DECONFLICTION AND SURFACE GENERATION FROM BATHYMETRY DATA USING LR B-SPLINES

IQMULUS WORKSHOP
BERGEN, SEPTEMBER 21, 2016

Vibeke Skytt, SINTEF

Jennifer Herbert, HR Wallingford

The research leading to these results has received funding from the European Union Seventh Framework Programme (*FP7/2007-2013*) under *grant agreement* n° 318787.

GEOSPATIAL DATA

- Types
  - Topography – land
  - Bathymetry – subsea
  - Subsurface phenomena – ground water, geological layers

- Acquisition methods
  - LIDAR
  - Photogrammetry
  - Sonar

Modern acquisition methods typically provide huge amounts of data

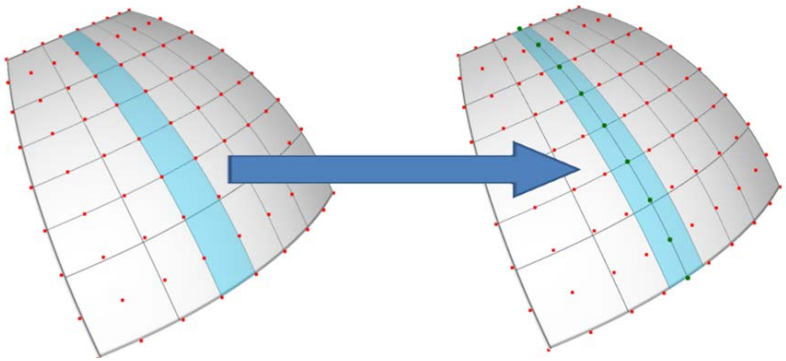
- Surface representation for geospatial phenomena
  - Raster (GeoTiff): Digital elevation model, uniform height map
    - Inverse weighted interpolation, natural neighbour interpolation, radial basis functions, …
  - Triangulations: Possible to interpolate, able to represent detail, may get huge
  - Splines: Approximation, smooth, not common

"BIG DATA" IS MORE THAN FILE SIZE

Indicator	Comment	In this context
Volume	Modern data acquisition technology creates high volumes of data.	Divide and conquer: Tiling and stitching
Variation	Topography and bathymetry data gained by different equipment and with different quality properties give heterogeneous point clouds.	Select the most appropriate data for surface generation: deconfliction
Velocity	Data from the same area changes in time	Select the most appropriate data for surface generation: deconfliction
Data reduction	Make acquired data more manageable by applying a different and more lean storage format	Spline approximation

SPLINE SURFACES

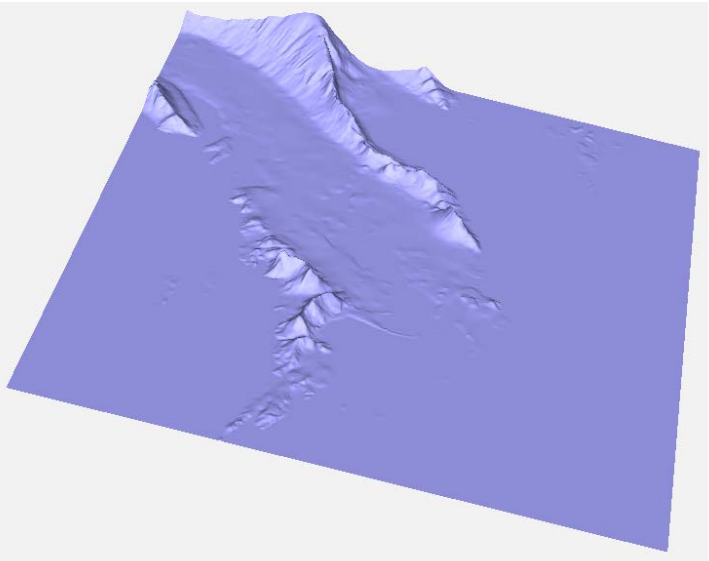
- Over dominantly smooth domains, splines provide compact representations
- Spline surfaces are piecewise polynomials typically defined on a regular domain
- The polynomial patches join with specified smoothness
- Are numerically stable: non-negative basis functions, partition of unity, linear independence of basis functions, …
- Exist methods for
  - Surface generation, e.g. approximation of scattered data points
  - Refinement – increase the number of polynomial patches
  - Interrogation, e.g. intersections with planes to produce iso curves



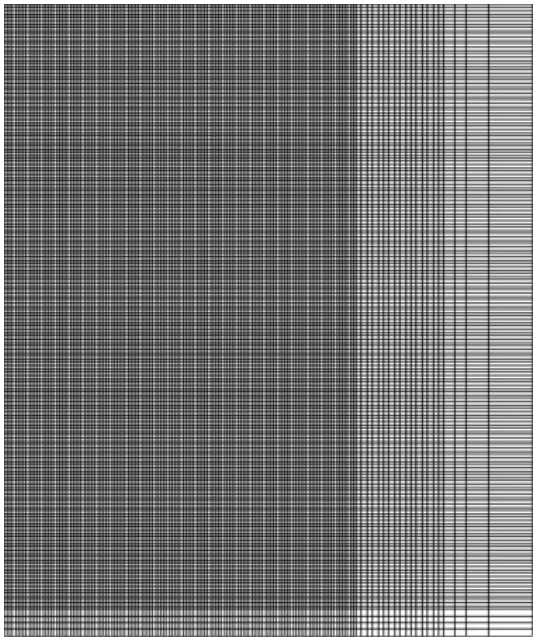
Local refinement is not possible for tensor-product splines

LOCALLY REFINABLE SPLINE SURFACES

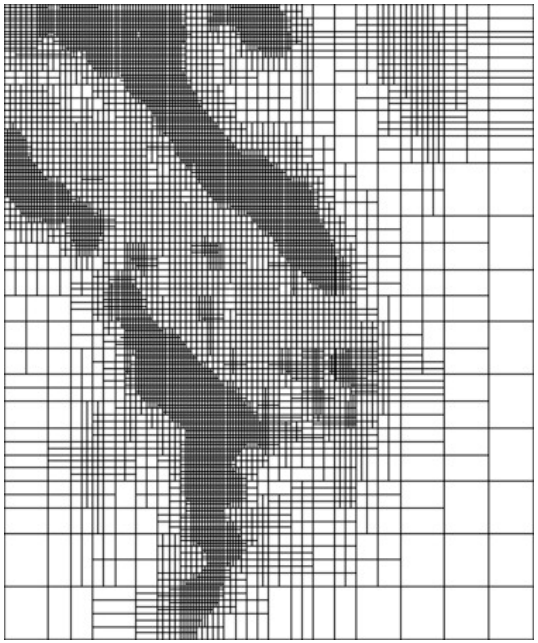
- Support local refinement
- Can model much larger regions with better accuracy
  - Hierarchical splines
  - T-splines
  - **LR B-splines**



Surface



Polynomial patches, tensor-product spline surface



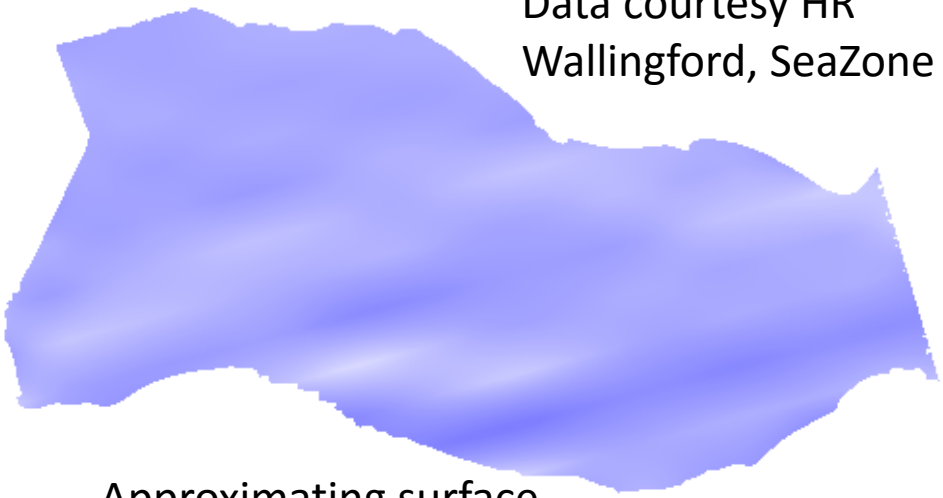
Polynomial patches, LR B-spline surface

ITERATIVE ALGORITHM FOR LR B-SPLINE SURFACE APPROXIMATION

- **Input:** point cloud, threshold, maximum number of iterations
- **Algorithm:**
  - Make a lean approximation on a regular grid
  - Compute distances between points and surface approximation
  - While (max err > threshold AND number of iterations < max number)
    - Refine the surface by inserting knots in regions where the error is above threshold
    - Compute approximation given the current degrees of freedom
- **Output:** LR B-spline surface, accuracy information

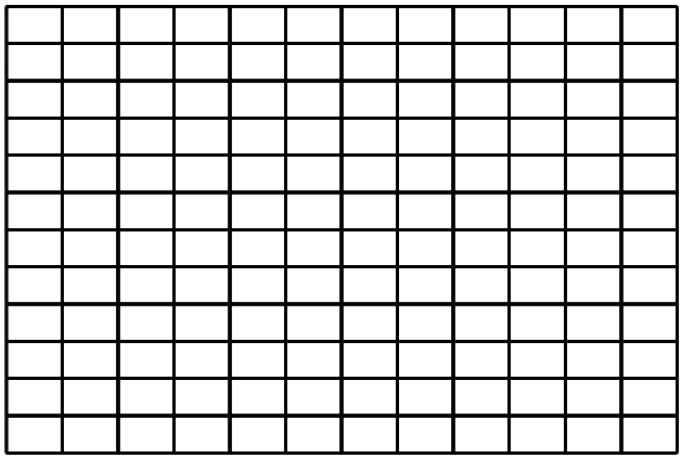
APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. INITIAL SURFACE.

