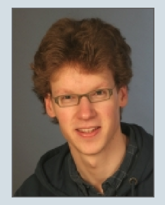
MADALGO – CENTER FOR MASSIVE DATA ALGORITHMICS



Freek van Walderveen
Aarhus University

Danmarks Grundforskningsfond
Danish National Research Foundation

Cleaning Sonar Terrain Data

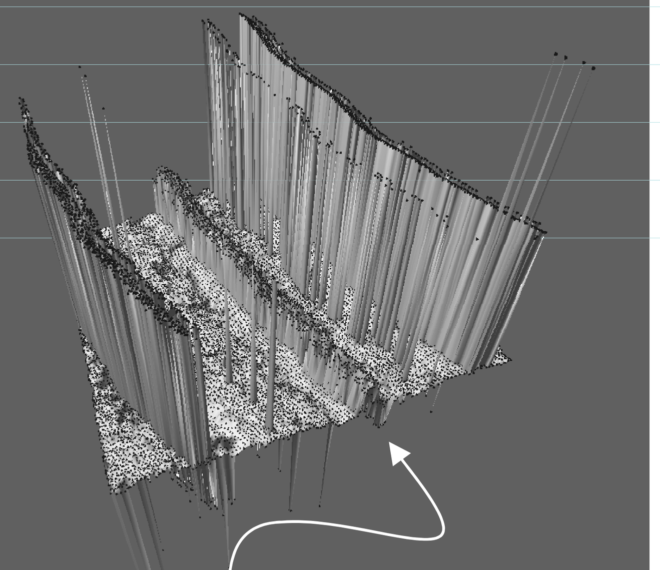
Sonar Scans of the Seabed

Why?

- Creating height maps, nautical depth charts.
- Terrain analysis: search for oil, inspection of pipelines.

How?

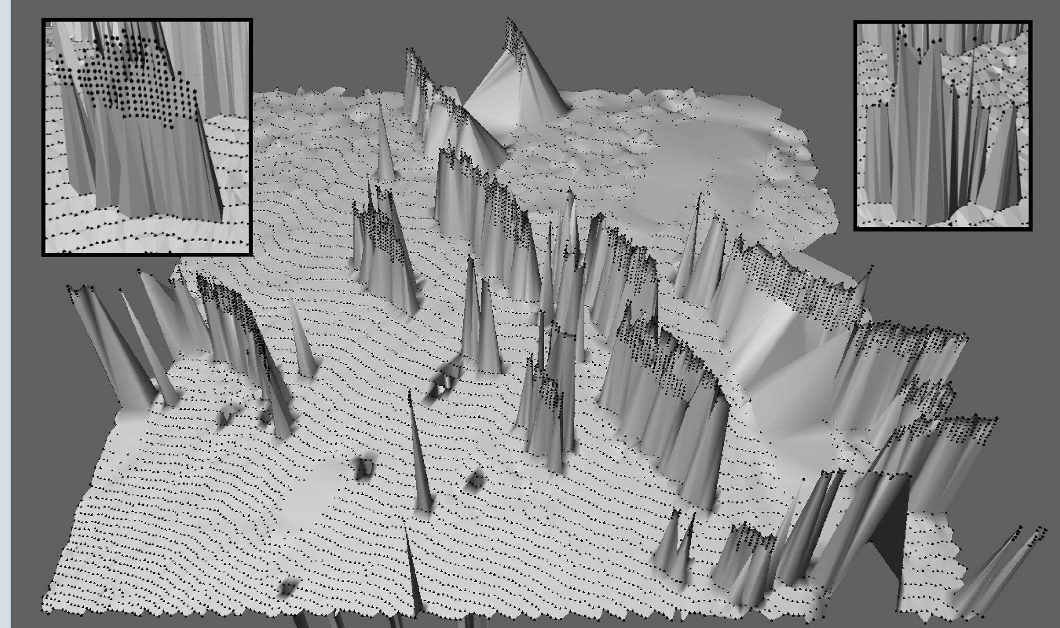
- LIDAR for land and shallow waters.
- Multibeam echo sounders (MBES).
  - Mounted under survey vessels.
  - Measures depth using sonar.
  - Each measurement produces a 3D point.
  - High resolution ⇒ large data sets.



