Sea Ice: Data, Modelling and Uncertainty

Ben Robinson

University of Bath
Postgraduate Seminar Series

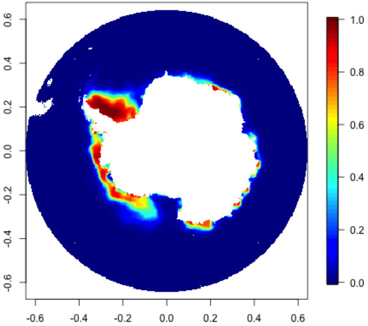
February 25, 2016

Developing a model for sea ice

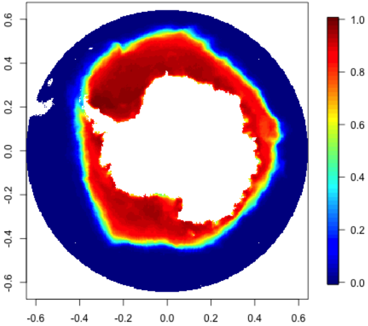
Developing a model for sea ice

The data

- Several data sources
- Most complete and reliable is data from satellites, collected by passive microwaves
- This is available from 1979-Present



(a) February mean



(b) August mean

The data

- Further back in time we only have observations of the ice edge
- Eventually only have point observations

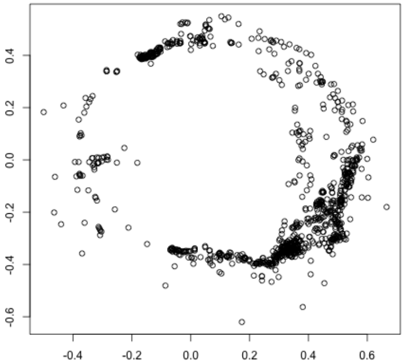


Figure: Ship observations from 1922-1953

The Problem

Our aim is to develop a model for concentrations of sea ice, which can be used to reconstruct historical sea ice concentrations and quantify uncertainty.

The Problem

Our aim is to develop a model for concentrations of sea ice, which can be used to reconstruct historical sea ice concentrations and quantify uncertainty.

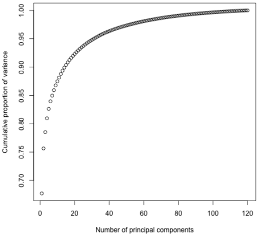
Challenges are:

- Multiple data sources, more sparse further back in time
- Need to account for
  - seasonality
  - long-time variation
  - spatial correlations
  - physical constratints
- Concentrations constrained between 0 and 1

Principal Component Analysis

Use singular value decomposition to identify main causes of variation around the mean:

- Columns of X contain ice concentrations for each month at all locations not on land, with row means subtracted
- Singular value decomposition is $X = U \Sigma V^T$
- Left singular vectors - i.e. columns of U - give principal components
- Squares of diagonal entries of Σ are proportional to variance in direction of each principal component



Principal Component Analysis

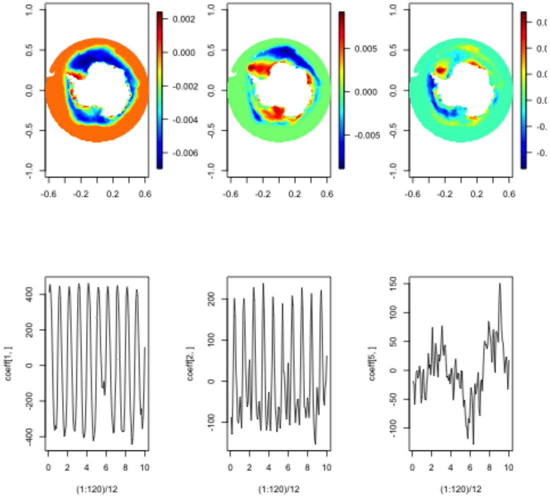


Figure: The first, second and fifth principal components.

A Simple Model

$$g(y(s,t)) = \mu(s,\tau) + U(s)\beta(t) + z(s,t) + \epsilon(s,t),$$

A Simple Model

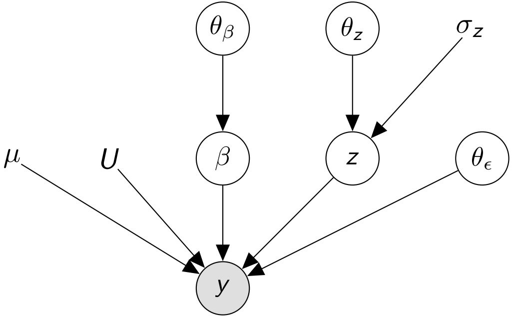
$$g(y(s,t)) = \mu(s,\tau) + U(s)\beta(t) + z(s,t) + \epsilon(s,t),$$
$$\text{Var}[z(s,t)] = \sigma_z(s,\tau)^2,$$

A Simple Model

$g(y(s,t)) = \mu(s,\tau) + U(s)\beta(t) + z(s,t) + \epsilon(s,t),$
$\text{Var}[z(s,t)] = \sigma_z(s,\tau)^2,$
where t is year and month, τ is month within year and g is defined by
$g(y) = \log(y) - \log(1-y)$, with inverse
$g^{-1}(y) = (1 + \exp(-y))^{-1} \in (0,1).$

A Simple Model

$g(y(s,t)) = \mu(s,\tau) + U(s)\beta(t) + z(s,t) + \epsilon(s,t)$,
$\text{Var}[z(s,t)] = \sigma_z(s,\tau)^2$,
where t is year and month, τ is month within year and g is defined by $g(y) = \log(y) - \log(1-y)$, with inverse
$g^{-1}(y) = (1+\exp(-y))^{-1} \in (0,1)$.