Data courtesy HR Wallingford, SeaZone

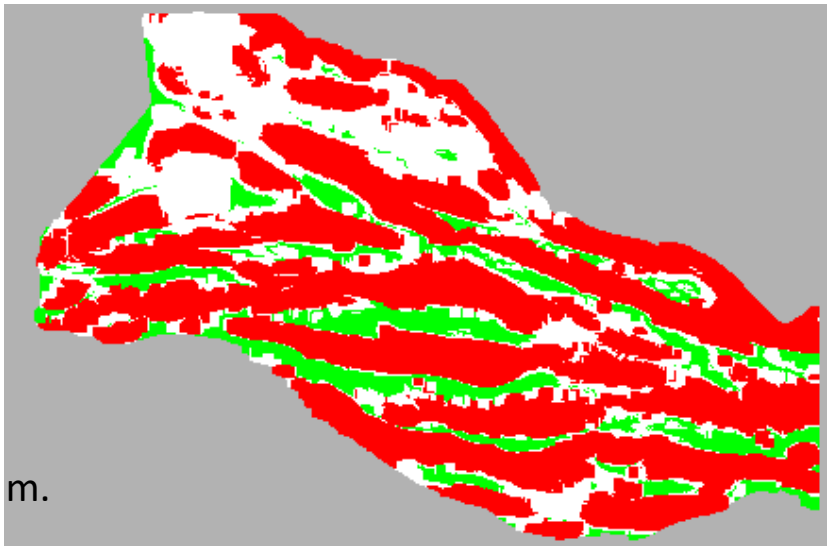


Approximating surface

Number of points	14.6 mill
No. of coefs.	196
Surface file size	26 KB
Max. dist	12.8 m.
Average dist	1.42 m.
No. of points, dist > 0.5 m	9.9 mill



Polynomial patches in the parameter domain of the surface



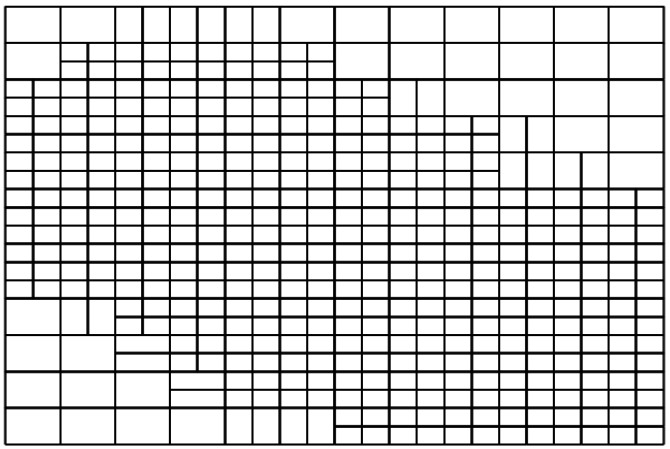
Elevation interval: ~50 m.

Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

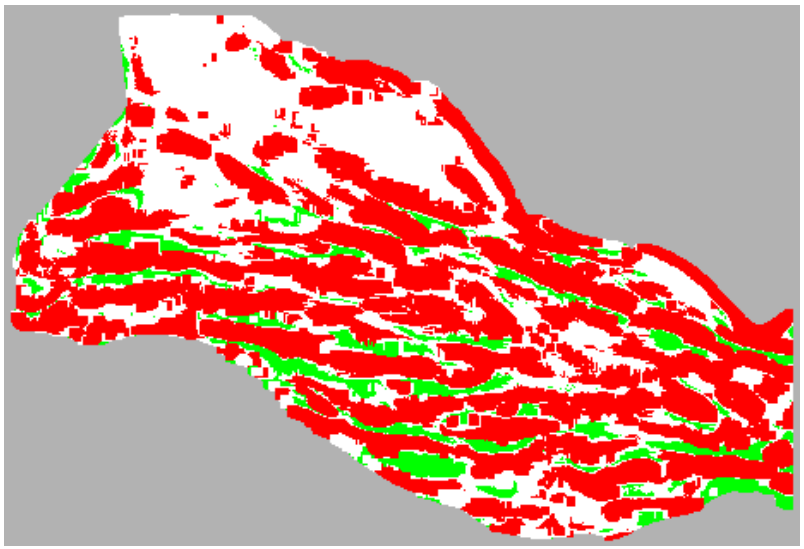
APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER ONE ITERATION.

Approximating surface

Number of points	14.6 mill
No. of coefs.	507
Surface file size	46 KB
Max. dist	10.5 m.
Average dist	0.83 m.
No. of points, dist > 0.5 m	7.3 mill



Polynomial patches in the parameter domain of the surface

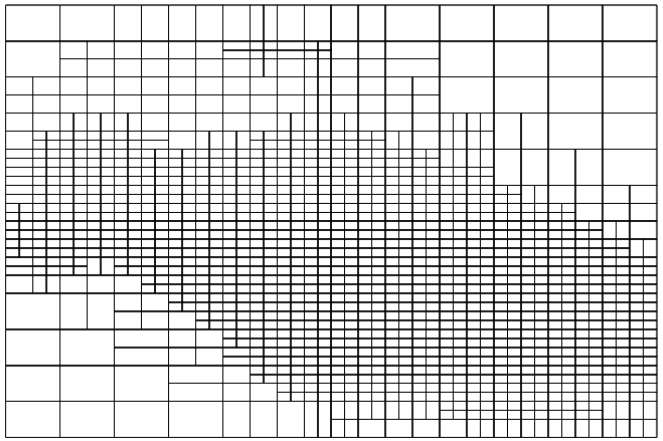


Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

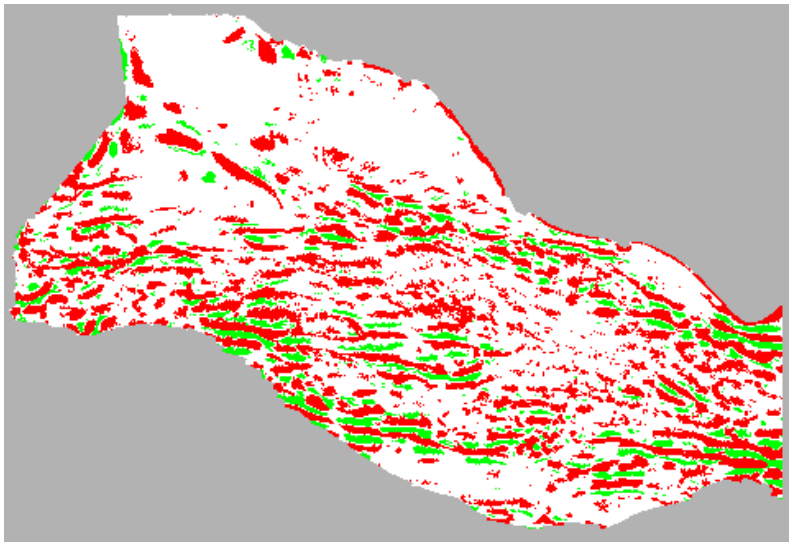
APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER TWO ITERATIONS.

Approximating surface

Number of points	14.6 mill
No. of coefs.	1336
Surface file size	99 KB
Max. dist	8.13 m.
Average dist	0.41 m.
No. of points, dist > 0.5 m	3.9 mill

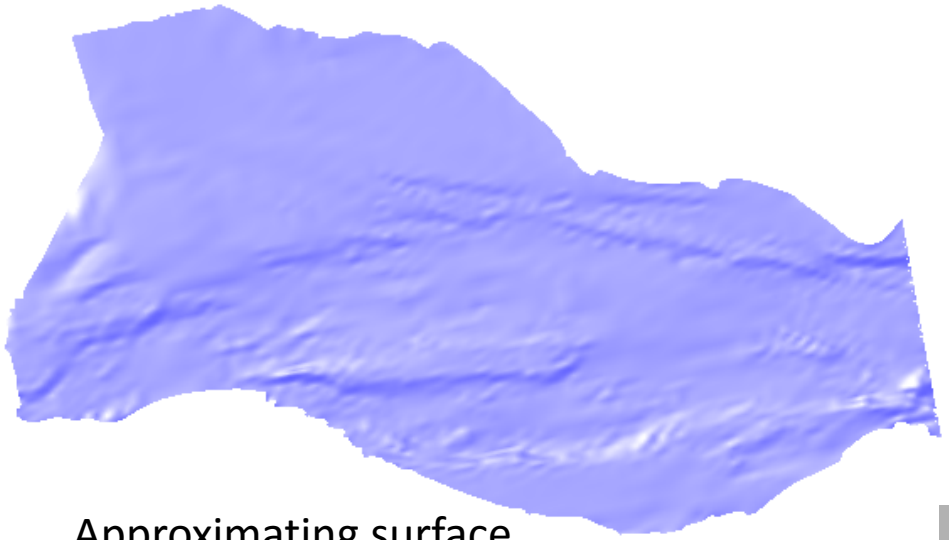


Polynomial patches in the parameter domain of the surface



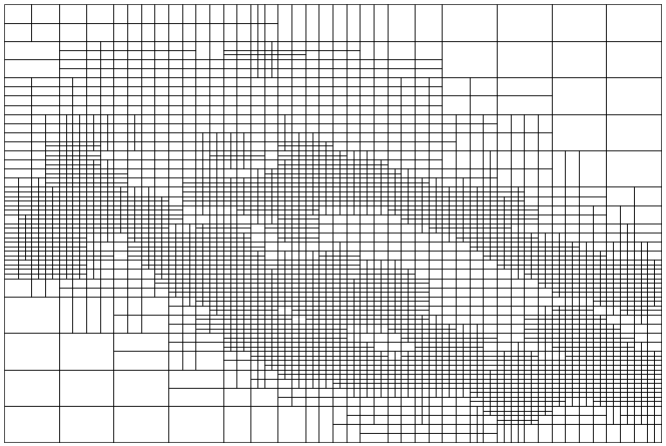
Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER 3 ITERATIONS

