

# An ocean of uncertainty

Liam Brannigan  
Dublin, March 2018

# Plan

- Bayesian machine learning
- Bayes Theorem is your friend
- Code & visualisation examples
- Oceanography example
- Summary

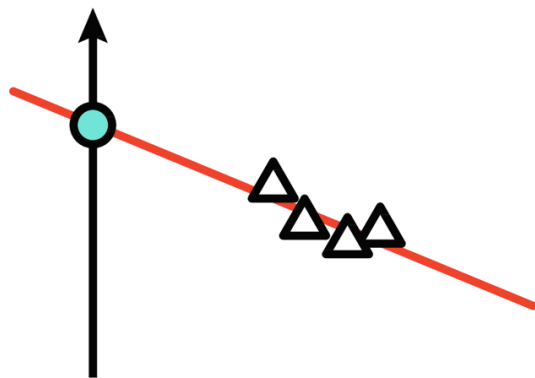
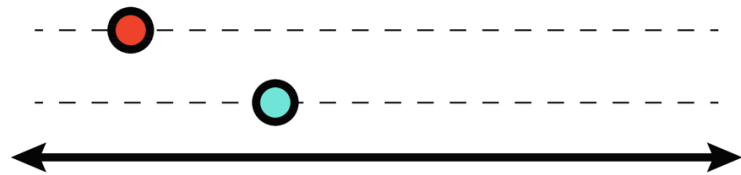
# About me

- Maths undergrad in Edinburgh
- Ocean physics PhD in Oxford
- Postdoc in Stockholm
- Looking for data science role in Dublin
- @braaannigan

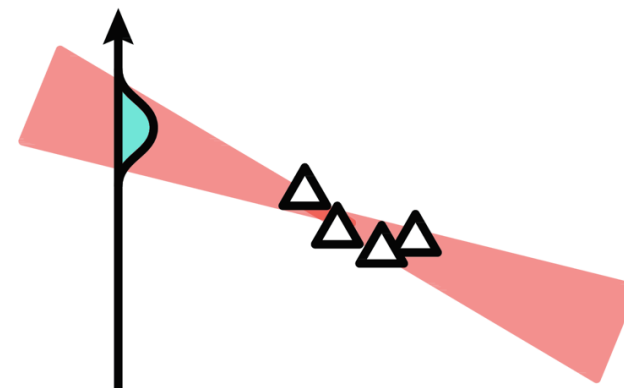
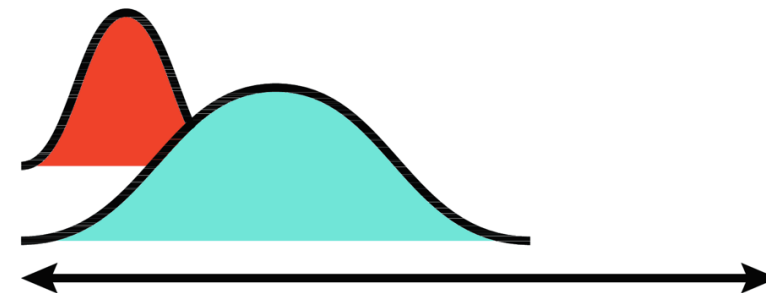
# What is Bayesian machine learning?

# Linear regression

Non-Bayesian

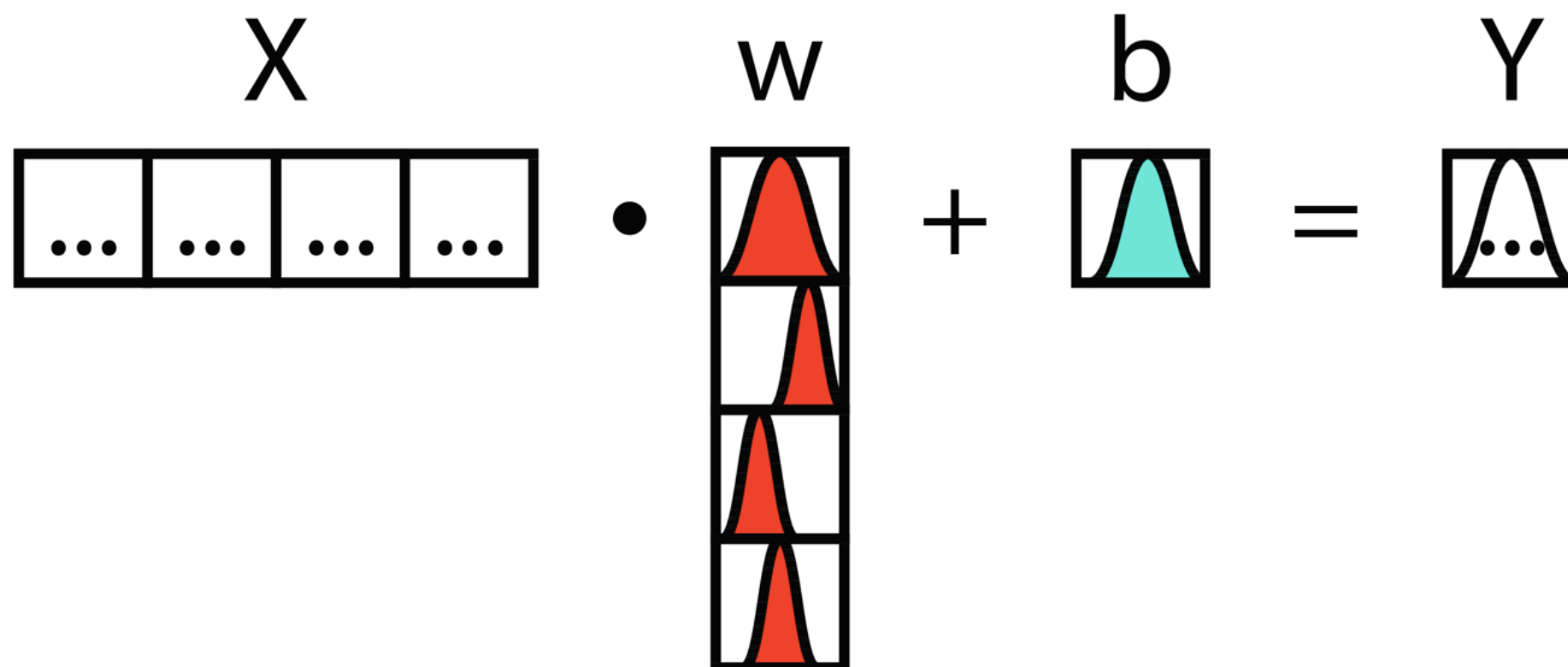
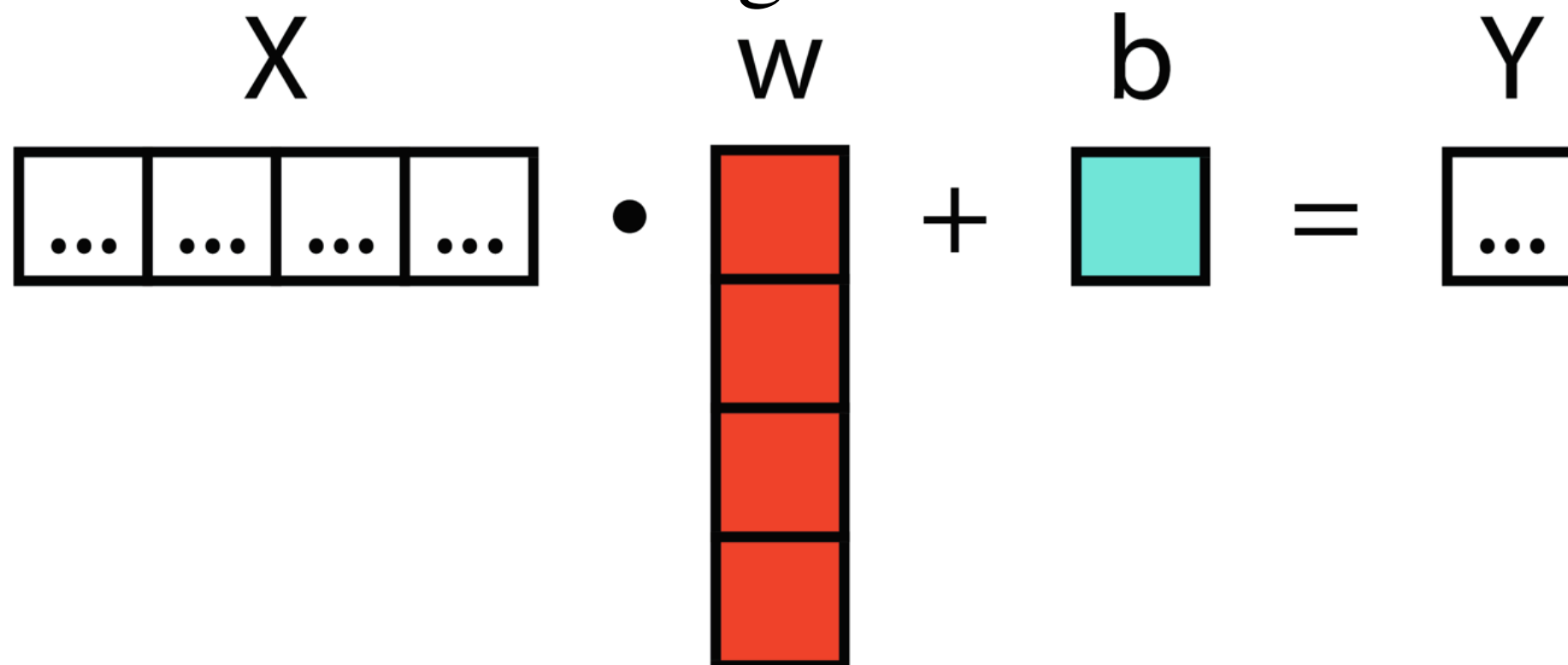


