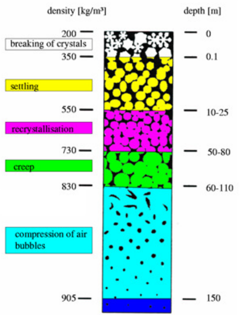
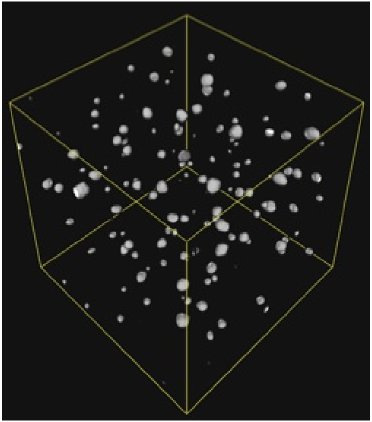
Spatial point processes: anisotropy, 3D and missing information

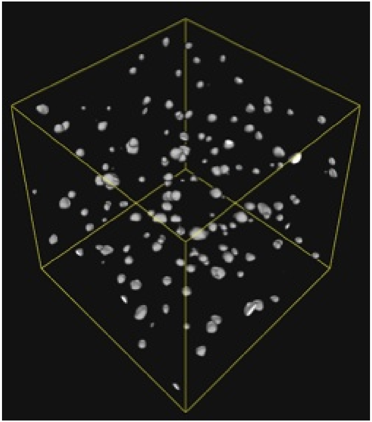
Problem

- Polar ice has information on the climate of the past
- To be able to interprete the ice core records, one has to know how old the ice is
- Question: How can we estimate the deformation history in polar ice?
- Method: Polar ice is compacted snow. If we go deep enough, the air pores are isolated in the ice.
  → Study the anisotropy (deformation) of these air inclusions in the ice samples from a deep ice core at different depths.



System of air pores in ice samples from depth 353m (left) and 505m (right)





Assumptions

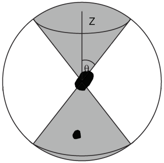
- We concentrate on pore centers (not shape of the pores).
- Pore centers a realization of a regular process
- The ice is compressed in z direction and streched in the lateral direction leading to a specific type of anisotropy of the air pores.
- Volume preserving compression

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \frac{x}{\sqrt{c}} \\ \frac{y}{\sqrt{c}} \\ cz \end{pmatrix}, 0 < c < 1$$

where the factor c gives the total "thinning" of the ice

Which directions to look at?

- Anisotropy in z-direction
  $\rightarrow$ a complete partitioning of the ball is not necessary
- Investigate the point patterns in double cones aligned along coordinate axes
- Compare observations (in z direction and in x and y direction)
- Large differences indicate anisotropies



Directional summary statistics in 3D:

- Directional Ripley's K function K_{dir}: Expected number of pores within the double cone C centred in a typical pore divided by the pore intensity
- Directional nearest neighbour distance function G_{dir}: Cumulative distribution function of the distance between a typical pore and its nearest neighbour in the cone centred at the core

Isotropy tests

- $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ estimators of one of the summary statistics introduced above with respect to the x-, y-, and z-direction.
- In the isotropic case, all three estimates will look similar. For the pressed pattern only $\hat{S}_x$ and $\hat{S}_y$ should be similar but show a clear deviation from $\hat{S}_z$.
- n point patterns $y_1, \ldots, y_n$

Isotropy tests

A test can be based on a comparison of the test statistics

$$T_{xy,i} = \int_{r_1}^{r_2} |\hat{S}_{x,i}(r) - \hat{S}_{y,i}(r)| \, dr,$$

and

$$T_{z,i} = \min \left(\int_{r_1}^{r_2} |\hat{S}_{x,i}(r) - \hat{S}_{z,i}(r)| \, dr, \int_{r_1}^{r_2} |\hat{S}_{y,i}(r) - \hat{S}_{z,i}(r)| \, dr \right),$$

where $[r_1, r_2]$ is a given interval.