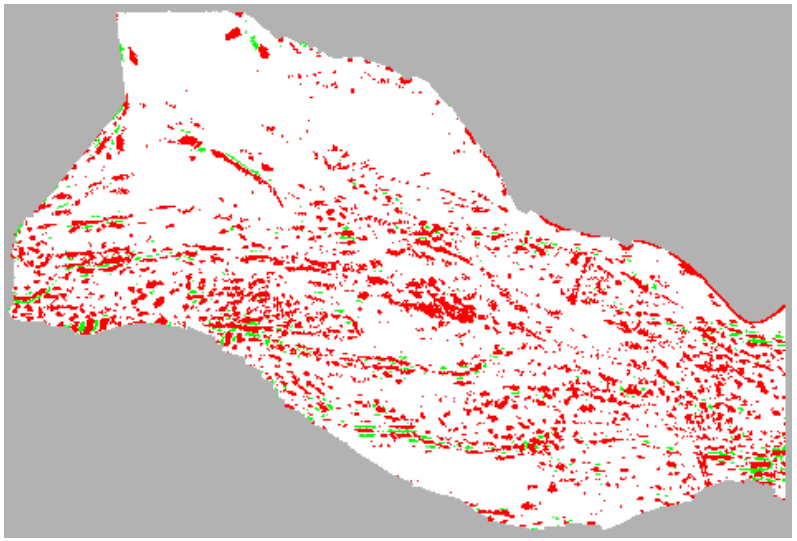


Approximating surface

Number of points	14.6 mill
No. of coefs.	3563
Surface file size	241 KB
Max. dist	6.1 m.
Average dist	0.22 m.
No. of points, dist > 0.5 m	1.4 mill

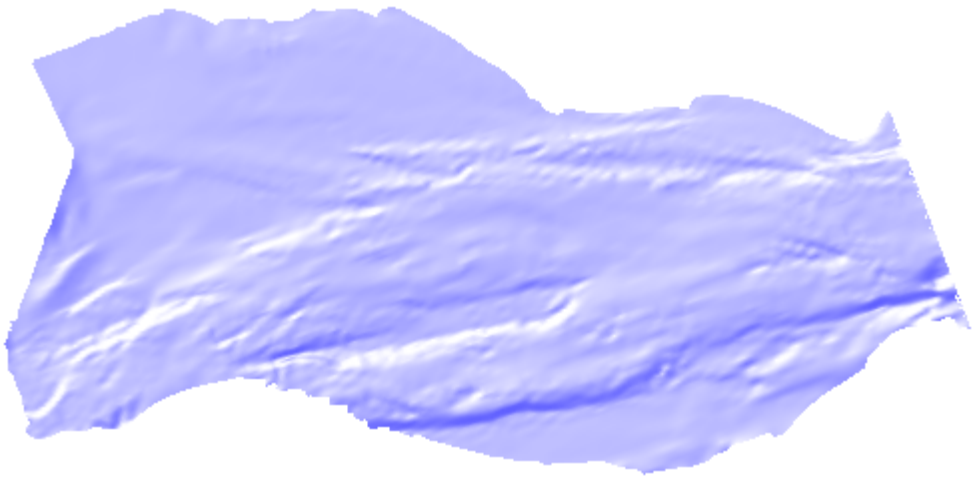


Polynomial patches in the parameter domain of the surface



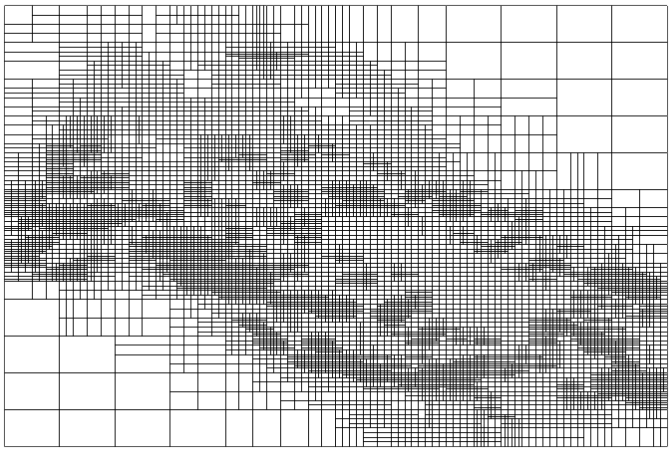
Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER 4 ITERATIONS

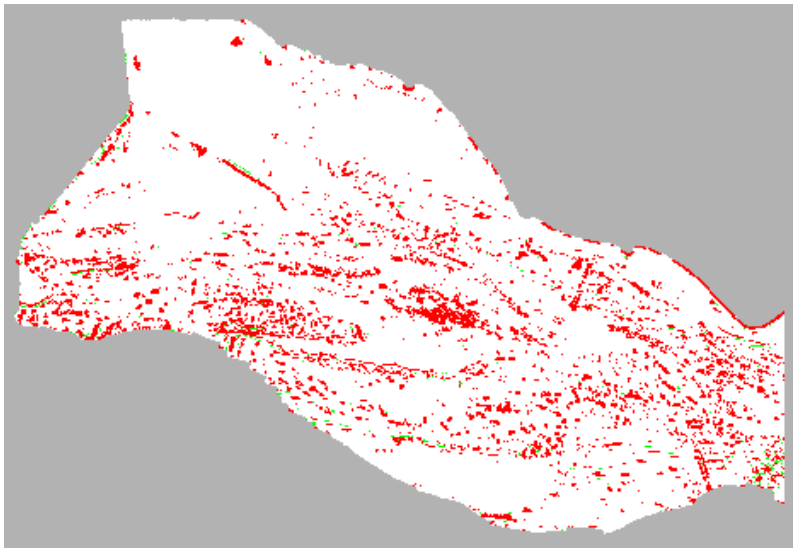


Approximating surface

Number of points	14.6 mill
No. of coefs.	9273
Surface file size	630 KB
Max. dist	6.0 m.
Average dist	0.17 m.
No. of points, dist > 0.5 m	0.68 mill

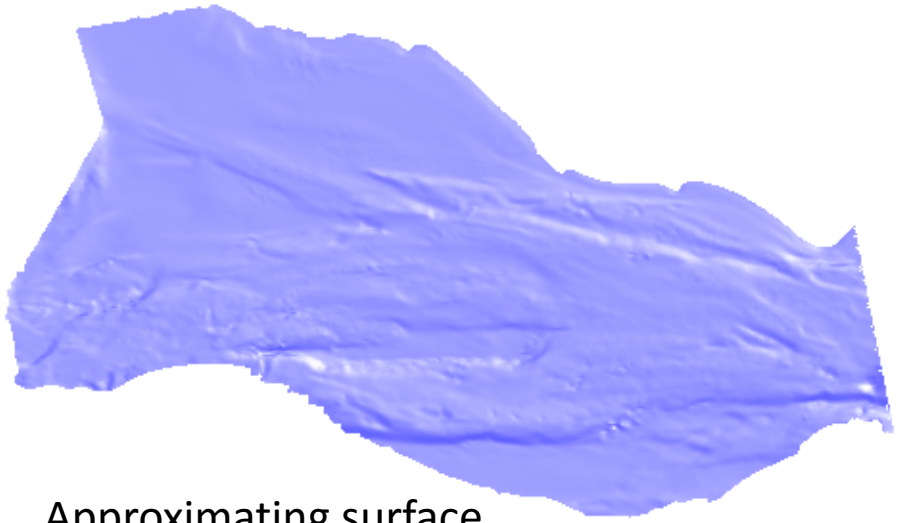


Polynomial patches in the parameter domain of the surface



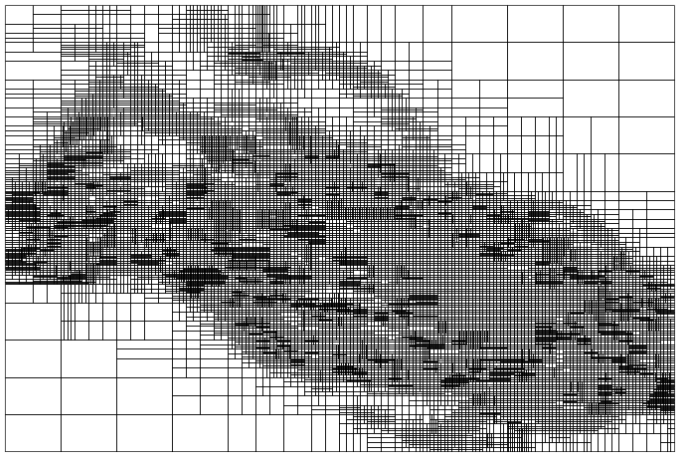
Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER 5 ITERATIONS.