Bayesian



<https://ericmjl.github.io/bayesian-deep-learning-demystified>

# Linear regression

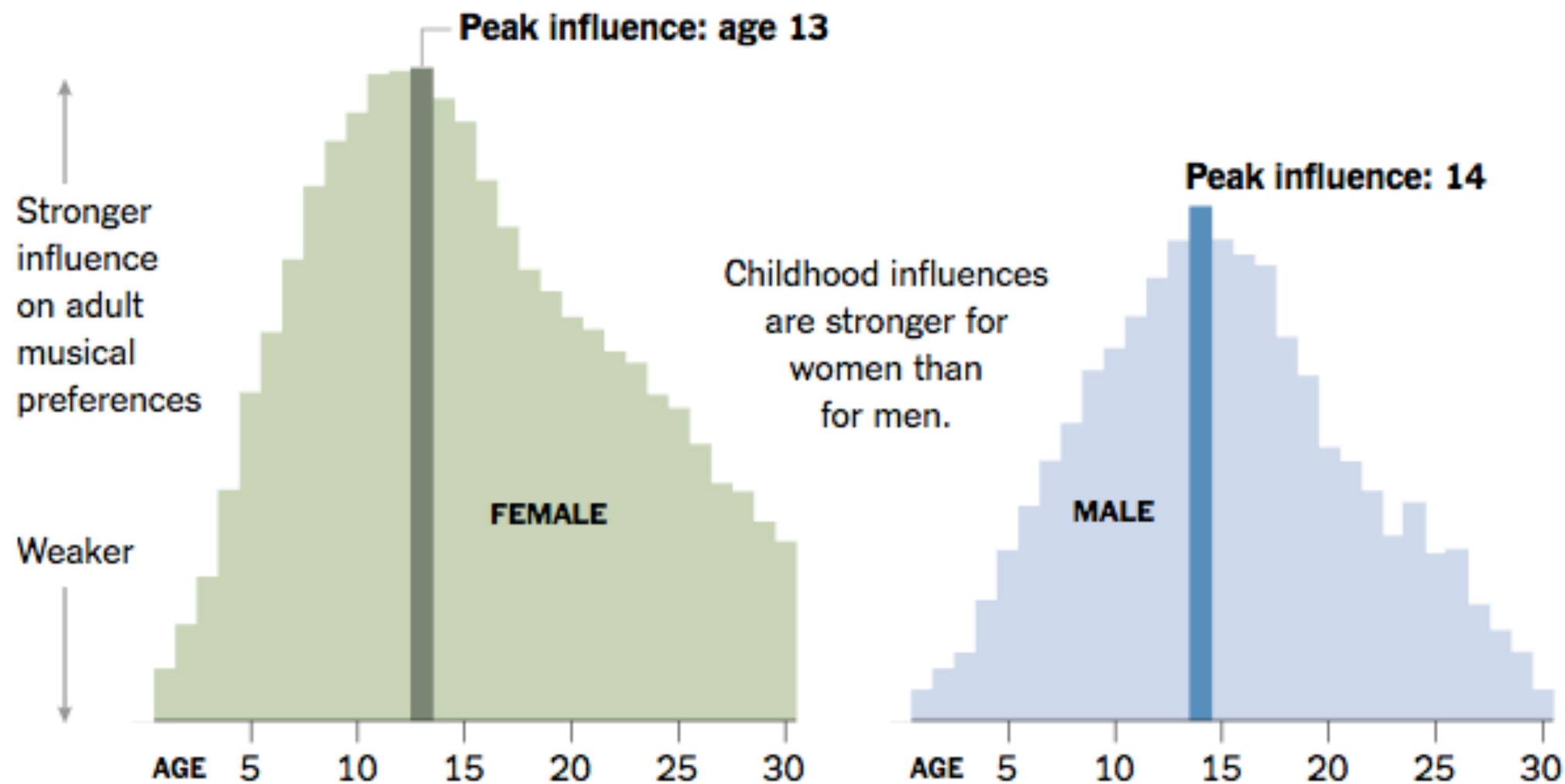


<https://ericmjl.github.io/bayesian-deep-learning-demystified>

# Aim of Bayesian machine learning

To infer the Probability Distribution Function (PDF) for your parameter given:

- your prior knowledge and
- your data

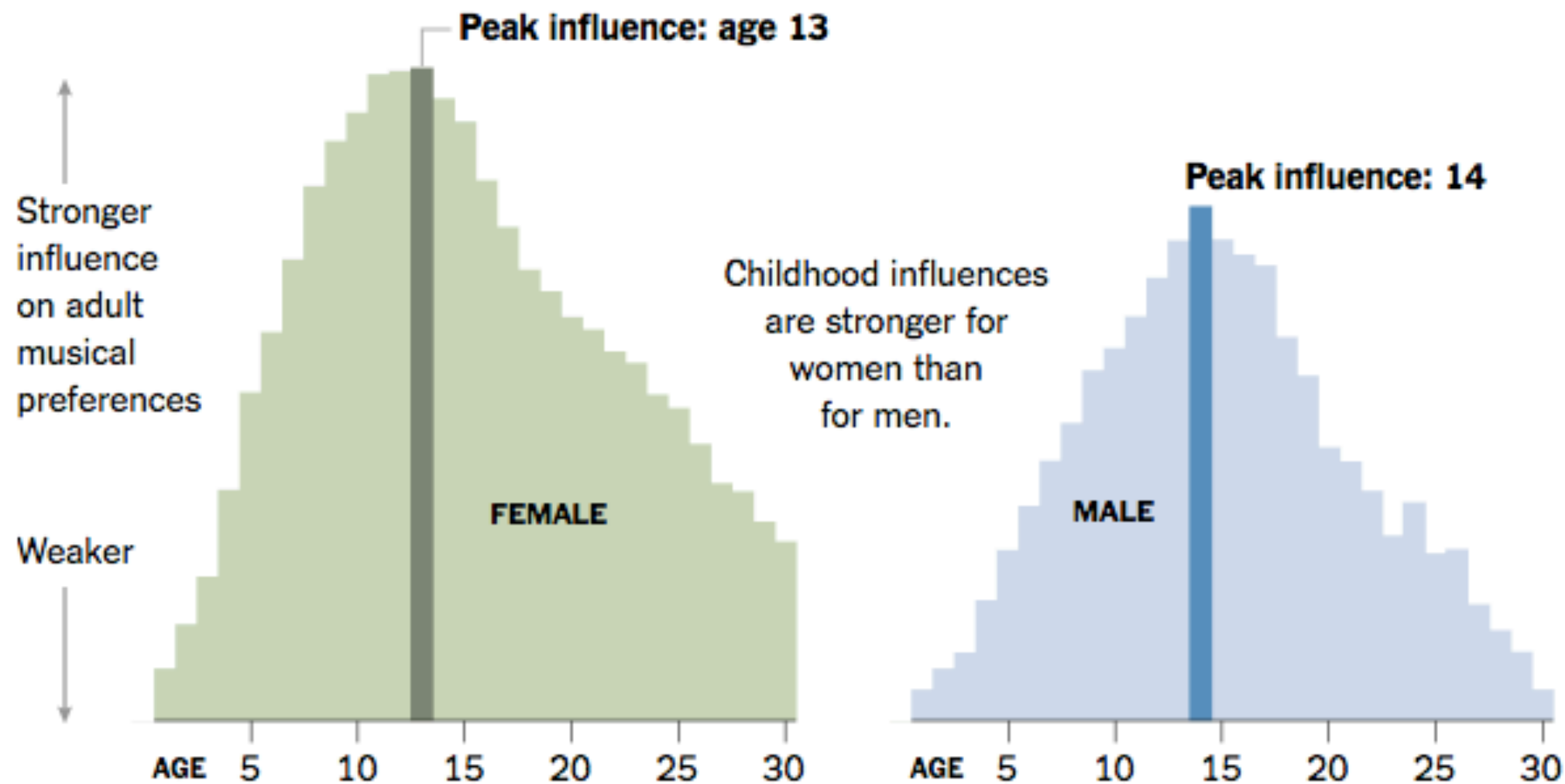


By The New York Times | Source: analysis of Spotify data by Seth Stephens-Davidowitz

Bayesian analysis helps us to think about the full distribution of data, rather than falling in the trap of just looking at the peak...

<https://www.nytimes.com/2018/02/10/opinion/sunday/favorite-songs.html>





By The New York Times | Source: analysis of Spotify data by Seth Stephens-Davidowitz

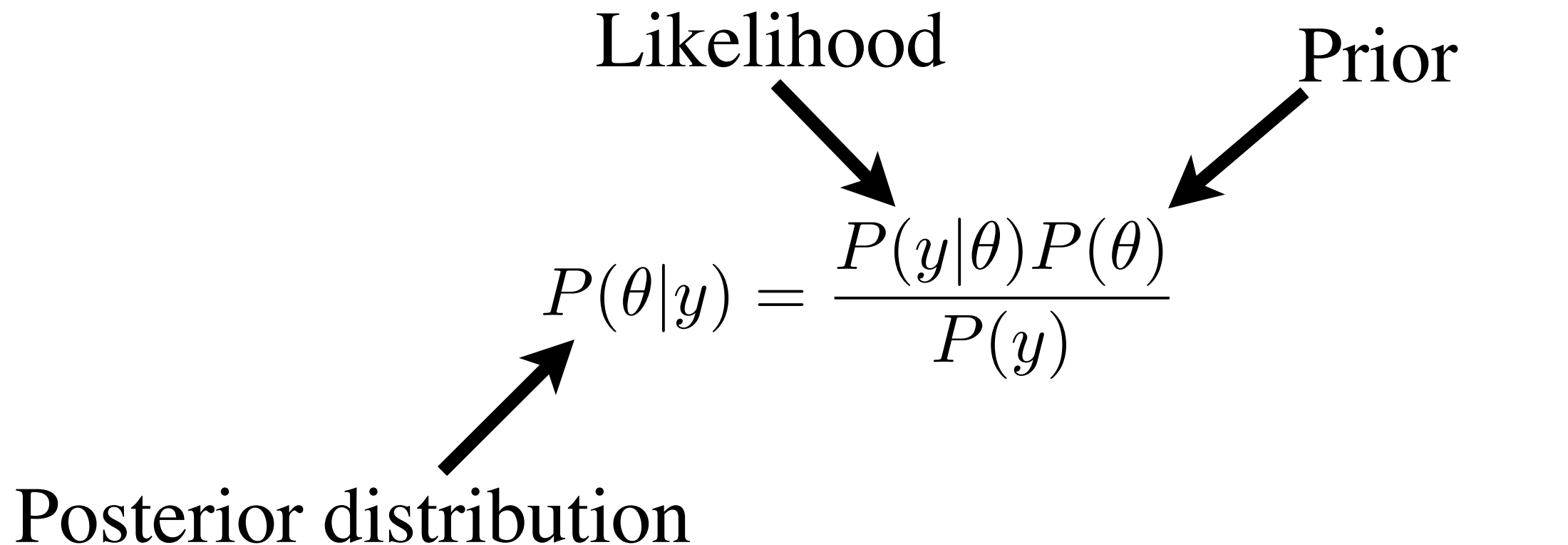
This research tells us that the majority of us, when we are grown men and women, predictably stick with the music that captured us in the earliest phase of our adolescence.