The isotropy hypothesis for a certain sample y_i is rejected at significance level α if the corresponding value $T_{z,i}$ is larger than $100(1-\alpha)\%$ of the estimated $T_{xy,i}$ values.

Alternatively a Monte Carlo test

Estimation of the pressing factor

- Rescale the pattern by $(\sqrt{d}, \sqrt{d}, \frac{1}{d})$ (with $d \in [0.6, 1.1]$)
- Compute

$$T_{\sum,d} = \int_{r_1}^{r_2} \Big(|\hat{S}_{x,d}(r) - \hat{S}_{z,d}(r)| + |\hat{S}_{y,d}(r) - \hat{S}_{z,d}(r)| + |\hat{S}_{x,d}(r) - \hat{S}_{y,d}(r)| \Big)\, dr$$

  for the pattern rescaled by the factor d. Here, $\hat{S}_{x,d}$, $\hat{S}_{y,d}$, and $\hat{S}_{z,d}$ are estimators of one of the two summary statistics with respect to the x, y, and z direction.
- The pressing factor is then estimated by

$$\hat{c} = \operatorname{argmin}_d T_{\sum,d}$$

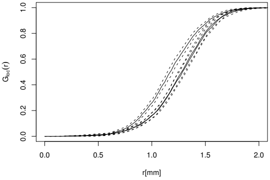
  i.e. the value of d leading to the most isotropic pattern

Ice data

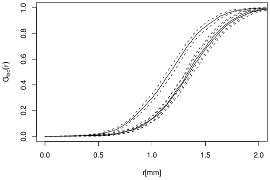
- Tomographic images of ice samples from Antarctica (imaged inside a cold room at -15° C (5° F))
- A sample is a cylinder which has height 15mm and diameter 15mm
- Ice samples from depths 153m, 353m and 505m
- 14 samples at each depth
- Number of air pores per sample is between 329 and 733

Ice data: unisotropy measured by G_{dir}

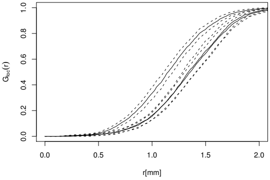
Pooled G_{loc} estimated per depth and pointwise confidence bands (95%) using the sample variances and quantiles from the T-distribution



153m



353m



505m

Ice data: estimated pressing factors

Compression stronger deeper down

- 153m: 0.810 (based on G_{loc}) and 0.820 (based on K_{dir})
- 353m: 0.630 (based on G_{loc}) and 0.641 (based on K_{dir})
- 505m: 0.534 (based on G_{loc}) and 0.545 (based on K_{dir})

New data

- New CT device especially designed for ice cores
- Two samples: 150m ($2.8 \times 2.8 \times 52.5$ cm^3, 190.000 pores) and 951m ($2.8 \times 2.8 \times 59$ cm^3, 250.000 pores), size approximately
- Samples contain relaxation (extra) bubbles
  1. lattice defects caused by inclusions or pollutants in the ice
     - smaller than real bubbles
  2. clathrate inclusions: air comes from the existing bubbles which were pressed into the crystalline lattice of the ice
     - similar size range as the real bubbles have
     - only in depths below 800-900 m

  that do not give any information on the motion of the ice

Problem/question

Since the extra bubbles do not carry any information on the motion of the ice and disturb the directional analysis, we would like to remove them before doing the analysis

Is it possible to classify each bubble either as "real pore" or as noise? (missing information)

We (*) will adapt the Bayesian MCMC approach by Walsh and Raftery (2002) for our situation

(*) Claudia Redenbach, Johannes Freitag, Martina Sormani, Adrian Raftery, Tuomas Rajala

Walsh and Raftery (2002)

How to distinguish between mines and noise?

