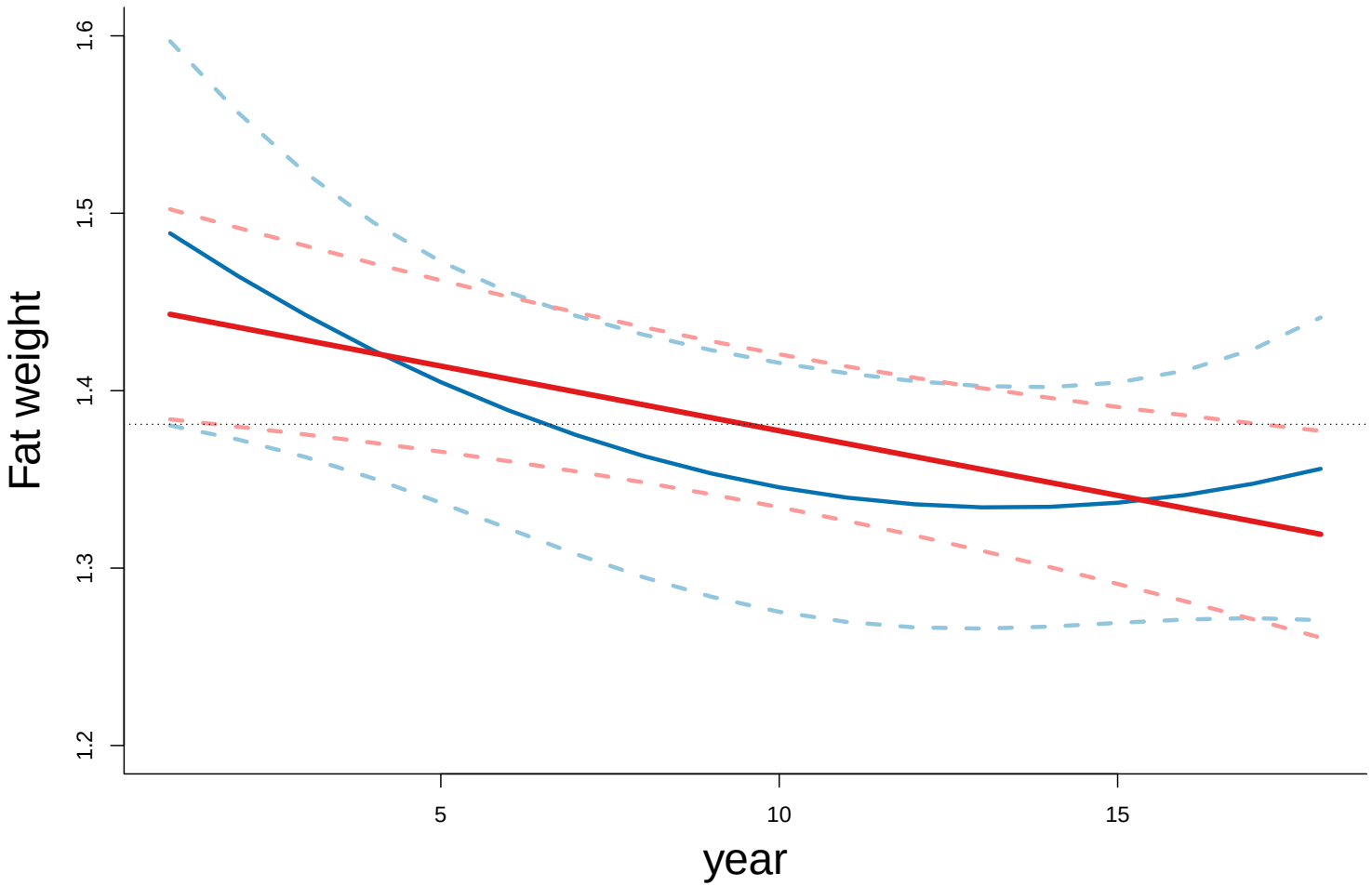
FIC results



Winning model:

$$Y \sim Year + BodyLength + Sex + Diatom + Date + Date^2 + Latitude + Sex * FetusLength + Sex * Diatom + Region + (1 + Date + Date^2 | Year)$$

Lines with bands



The decline in fat weigth over the years, in the wide model and in the winning model according to FIC. After model selection with FIC we get narrower confidence bands.

How large is the decline? Does it matter?

About 7.3 kg fat per year $\longrightarrow$ almost 5% decrease in fat reserves over 10 years.

- the whales are in the Antarctic to become fat;
- in a typical season, they gain fat at the rate of about 2.7 kg per day;
- over a 10 year period, they seem to "have lost" 73 kg fat, the equivalent of about 27 feeding days;
- during the breeding season (the winter?) they feed little. The fat-accumulation in the summer season is therefore important.
- reduced reproduction?

At the Scientific Committee meetings 2017 and 2018

Heated discussions (also in the months prior to the meeting), hundreds of pages of reports, notes of criticism and counter-notes, frequent demands for new analyses (usually with small changes).

Australian criticism of the Norwegian contribution

- FIC results in biased estimates;
- the model selected by FIC depends on the choice of the wide model;
- the selected model is biologically implausible;
- the FIC scores are uncertain - which model is *r*eally the best?