But it also adds one more piece to the central puzzle of my adult life: Why did I develop so abnormally?

<https://www.nytimes.com/2018/02/10/opinion/sunday/favorite-songs.html>

# Bayes Theorem

# Bayes Theorem is your friend



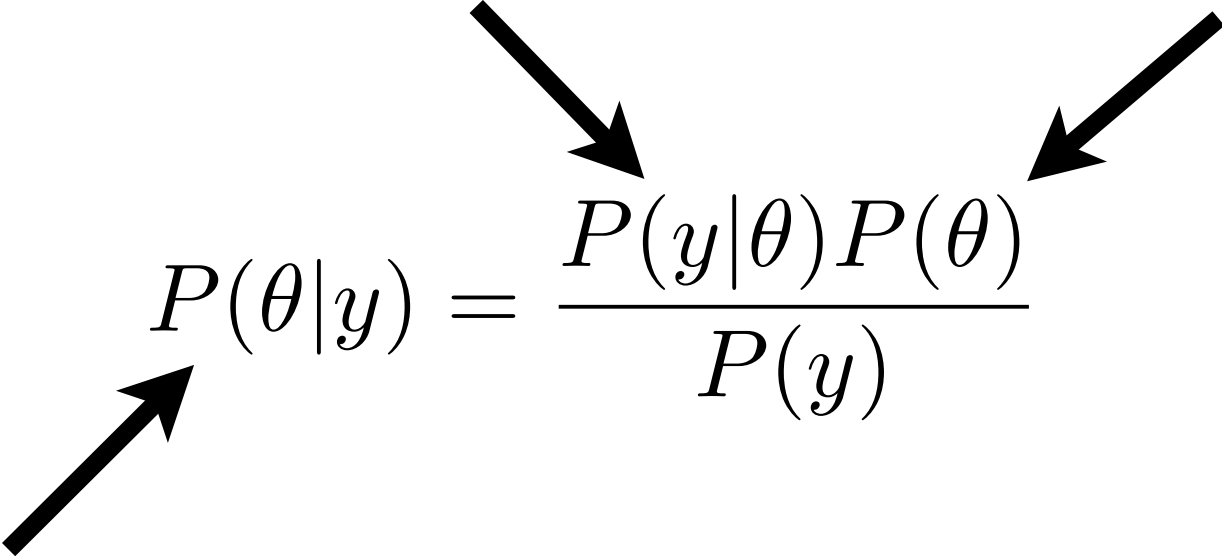
$\theta$  Parameter e.g. regression slope

$y$  Data

# Bayes theorem is your friend

What you learn  
from the data

What you  
knew beforehand


$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)}$$

What you know now

$\theta$  Parameter e.g. regression slope

$y$  Data

# Bayesian ML in python

<b>Name</b>	<b>Focus</b>	<b>Main developers</b>	<b>Backend</b>
PyStan	Statistical models	Scientists	Stan
PyMC3	Statistical models	Scientists	Theano->?
Edward	Deep learning	Google	TensorFlow
Pyro	Deep learning	Uber	PyTorch

# Examples

see [https://github.com/braaannigan/bayesian\\_intro](https://github.com/braaannigan/bayesian_intro)