A Simple Model

Observation level:

$$g(y)|\beta, z, \theta_\varepsilon \sim \mathcal{N}(\mu, \sigma_z \Sigma_z + \Sigma_\varepsilon).$$

Latent process level:

$$z|\theta_z \sim \mathcal{N}(0, \Sigma_z).$$

- z is a Gaussian random field - at any collection of points in space, distribution is jointly Gaussian
- Uniquely determined by mean and covariance functions
- Choose covariance to be:
  - Stationary - only depends on relative position of two points
  - Isotropic - only depends on Euclidean distance between two points

Gaussian Random Field

Simulate stationary isotropic gaussian random field, using SPDE model in INLA:

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} z(u) = \mathcal{W}(u).$$

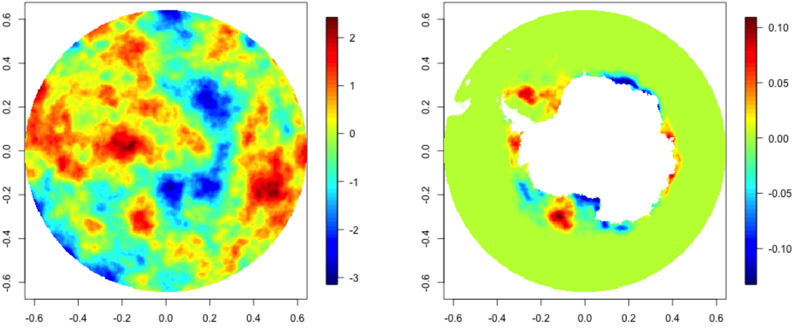
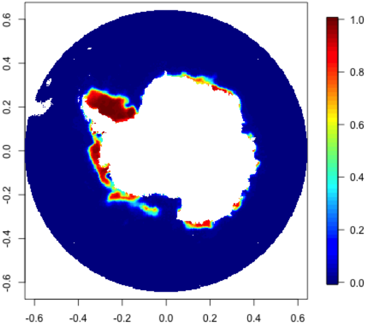
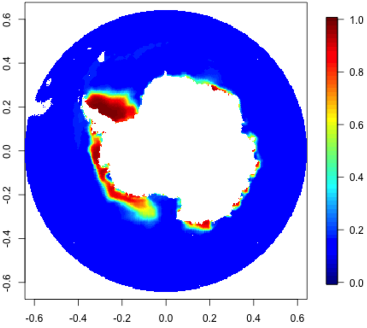


Figure: Stationary isotropic GRF (left) and GRF multiplied by standard deviation at each point in space (right)

February



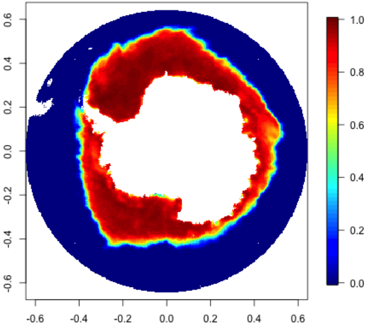
(a) Observed sea ice concentrations February 1984.



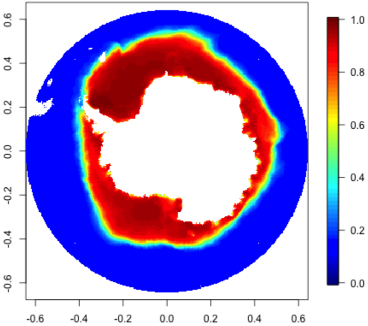
(b) Simulated February sea ice concentrations.

Figure: Southern hemisphere sea ice concentrations for February.

August



(a) Observed sea ice concentrations August 1984.



(b) Simulated August sea ice concentrations.

Figure: Southern hemisphere sea ice concentrations for August.

Changing the model

- Current model is unable to predict changes in ice edge different from the data set
- Change what we model to concentration at a distance from coastline
- Express the distance from the coastline d as a function of the ice concentration c and the point on the coastline x; e.g. $d(x, c) = \sup\{d : (\text{ice concentration at the point } x + d \cdot n) \geq c\}$.
- Coastline of Antarctica is not convex, so need to look at physical properties to determine direction

The Arctic

- Geographical features around the Arctic are more complicated:

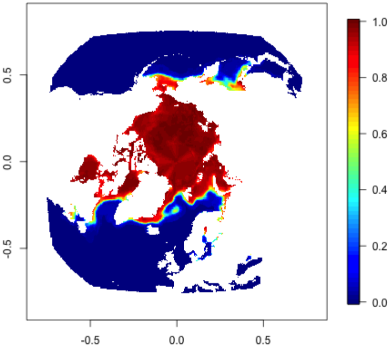
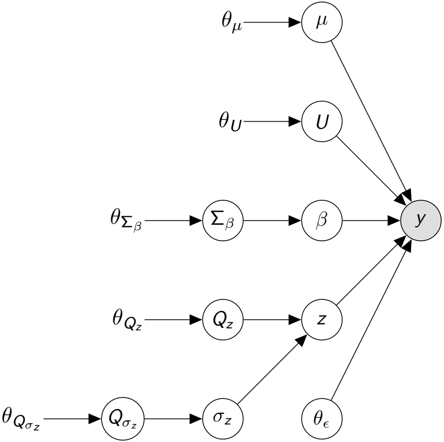


Figure: Observed ice concentrations in the northern hemisphere January 1981, where white space indicates land.

Improving the model



- Check model by using one decade of data to predict another, and vice versa
- Assign prior distributions to parameters and use Bayesian inference
- Model no longer jointly Gaussian means new computational challenges