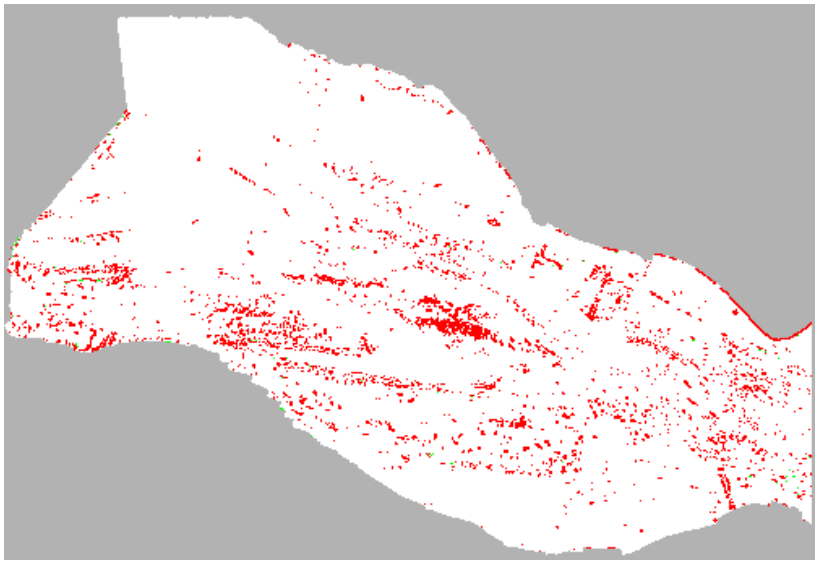


Approximating surface

Number of points	14.6 mill
No. of coefs.	23002
Surface file size	1.6 MB
Max. dist	5.3 m.
Average dist	0.12 m.
No. of points, dist > 0.5 m	244 850

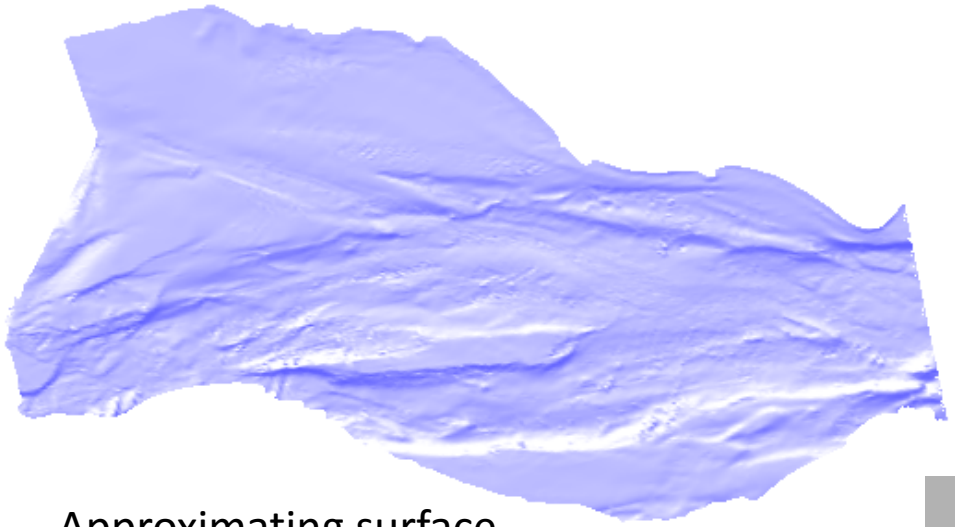


Polynomial patches in the parameter domain of the surface



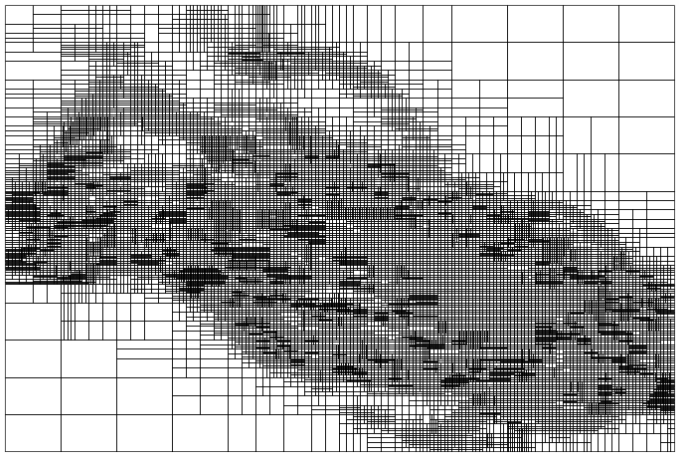
Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER 6 ITERATIONS

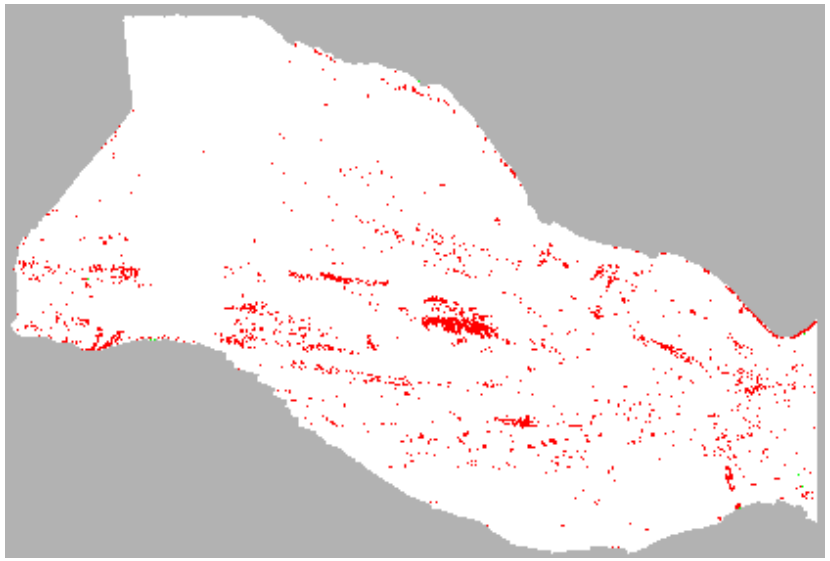


Approximating surface

Number of points	14.6 mill
No. of coefs.	52595
Surface file size	3.7 MB
Max. dist	5.4 m.
Average dist	0.09 m.
No. of points, dist > 0.5 m	75 832

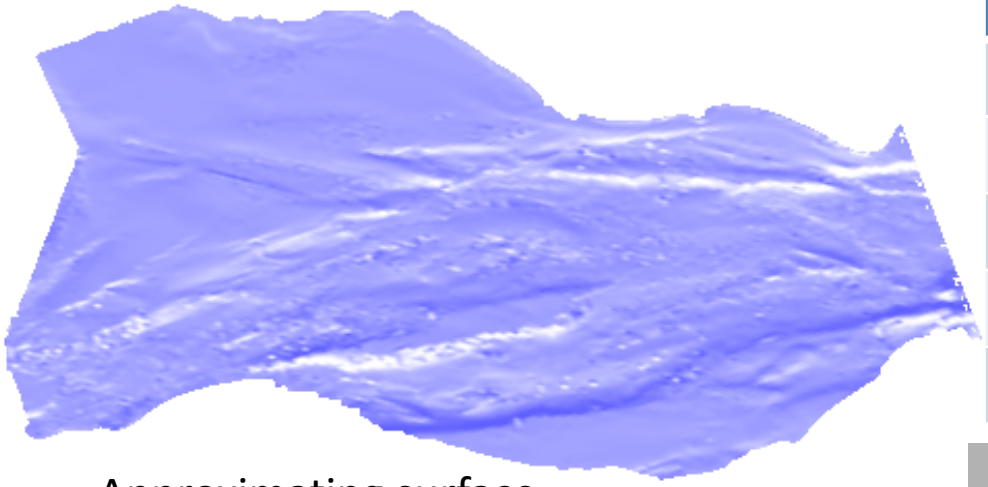


Polynomial patches in the parameter domain of the surface



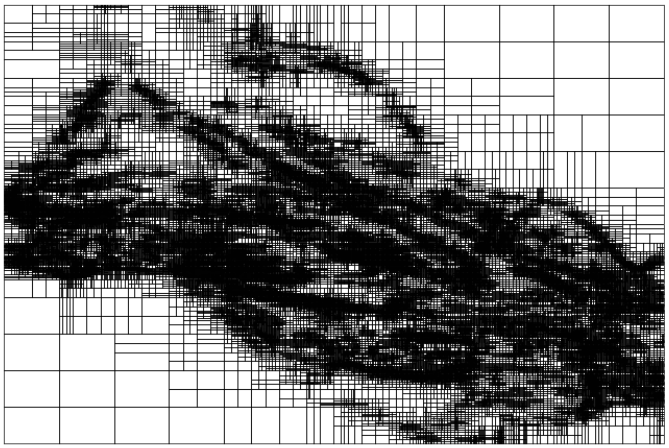
Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