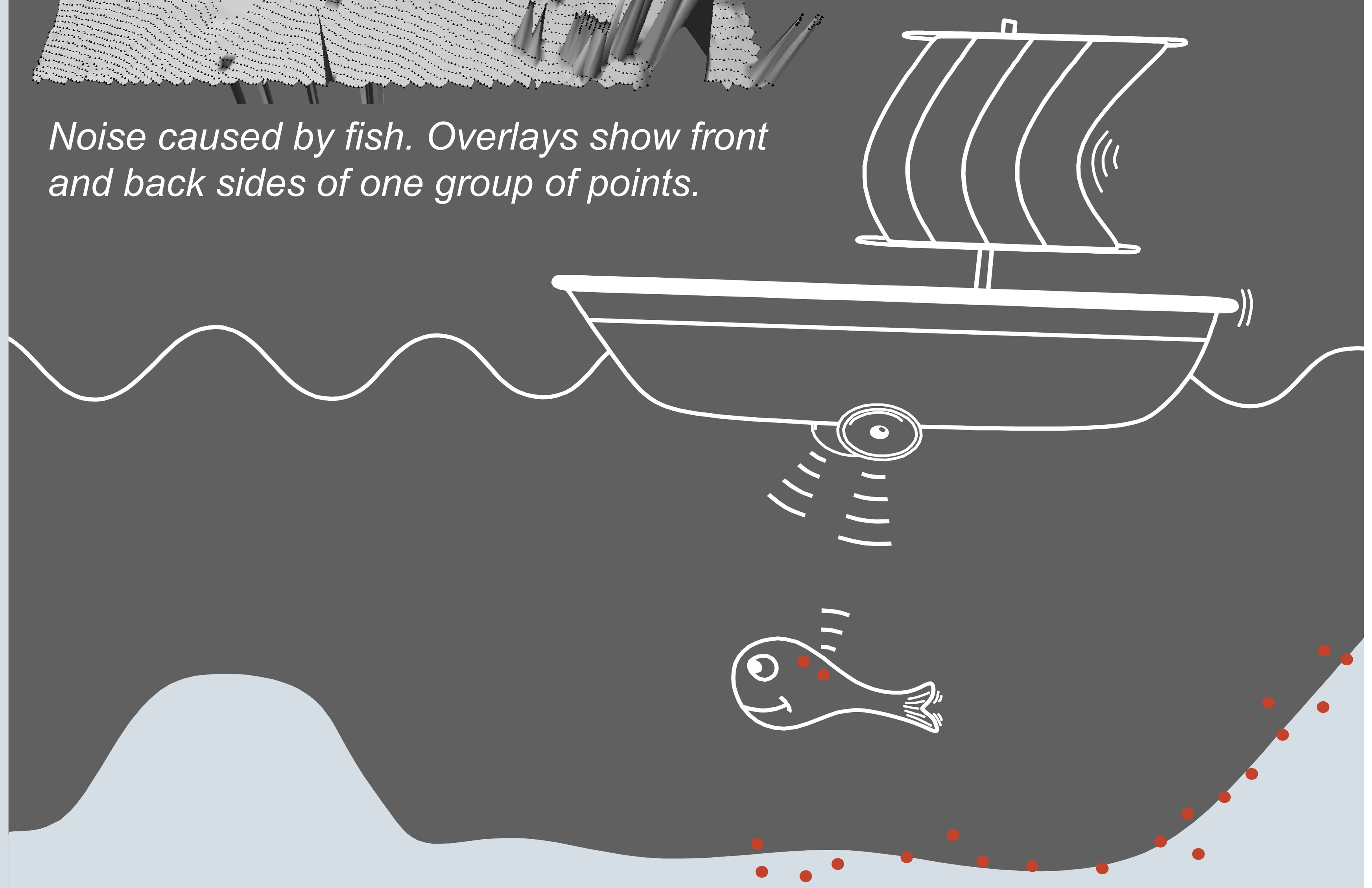
A pipeline surrounded by two ribbons of structural noise.