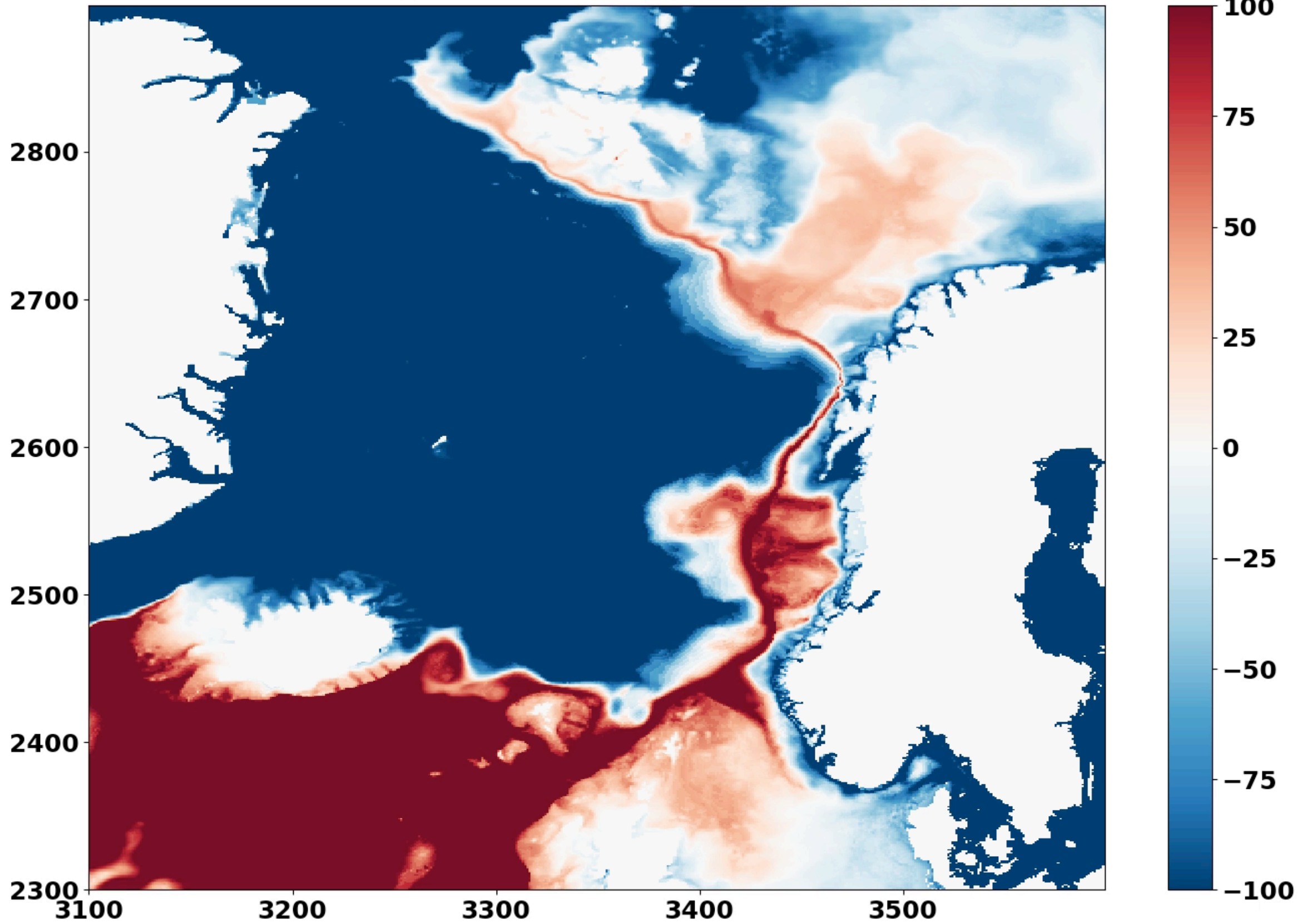
# Oceanography examples

@braaannigan

Thursday, 22 March 18



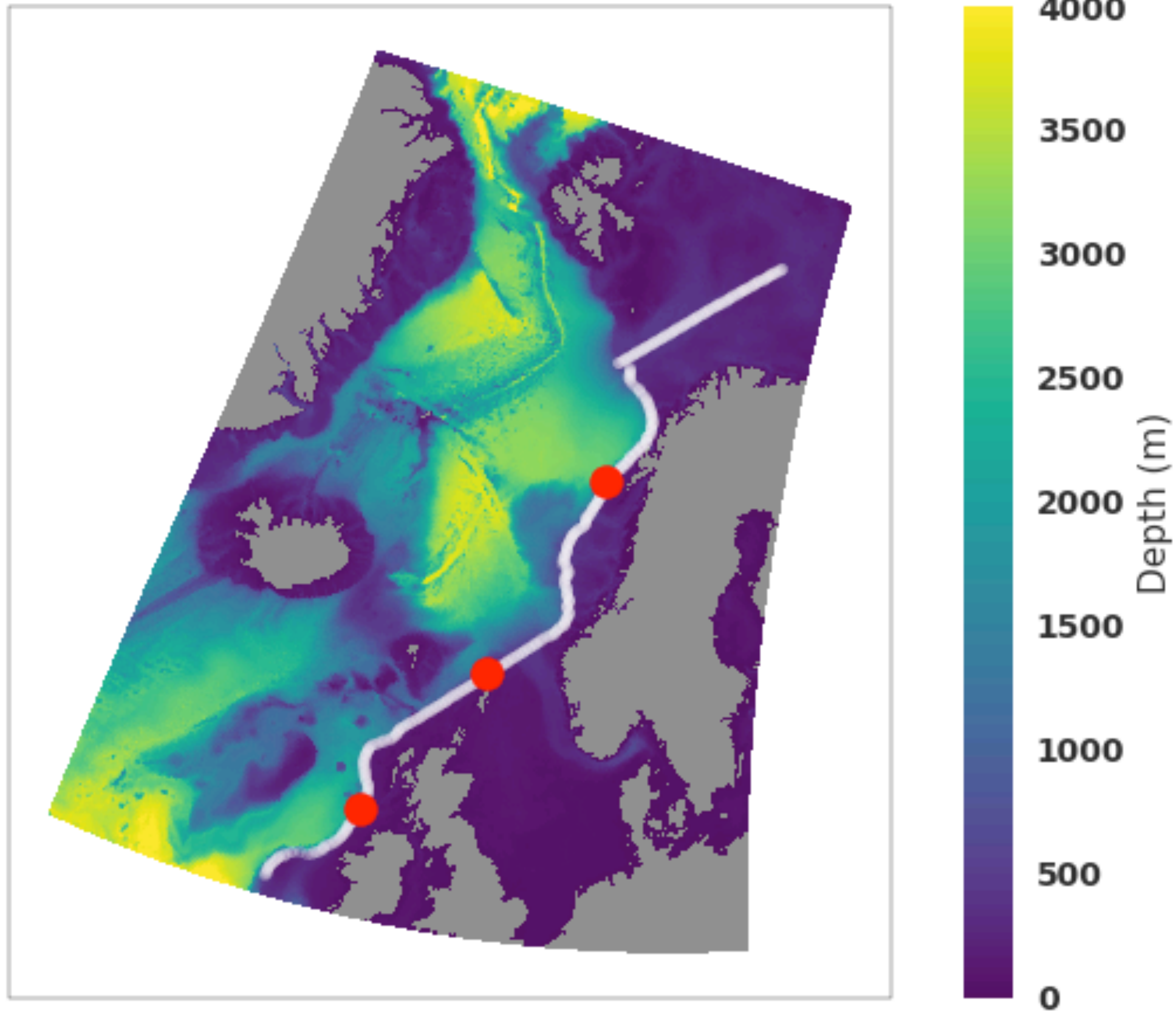