APPROXIMATION OF 280 MB POINT CLOUD. THRESHOLD 0.5 M. AFTER 7 ITERATIONS

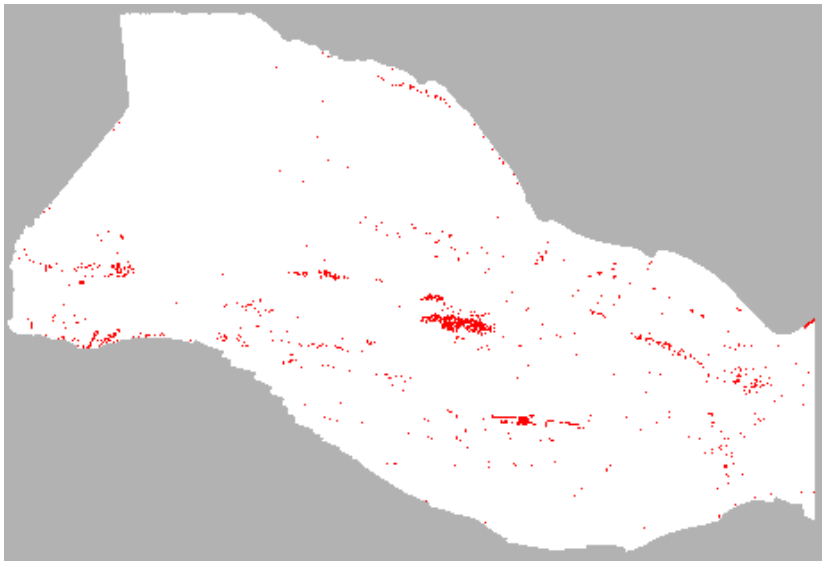


Approximating surface

Number of points	14.6 mill
No. of coefs.	99407
Surface file size	7.0 MB
Max. dist	5.3 m.
Average dist	0.08 m.
No. of points, dist > 0.5 m	20 148 (0.1%)

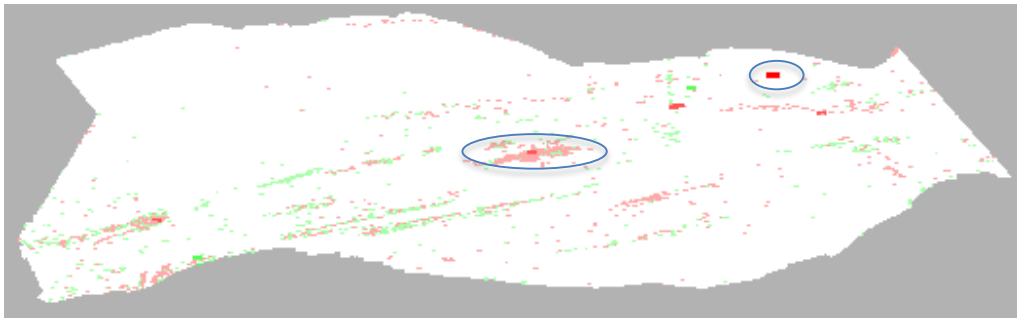


Polynomial patches in the parameter domain of the surface



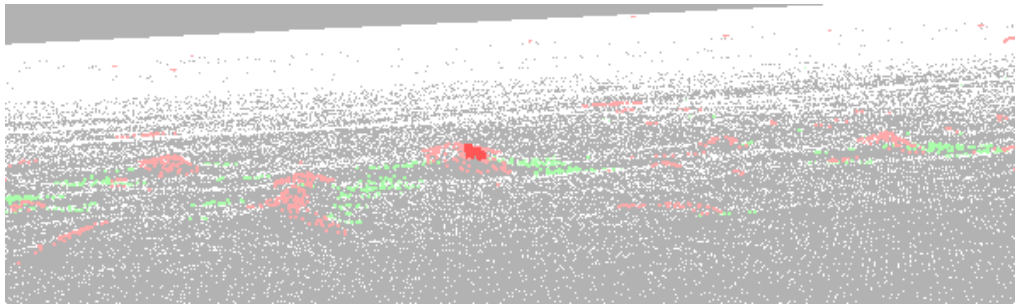
Distance field, white points less than 0.5 m distance, red points above, green (partly hidden) below surface

SUMMARY, ANALYSING THE DISTANCE FIELD



7 iterations. Increasing distance is visualized with larger points size and clearer colour

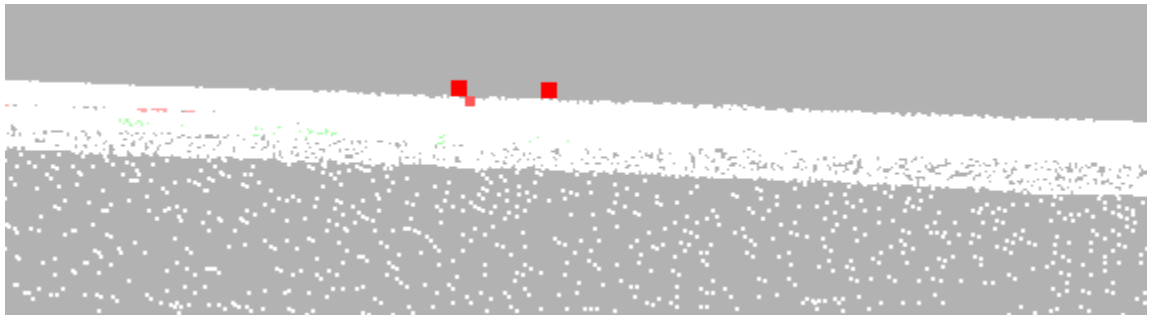
# it	Max dist	Av. dist	File size	# out
1	10.5 m.	0.83 m.	46 KB	7.3 mill
2	8.1 m.	0.41 m.	99 KB	3.9 mill
3	6.1 m.	0.22 m.	241 KB	1.4 mill
4	6.0 m.	0.17 m.	630 KB	0.68 mill
5	5.3 m.	0.12 m.	1.6 MB	244 850
6	5.4 m.	0.09 m.	3.7 MB	75 832
7	5.3 m.	0.08 m.	7.0 MB	20 148



Some details are not absolutely captured

Distance > 4 m.: 2 points
Distance > 2m.: 22 points

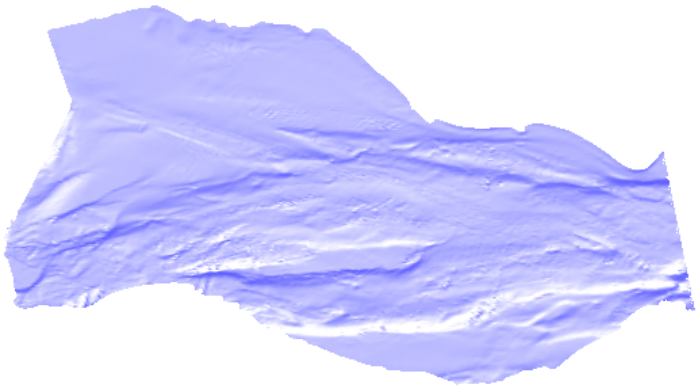
Elevation interval: ~50 m.



Outliers

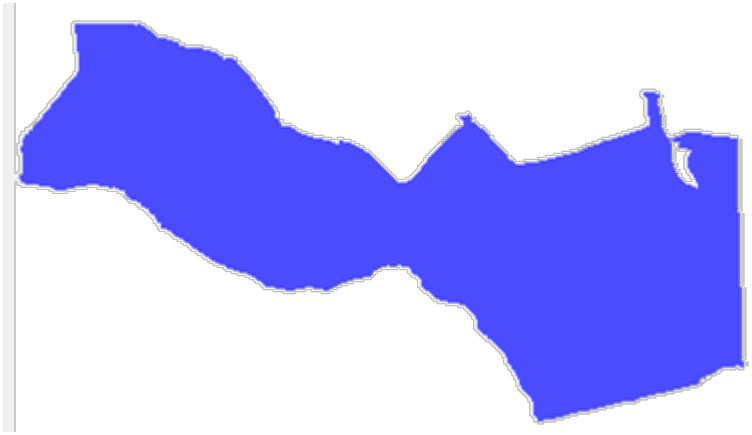
SURFACE SIZE VERSUS POINT CLOUD SIZE

- Original data set contains approx. 58 million points
- We perform successive thinning of the point cloud and approximate with fixed parameters:
  - 0.5 m threshold, 6 iteration levels
- Results are very stable showing that the resulting LR B-spline grid is more dependent on the features of the terrain than the number of points in a scan.



No. points	No. coefs.	Max. error	Average error	Average outside	Prop. OOT points
58 578 420	53454	5.55	0.092	0.66	0.56%
29 289 210	52 709	5.39	0.092	0.66	0.55 %
14 644 406	52 595	5.39	0.093	0.65	0.52 %
7 322 302	52 611	5.33	0.093	0.65	0.47 %
3 661 151	53628	5.25	0.093	0.65	0.41%
1 830 575	51 124	3.24	0.094	0.65	0.40 %