What's the problem?

- Not only features on the terrain are reported.
  - Data includes fish and other non-permanent objects in the water.
- Spurious measurements.
  - E.g. refraction in gas bubbles.
- Miscalibration of devices causes systematic errors.
- Sensor inaccuracy.



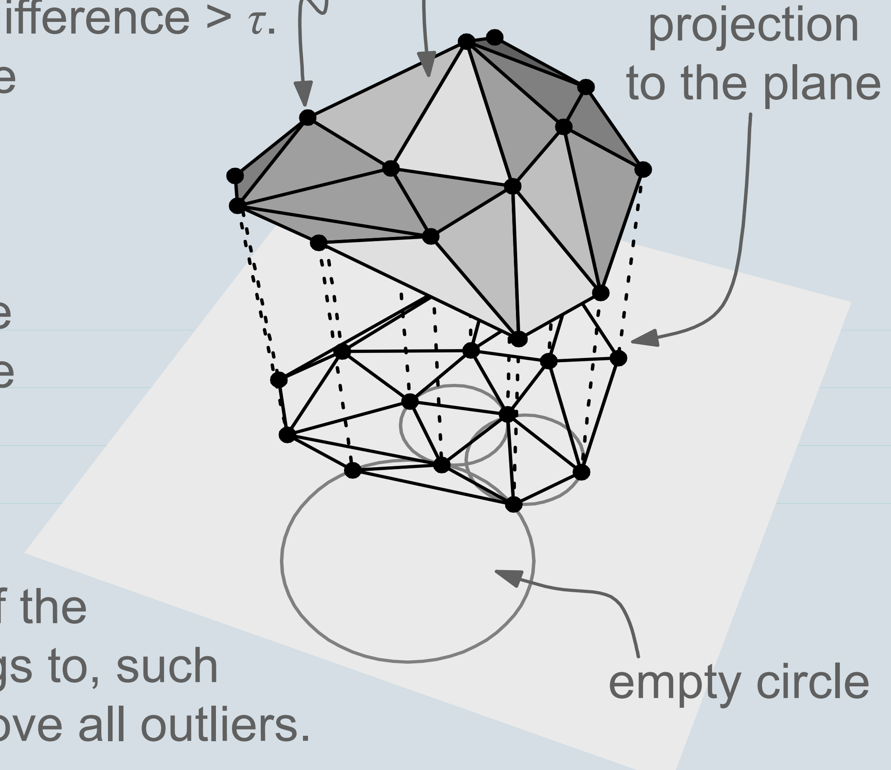
Noise caused by fish. Overlays show front and back sides of one group of points.



Our Solution

Our solution is based on a triangulation that considers only the horizontal coordinates of the 3D points (see picture below). This allows for an efficient implementation, while the edges between the points represent a kind of closeness in horizontal direction. Therefore, if two points are connected via an edge while they are far apart, say more than a distance τ in the vertical direction, they are probably part of different surfaces. From a high level, our algorithm then works as follows.

- Compute 2D triangulation.
- From the triangulation remove all edges whose endpoints have height difference > τ.
- In the resulting graph, find large connected components.