1970-01-16T13:00:00.000000000+0100



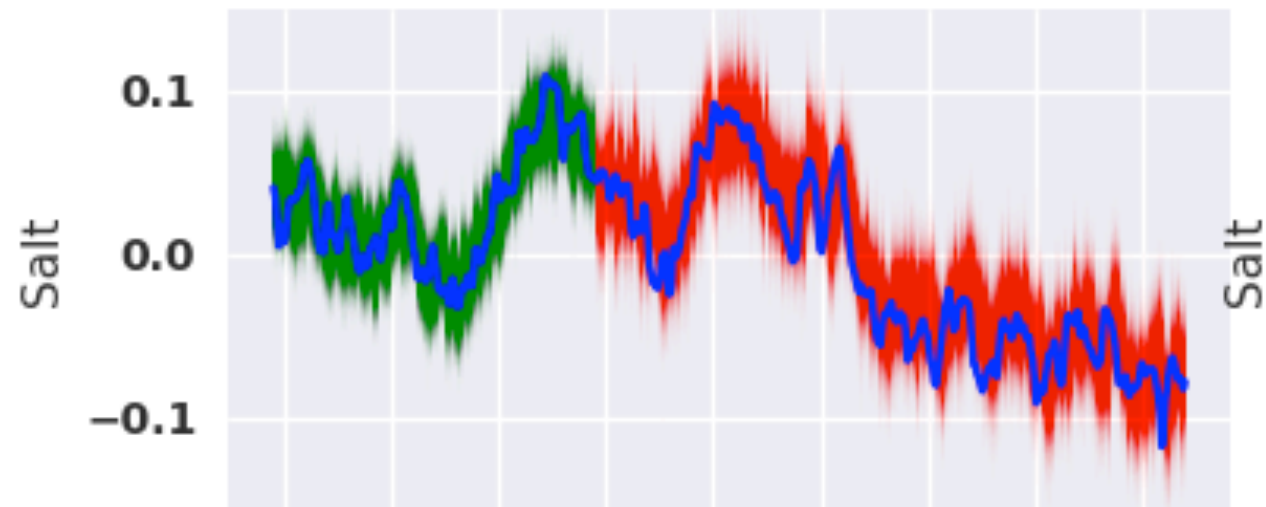
@braaannigan

Thursday, 22 March 18

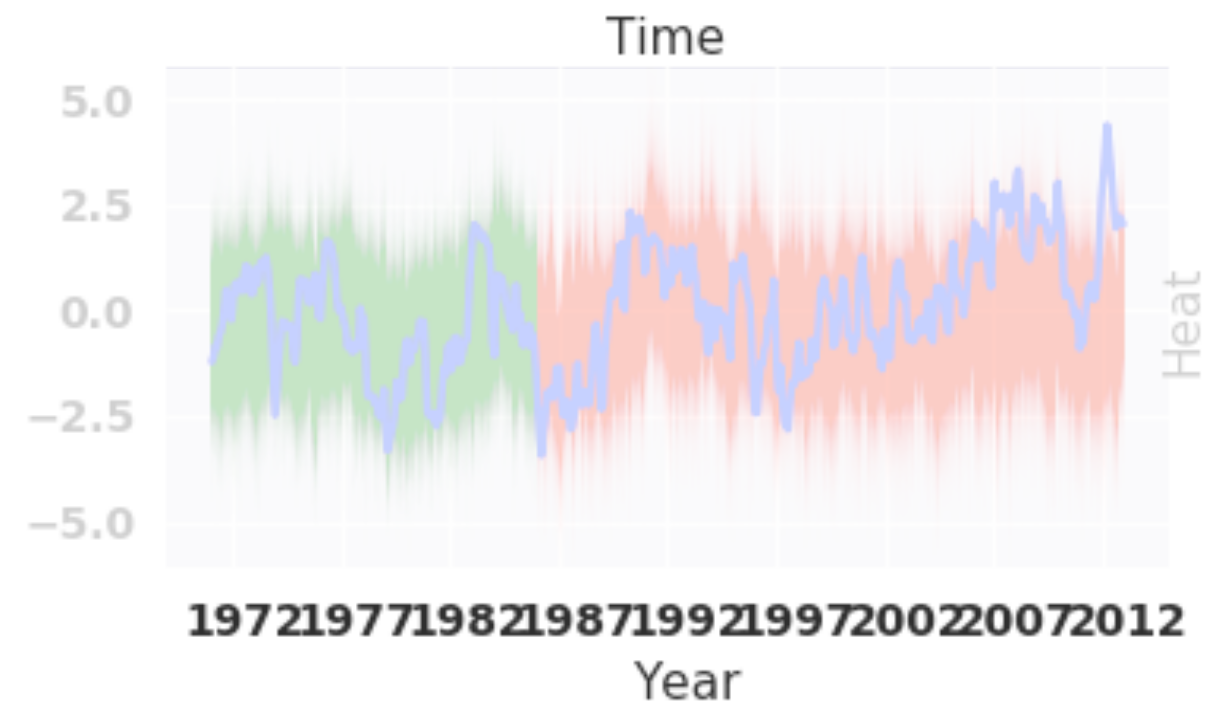