TILING

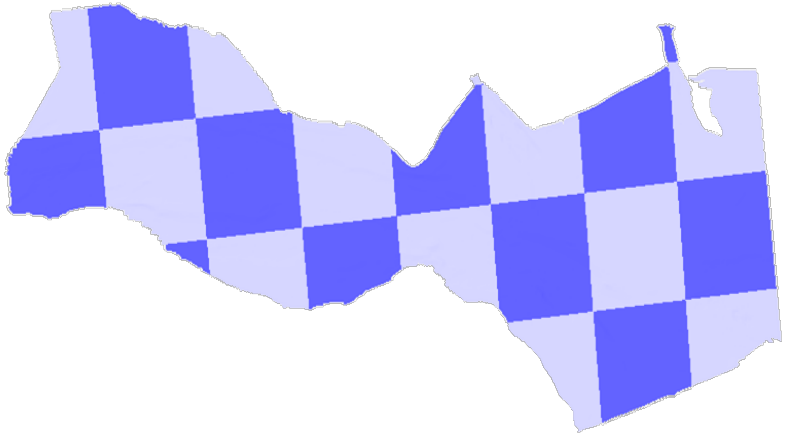


One data survey: 131 million points

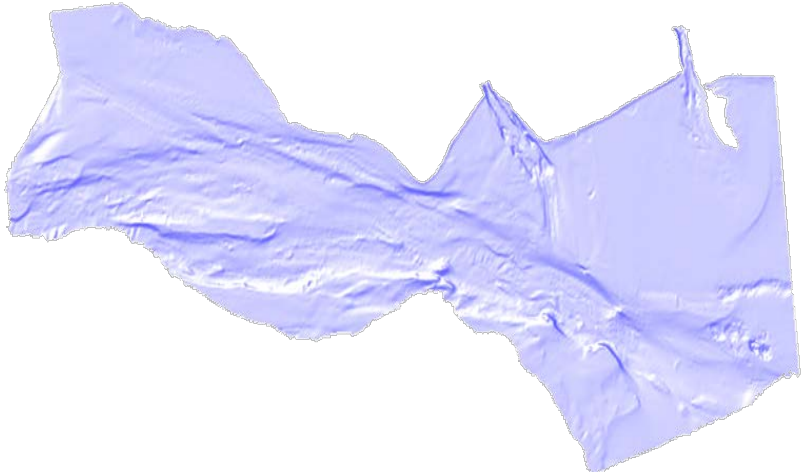
Many data sets are too large to approximate by a single surface

- Many operations for each single point (execution time)
- Memory limitations (risk of crash)
- Very large surfaces are cumbersome to handle (execution time of subsequent operations)

Tiling enables parallelization on computation nodes



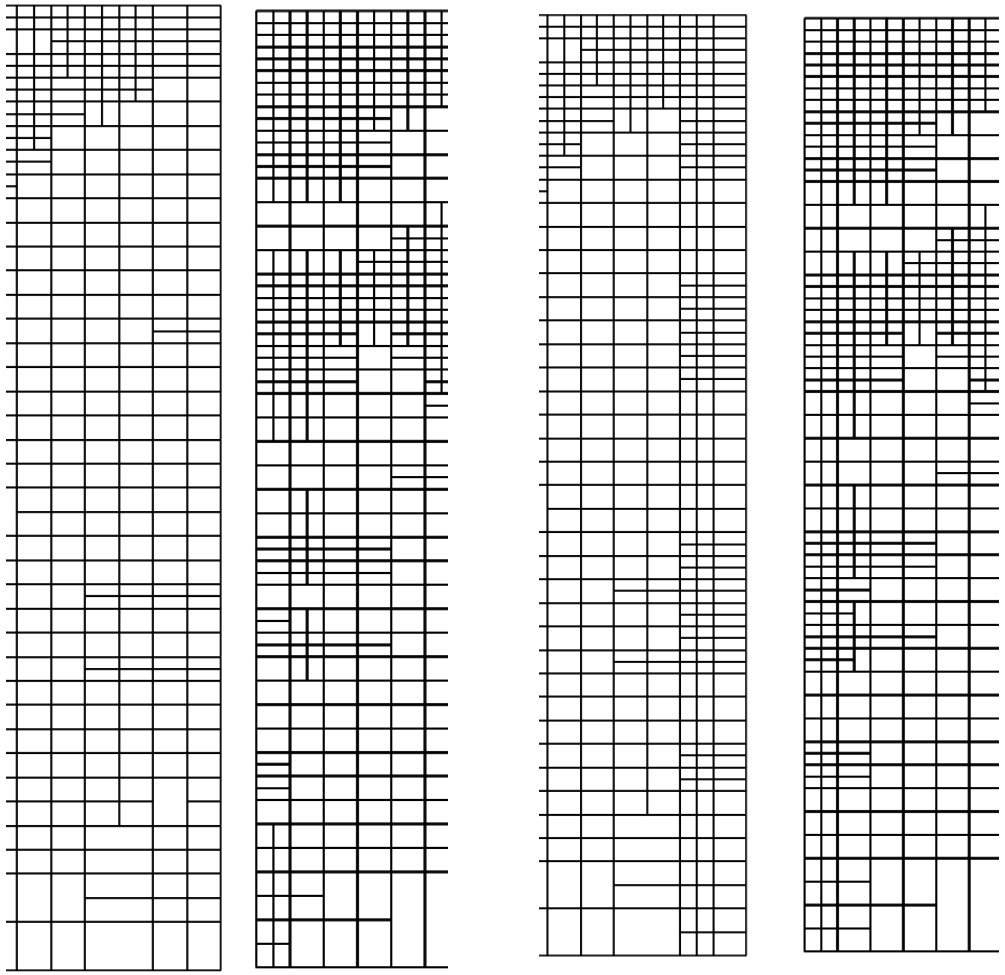
Surface set approximating data set after being split into regular tiles



C^1 seamless super surface

STITCHING

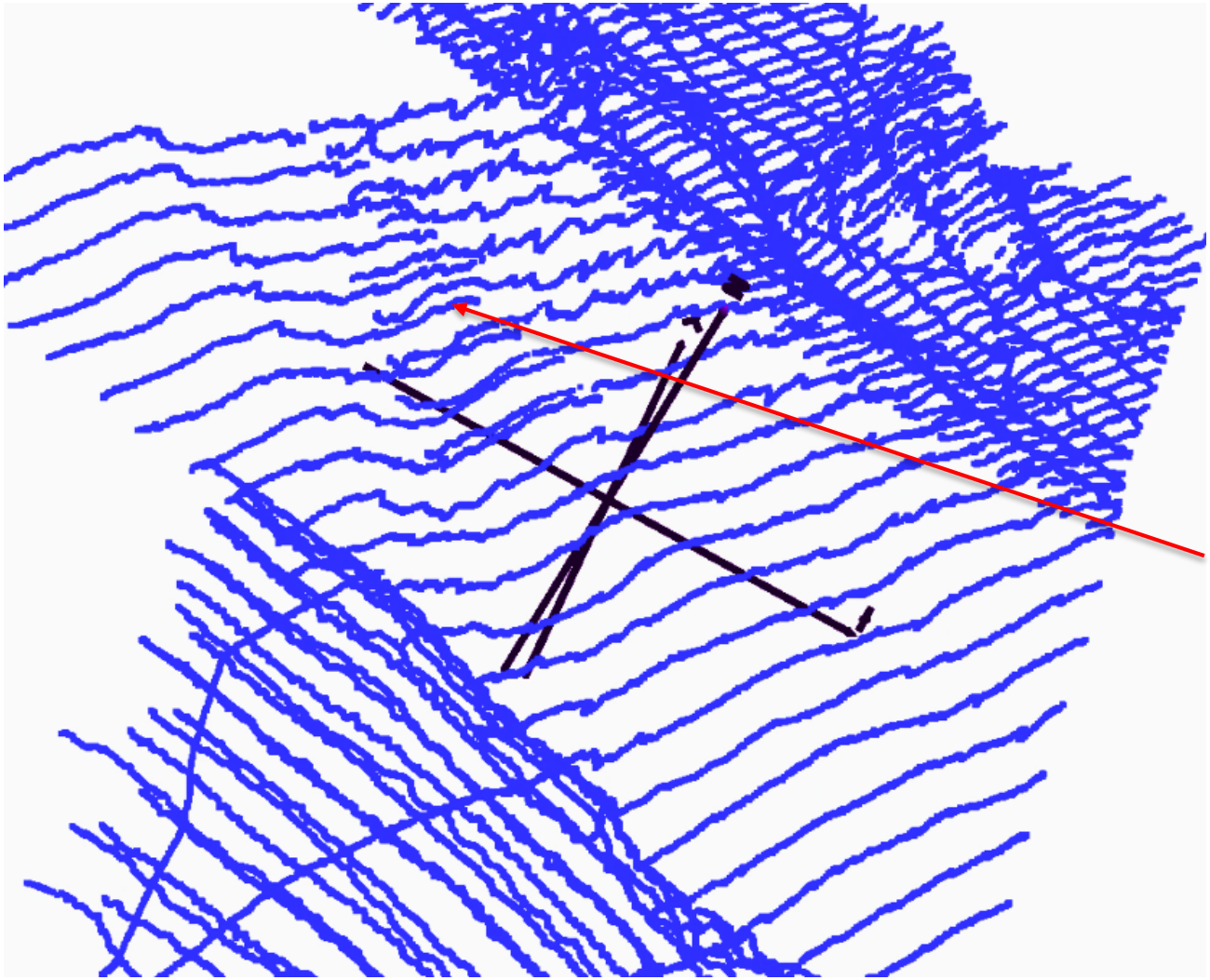
- Surfaces are created with an overlap and cut back to tile boundaries. This gives an almost seamless combined surface
- Stitching to get a completely seamless surface
- C^0 continuity is easy to achieve if the two boundary curves have the same polynomial structure
- C^1 continuity obtained through demanding similar polynomial patches at the boundary
- Local refinement makes this possible without globally increasing the data size of the surfaces