And lots of other things:

- the data should be split into different groups;
- the effects are not linear;
- treatment effect contrast vs sum-to-zero contrast (!?);
- the choice of optimiser;
- ...

End result

After days and days of exhausting discussions:

- **more or less status quo**;
- neither side were convinced;
- the conclusion from 2014 was not overturned;
- the Scientific committee decided that the question will not be discussed again before 2021 (!) at the earliest.

Reflections and speculations

Why was the body condition question discussed for so long? Why did it matter?

- Very strong conviction on both sides?
- Individual pride?
- Politics?

Reflections and speculations: Politics?

The Japanese scientific whaling has labelled as "fake science" by NGOs and other countries. Australia took Japan to the UN's International Court of Justice in 2013:
"The court did not rule whether the whales, whose meat is sold to cover the costs of the hunt, were caught commercially or not. But it cited evidence – supported by Japans own expert witness – that **catch sizes were not set to meet scientific objectives**. *Rather [...] a desire for certain catch sizes seemed to have determined the scientific plan."*(New Scientist, 2014)

Reflections and speculations: Politics?

The idea: **fake science cannot produce scientific findings**. So, conclusions from fake science must be contested.

The method: expand the uncertainty (in the words of T. Schweder).

"Crucially for the Japanese, the results have been published in a mainstream **western scientific journal** *– a move that has dismayed campaigners, who say it could offer scientific whaling a veneer of respectability, and bolster Japan's efforts to hunt more whales."* (The Guardian, 2008).

The results from the original paper were even picked up by various mainstream media.

Some media interest

National Geographic (2016),

A Rare Look at the Disappearing World of Antarctica's Whales

As the southern continent rapidly warms, some whale populations are booming—while others are suffering from lack of ice.

BY **DOUGLAS FOX**

PUBLISHED DECEMBER 6, 2016

CIERVA COVE, ANTARCTICA—Ari Friedlaender kneels at the front of our inflatable boat as it bobs on the ocean.

He holds a rifle, with two metal harpoons protruding from its muzzle. His gaze

"It is even possible that as their hunting grounds recede, the animals are finding less to eat. Over the last 28 [sic] years, the average amount of body fat carried by Antarctic minke has declined by 9 percent, and the average amount of food in their stomachs has fallen 31 percent, according to data collected by Japanese whalers."

Some media interest

The Guardian (2008),

Whaling

Whales losing blubber, claims controversial Japanese study

Data from Japan's widely condemned scientific whaling programme suggests a loss of fat over the past 20 years may be due to climate change, but some claim the research is unethical

David Adam, *Environment correspondent*

Tue 26 Aug 2008
09.52 BST

"The team says its study offers the first evidence that global warming could be harming whales, because it restricts their food supplies."
"This is not the first time that the Japanese scientific whaling programme has published results, but these are the most high-profile findings so far."

Reflections and speculations

The uncertainty game has been played in many fields:

- the relationship between smoking and lung-cancer;
- humans and climate change;
- is it healthy to eat farmed salmon?
- where do Norwegian wolves come from?

but the arguments are not necessarily wrong!

Statisticians are good at the uncertainty game.

Testimonial from the IWC scientific committee (not mine!)

"[...] it is somewhat frustrating to be **constantly bogged down arguing over the exact wording of a report** – that quite frankly hardly anyone will read, or making off-hand interventions that derail carefully set up initiatives or resolutions [...]. Some of the scientists are the IWC are **cunning and Machiavellian**, trying to insert documents into the meeting that bolster the case for whaling or try to increase the number of animals that can be taken. Some however are **painfully naive, and clueless** as to the politics that whirl round them. So those of us who are trying **to fight the good fight to conserve whales are not only battling against the forces of darkness** trying to argue that more whales can be killed, but also against oblivious scientists who do not understand how their words and statements will be used, or how their results will be twisted and misinterpreted."

The kraken in the aquarium, (Professor in marine conservation)

References

- Cunen, Walløe, and Hjort (2017). Decline in Energy Storage in Antarctic Minke Whales During the JARPA Period: Assessment via the Focused Information Criterion (FIC). *Report of the Scientific Committee of the International Whaling Commision*.
- Cunen, Walløe, and Hjort (2018). Focused Model Selection for Linear Mixed Models, with an Application to Whale Ecology. *Submitted for Publication*.
- Konishi, Tamura, Zenitani, Bando, Kato and Walløe (2008). Decline in energy storage in the Antarctic minke whale (balaenoptera bonaerensis) in the southern ocean. *Polar Biology*.
- Konishi and Walløe (2015). Substantial decline in energy storage and stomach fullness in Antarctic minke whales during the 1990s. *Journal of Cetacean Research and Management*.
- de la Mare, McKinlay, and Welsh (2017). Analyses of the JARPA Antarctic Minke Whale Fat Weight Data Set. *Report of the Scientific Committee of the International Whaling Commision*.
- McKinlay, de la Mare, and Welsh (2017). A Re-Examination of Minke Whale Body Condition as Reflected in the Data. *Report of the Scientific Committee of the International Whaling Commision*.
- McKinlay, de la Mare, and Welsh (2018). No Substantial Change in Antarctic Minke Whale Condition During the JARPA Years. *Report of the Scientific Committee of the International Whaling Commision*.
- Schweder (2001). Protecting whales by distorting uncertainty: non-precautionary mismanagement? *Fisheries Research*.