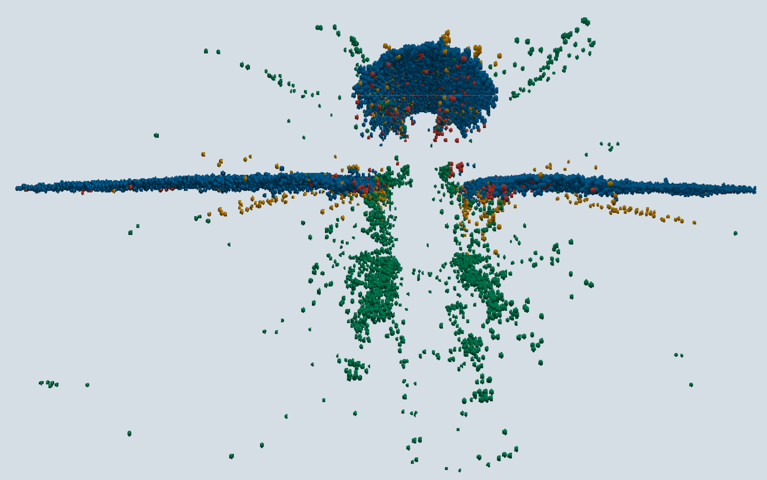


Now we have made a separation between points belonging to large surfaces and those that are single outliers or points belonging to small surfaces.

- Assign to each point the size of the connected component it belongs to, such that the user can instantly remove all outliers.

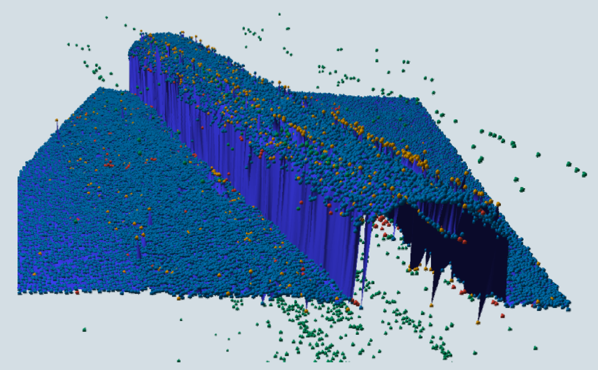
All of the above steps are implementable in such a way that they use the hard disk efficiently. Therefore, our algorithm scales well and is able to handle large data sets.

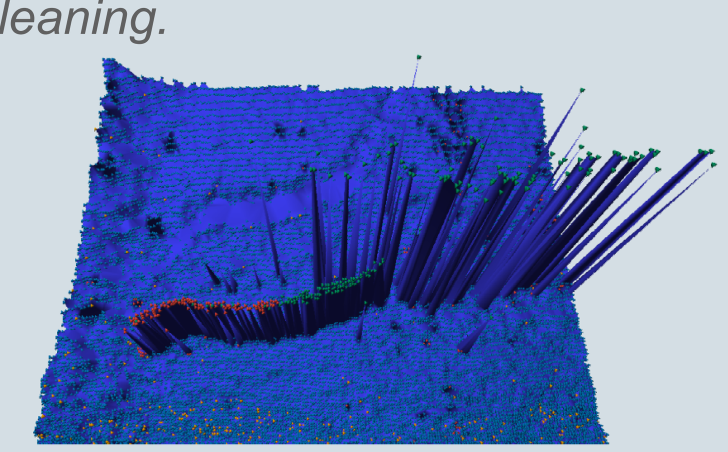
Our Method Compared to Manual Cleaning



- *Blue: reconstructed surface.*
- *Green: points removed by both manual cleaning and our algorithm.*
- *Red: points kept during manual cleaning, but removed by our algorithm.*
- *Orange: points removed by our algorithm, but kept during manual cleaning.*

Top: Cross-section of a pipeline. Bottom: Same pipeline in perspective.