before

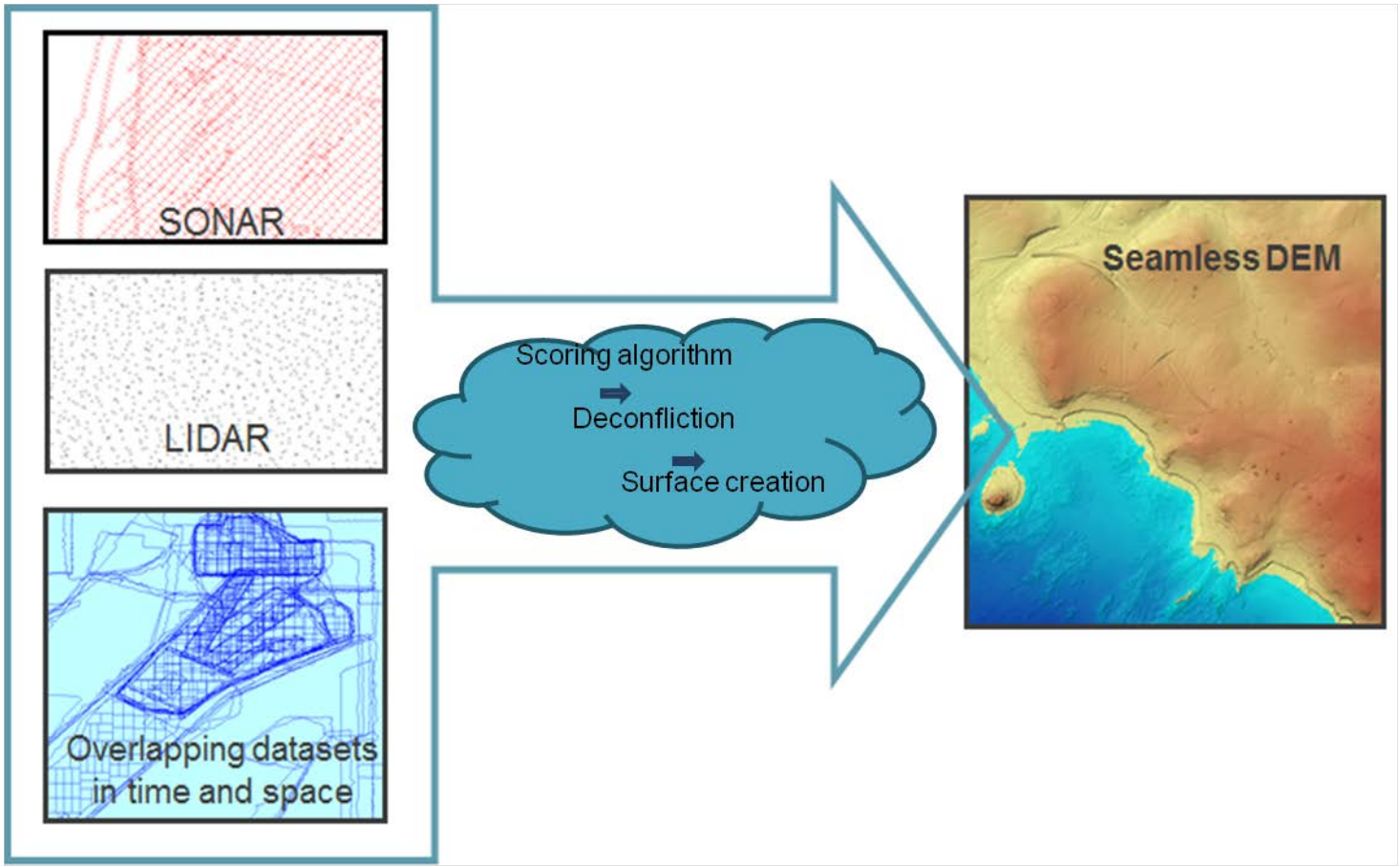
after

INCONSISTENCIES IN THE DATA SET



Not easily modelled by a smooth surface. But is this behaviour correct?

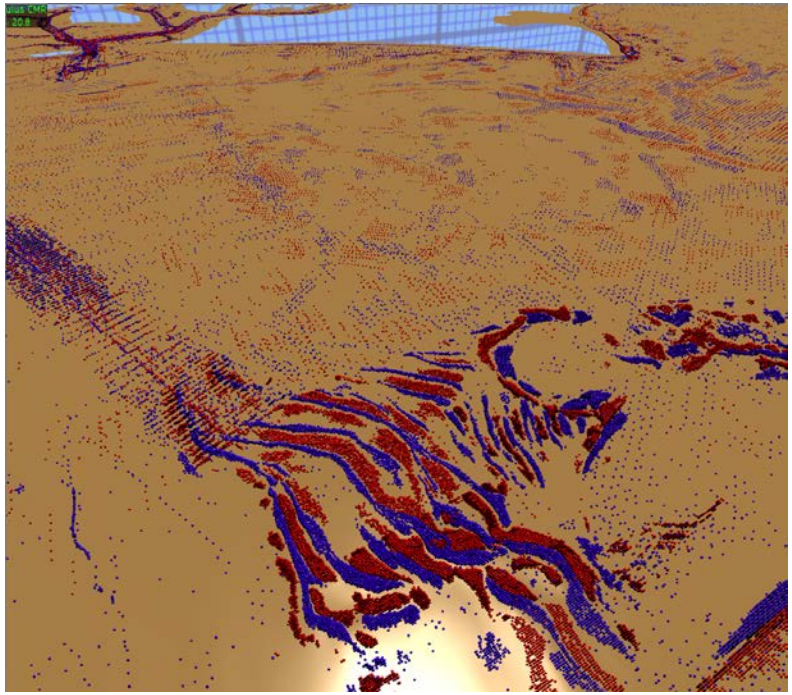
SURFACE GENERATION WORKFLOW WITH HETEREOGENOUS POINT CLOUDS



PREPARE FOR DECONFLICTION

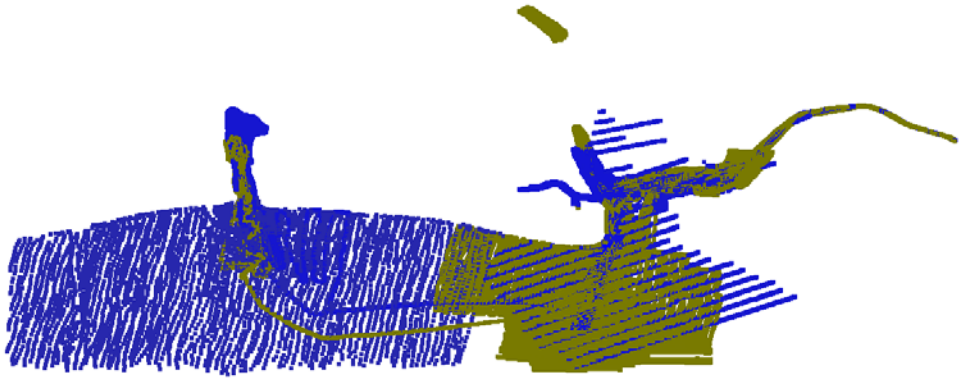
- Collections of bathymetric surveys are an example of a potentially ‘big data’ source structured as a point clouds
- Individual surveys vary both spatially and temporally and can overlap with many other similar surveys
- Depth soundings differ greatly between surveys
- A strategy needs to be employed to determine how to create an optimal bathymetric surface using all of the relevant, available data, i.e. the *best* surface
- Prioritise the surveys based on included metadata and generated statistics
  - Most recent data
  - Point density
  - Survey technique
  - Data type
- Select the *best* data for surface generation

High data volumes, heterogeneous point clouds

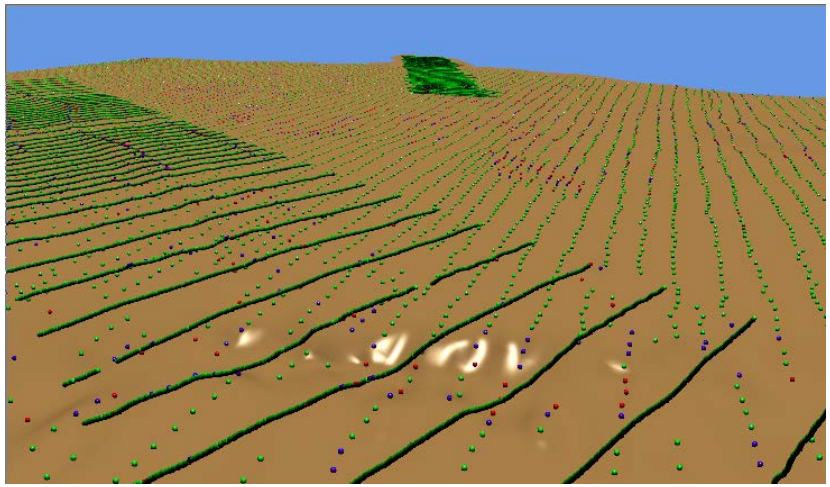


DECONFLICTION

Deconfliction is about removing points that don't fit in



- Given
  - A number of overlapping data surveys
  - A priority score for each survey
  - Can not expect complete consistency between the various data surveys
  - The combined point cloud will typically have a very hetereogenous pattern

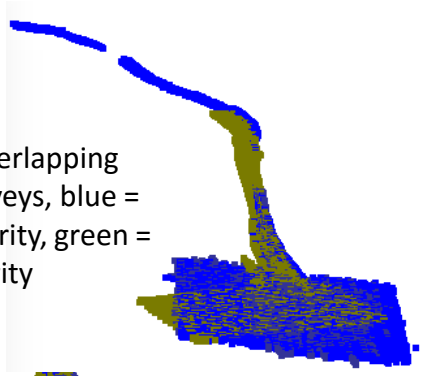


- Want
  - A consistent point cloud for surface generation
  - Avoid large areas without points
  - Avoid large jumps in height between adjacent cleaned data surveys