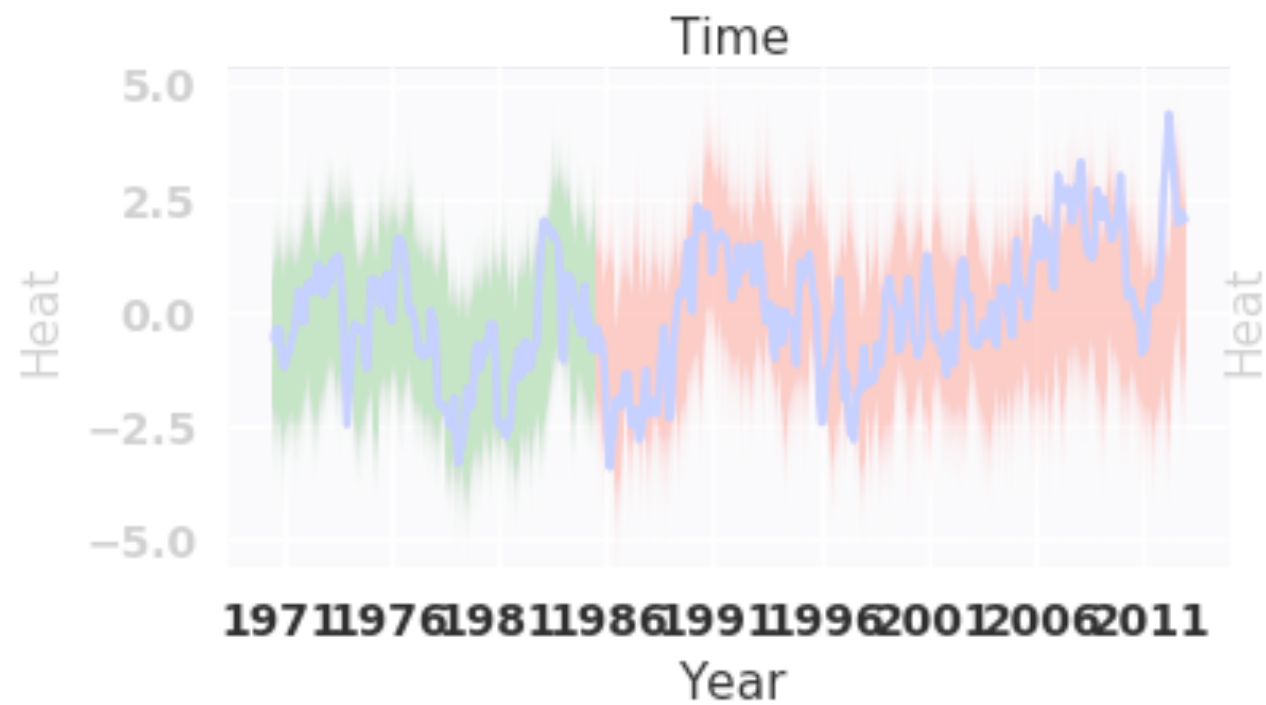
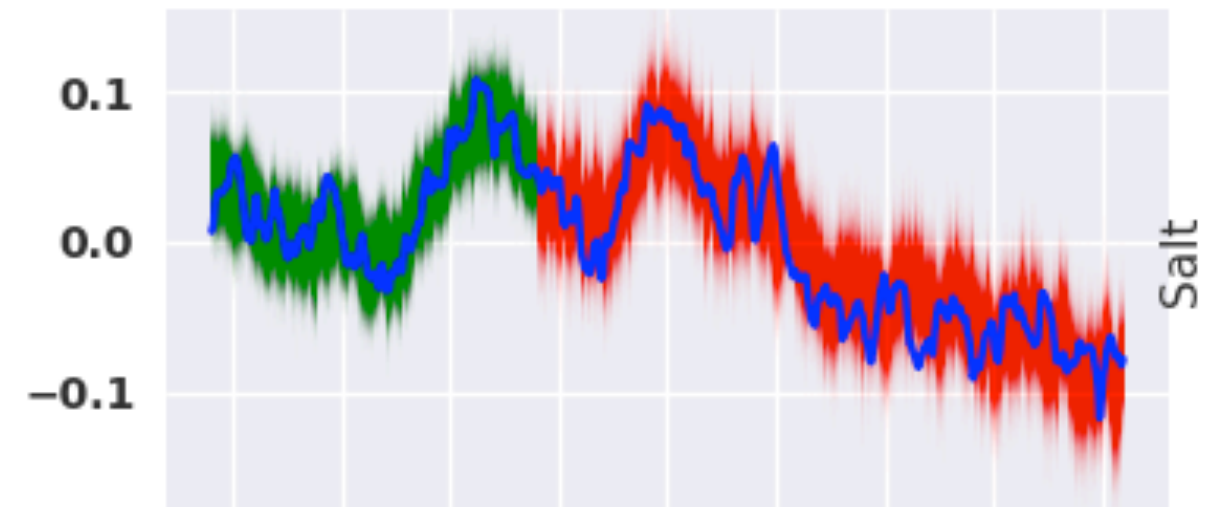
## Bathymetry with along-shelf track