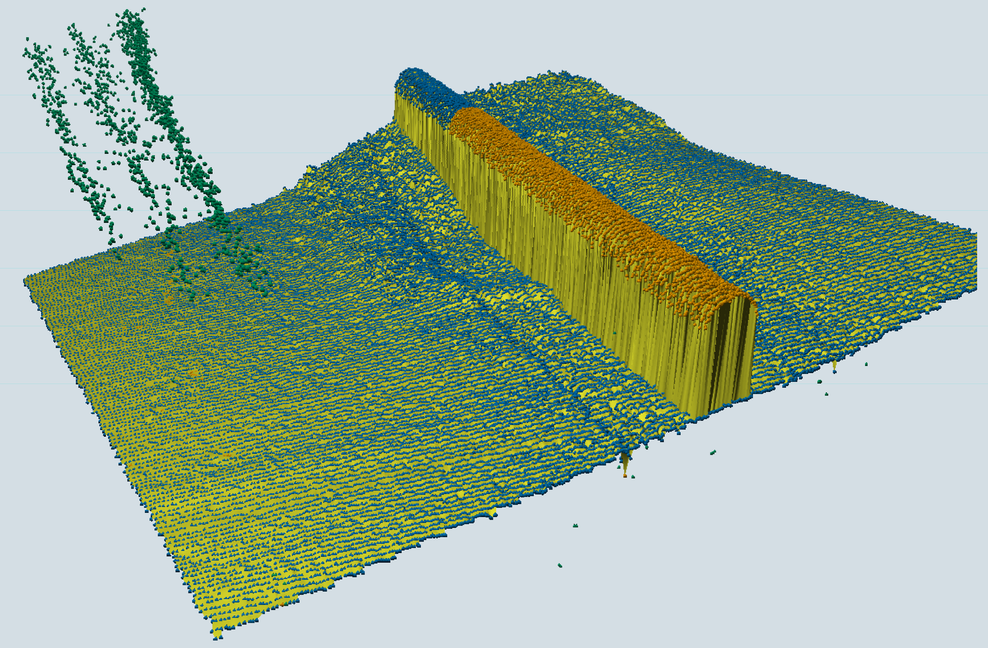




Red points: overlooked by operator?

Dealing with Overhangs

Real sonar data includes over-hangs: situations where one surface appears above another one, allowing the sonar to obtain points on both surfaces.