- Poisson noise, mines located approximately in parallel rows
- MCMC algorithm to estimate the model and to obtain posterior probabilities for each point to be a mine (points classified)

Instead of mines in parallel rows surrounded by Poisson noise, we have real bubbles located according to a regular process surrounded by Poisson noise

Model assumptions

- Real pores modeled by Strauss process (regular process) with density
$$f(x) = \alpha\beta^{n(x)}\gamma^{s(x)},$$
where $\beta > 0$ is the "intensity" parameter, $0 < \gamma \leq 1$ the strength of interaction, $n(x)$ the number of points in the realization, $s(x)$ the number of r close pairs of points, and r the interaction radius
- Poisson process with intensity λ_0 for noise bubbles
- The complete point process Y is a superposition of the Strauss process Y_1 and the Poisson process Y_0
- Parameter vector $\theta = (\lambda_0, \beta, \gamma, r)$

Model assumptions

- The data consist of n bubbles, n_1 Strauss (real) bubbles and n_0 Poisson (noise) bubbles
- $Z \in \{0, 1\}^n$ is a random vector such that

$$Z_i = \begin{cases} 1 & \text{if the } i\text{th bubble belongs to the Strauss pattern} \\ 0 & \text{if the } i\text{th bubble belongs to the Poisson pattern} \end{cases}$$

Aims

1. estimate the vector Z (classification into real (Strauss) and noise (Poisson) bubbles)
2. estimate the vector of parameters $\theta = (\lambda_0, \beta, \gamma, r)$

Bayesian approach: prior distributions

Uniform priors to the parameters in θ and assume that they are independent

Prior for Z

$$\pi(Z|\theta, N, A) = \pi(n_0|\theta, N, A) = \binom{N}{n_0}\left(\frac{\lambda_0}{\lambda_0+\lambda_1}\right)^{n_0}\left(\frac{\lambda_1}{\lambda_0+\lambda_1}\right)^{n_1},$$

where λ_0 is the intensity of the Poisson process and λ_1 the intensity of the Strauss process, which can be approximated by a function of the parameters β, γ and r (Baddeley and Nair, 2012).

MCMC algorithm

- Select initial parameters according to the prior distributions
- Iterate the following steps
  - Update θ: Pick one of the parameters θ_i in θ at random and propose to change it to $\theta_i' = \theta_i \exp(X)$, where $X \sim N(0, \tau^2)$.
  - Update Z: Pick one data point at random and propose to move it to the other process. Accept with the probability given by the Hastings ratio.

Simulation study

- We started by estimating Z when $\theta = (\lambda_0, \beta, \gamma, r)$ is known
- Simulation study in the unit square (2D)
- A bubble is declared "true" if the estimated posterior probability is "high".

First results

- Approximately 25% of the bubbles are misclassified
- Often, the misclassified bubbles are close pairs of bubbles, where one is a real bubble and the other is a noise bubble
- Therefore, the spatial pattern of real bubbles remains the same (which is important given the anisotropy analysis that we would like to perform).

What remains?

- Simultaneous estimation of Z and θ
- Go from 2D to 3D
- Does this work for the ice data? We believe that
  - 20-30% of the bubbles are noise bubbles
  - we may be able to estimate the interaction radius r of the Strauss process
- Should size information be included?
- Anisotropy analysis of the ice after the noise bubbles have been removed

To conclude

So far, we have

- taken into account anisotropy (old data)
- taken care of missing information (on going)

Some further challenges

- There may be some additional anisotropy in the xy plane due to flow
  $\rightarrow$ A new type of transformation may be needed
- Data shows some layering of the ice (intensity of the air pores varies from layer to layer) due to seasonal changes
  $\rightarrow$ Inhomogeneous models

References

- Redenbach, C., Särkkä, A., Freitag, J. and Schladitz, K. Anisotropy analysis of pressed point processes. *Advances in Statistical Analysis* 93(3) (2009) 237-261.
- Walsh, D.C.I., Raftery, A.E. Detecting mines in minefields with linear characteristics. *Technometrics* 44 (2002) 34-44.