Model performance with lead of 6 months

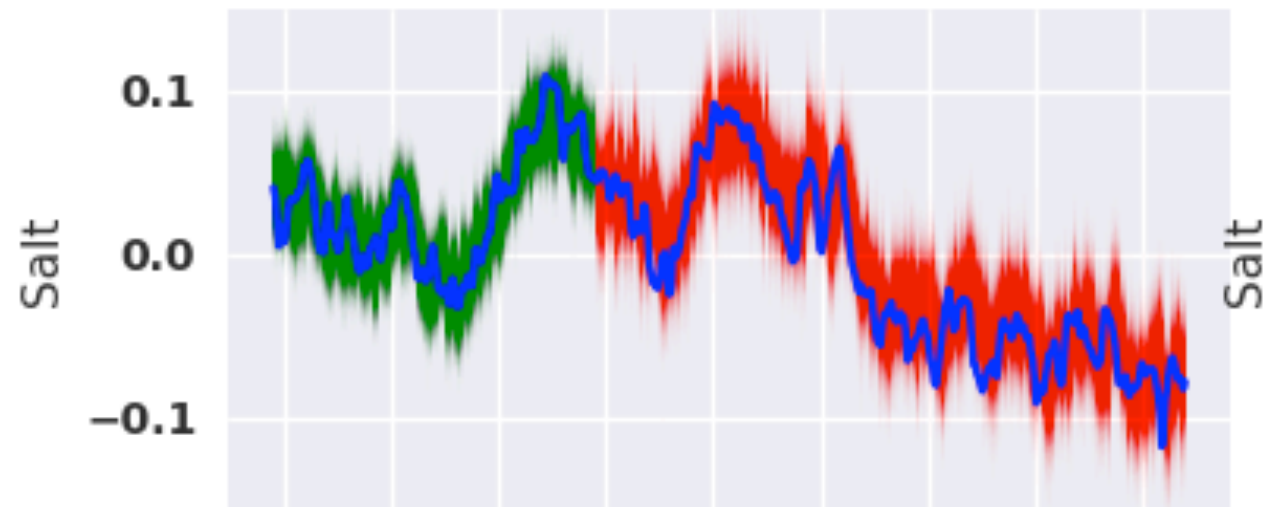


Model performance with lead of 12 months

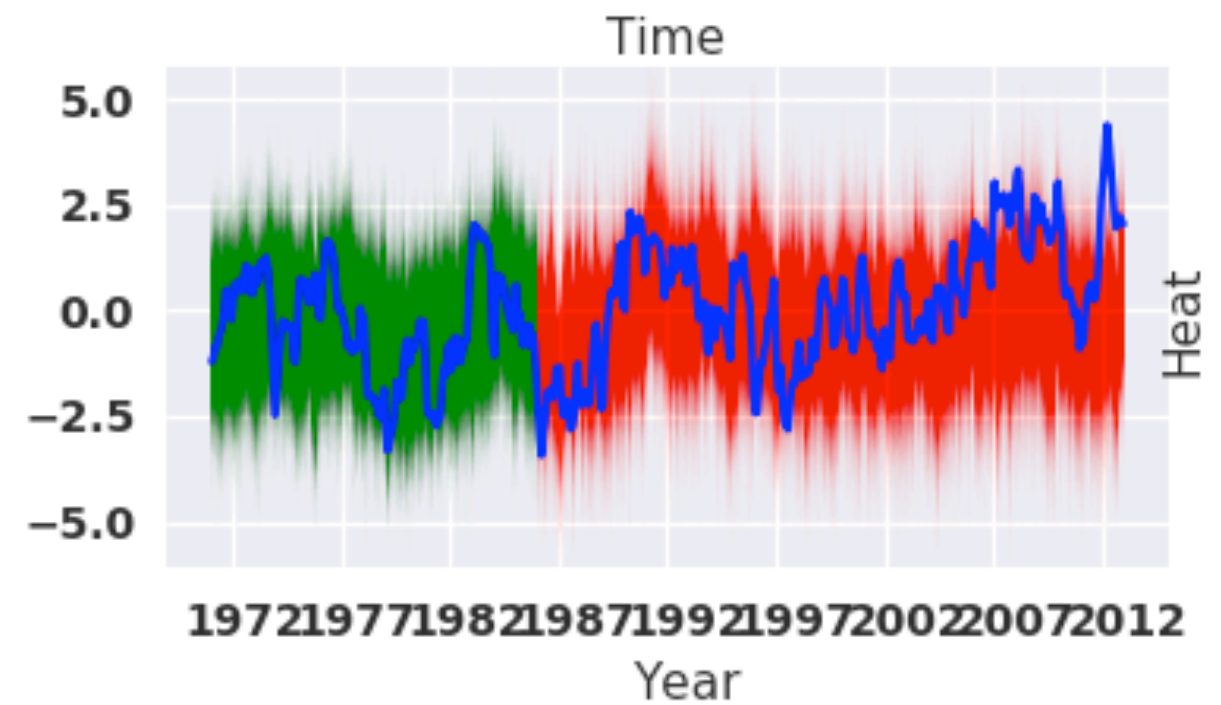
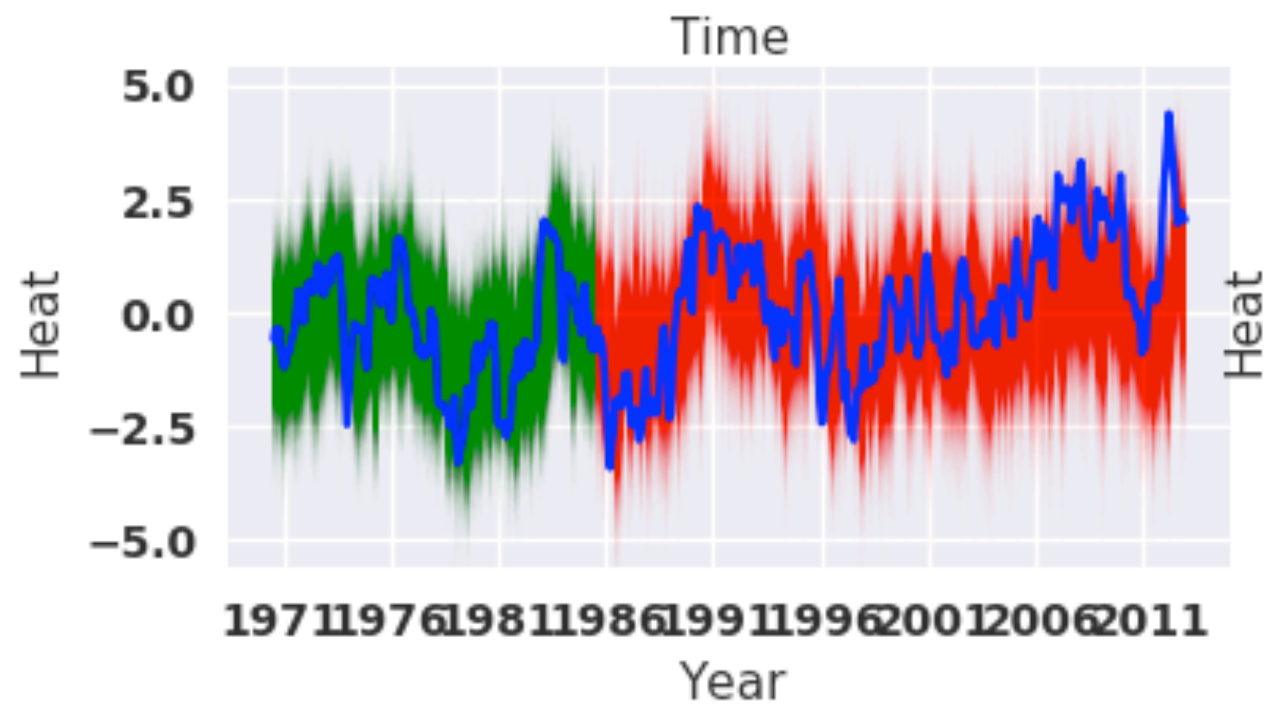
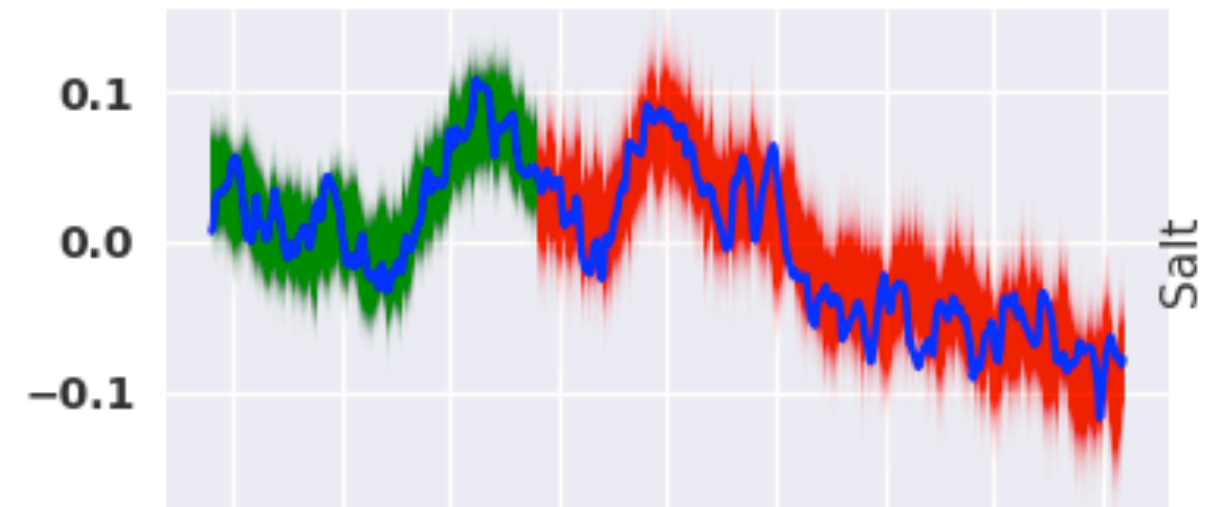




Model performance with lead of 6 months



Model performance with lead of 12 months



# Summary

- Bring prior knowledge into your ML
- Inform your analysis with uncertainty
- Numerous packages to do Bayesian analysis in python

# Next steps

Check out this tutorial:

[https://github.com/braaannigan/bayesian\\_intro](https://github.com/braaannigan/bayesian_intro)

or this textbook/lecture series

Statistical Rethinking (Richard McElreath)