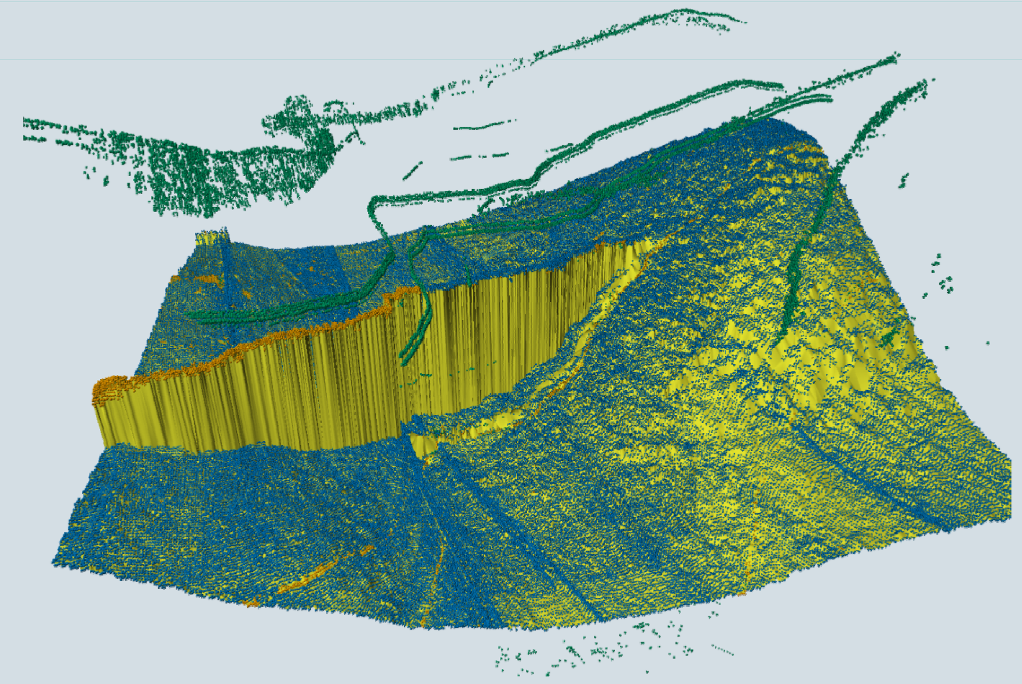


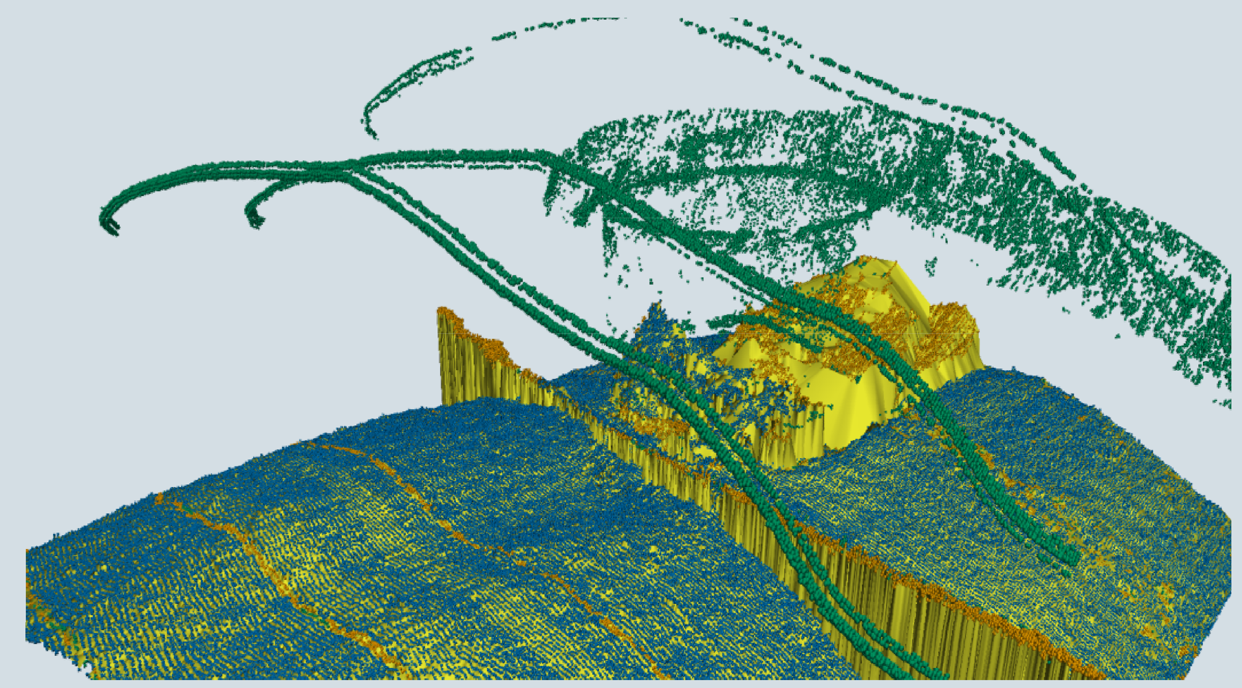
In the picture to the right, the orange section of the pipe-line does not lie on the seabed, leaving a gap, such that points on the seabed are also reported.

Due to the 2D triangulation, the points on the pipe appear in small components and will be removed early. This problem can be solved by augmenting the triangulated graph.

Pipeline spanning a valley.

Green points represent noise, removed by both algorithms. Orange points are removed by the original algorithm, even though they should have been kept, which is what the extended algorithm indeed does.





References

[1] Arge, Larsen, Mølhave, van Walderveen. *Cleaning massive sonar point clouds.* ACM SIGSPATIAL GIS 2010.

MADALGO – Center for Massive Data Algorithmics, a Center of the Danish National Research Foundation