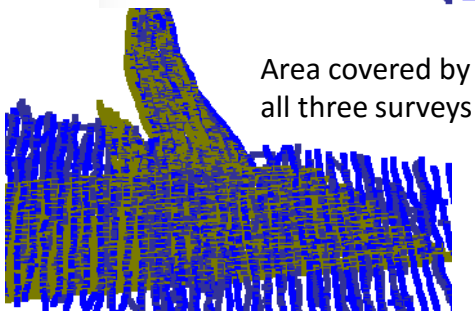
DECONFLICTION ALGORITHM

- Is performed tile by tile. Can be run in parallel
- Create reference surface: Approximate combined point cloud by low resolution LR B-spline surface
- For each data survey
  - Distribute data points with respect to the polynomial patches of the LR B-spline surface
  - Compute the distances between the data points and the surface
  - Compute statistics: distance interval, average distance, standard deviation, number of points above and below surface, …
- For each polynomial patch
  - Use the distance statistics and the priority score of the data surveys to decide whether to keep or remove a particular group of points
- If in doubt: Use decision on adjacent polynomial patches to make a decision
- Post processing: Update reference surface with respect to cleaned (deconflicted) point set for a more accurate surface representation

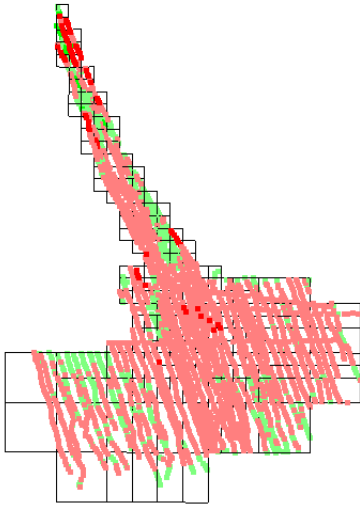
DECONFLICTION ALGORITHM EXAMPLE



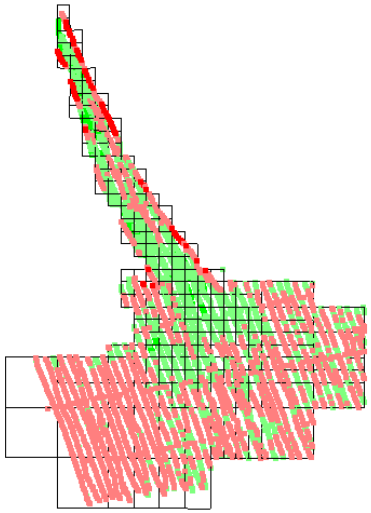
Three overlapping data surveys, blue = high priority, green = low priority



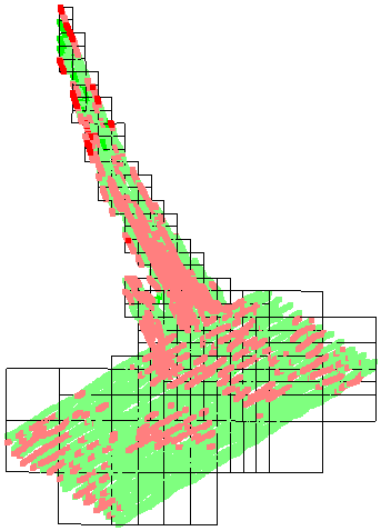
Area covered by all three surveys



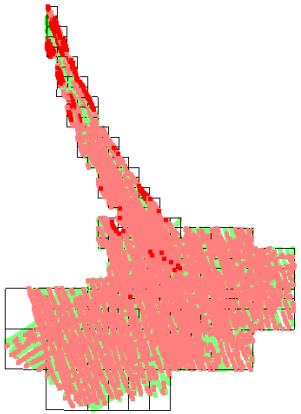
Survey set 1, medium priority



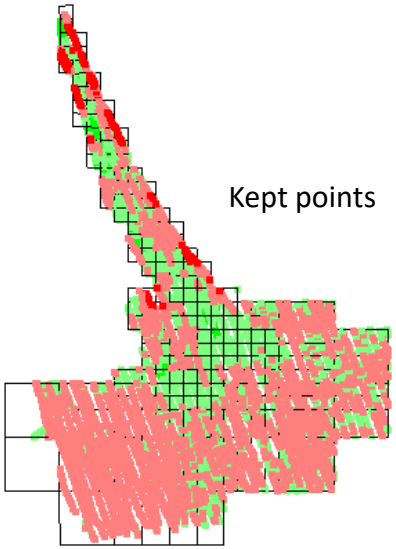
Survey set 2, high priority



Survey set 3, low priority



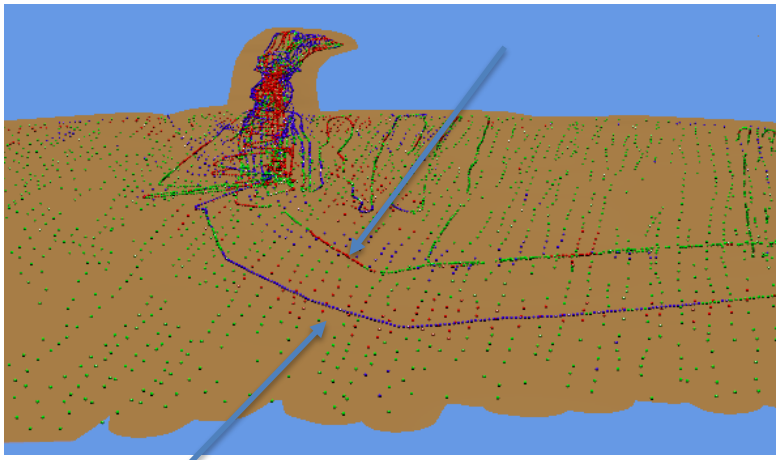
All data points classified by their distance to the reference surface, red points above and green points below. Clearer colour means larger distance



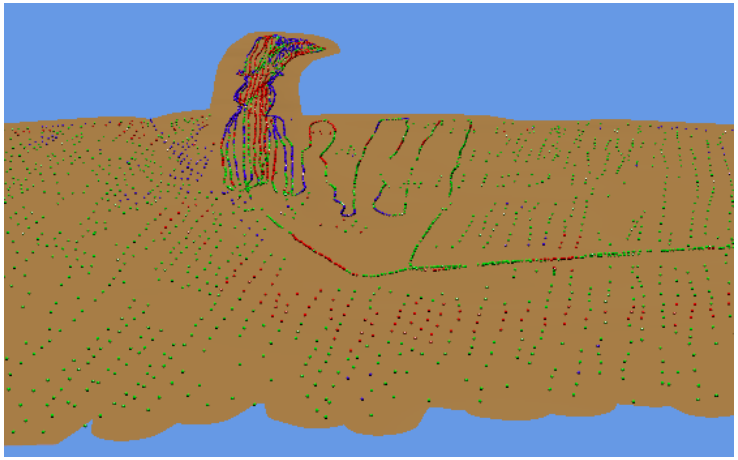
Kept points

Removed points

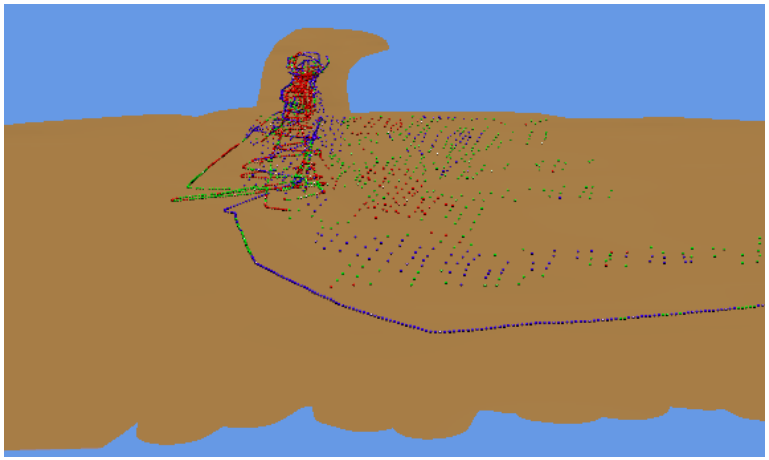
DECONFLICTION AND POST DECONFLICTION SURFACE GENERATION



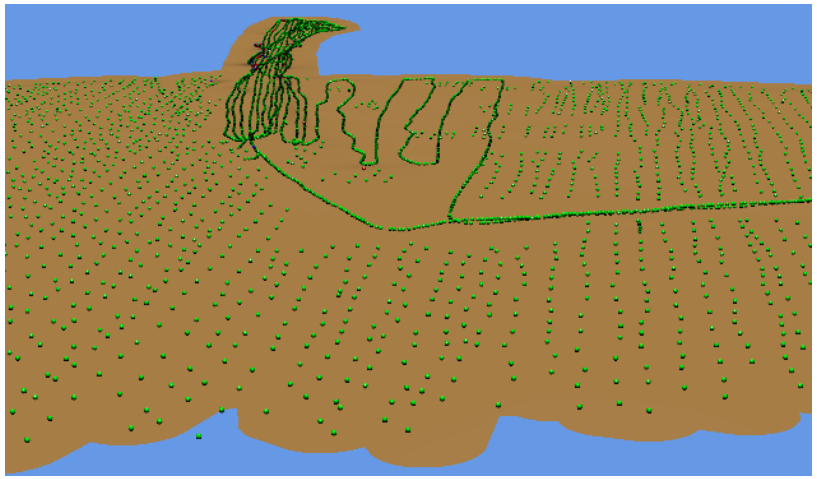
A detail with several overlapping data sets. The green points lie within the threshold from the reference surface, the blue points are above and the red ones below.



Reference surface and remaining points after deconfliction



Reference surface and points removed by the deconfliction



Final surface and the points used to create it. Most of the points lie within the threshold

SUMMARY

- Geospatial data processing
- Current data acquisition methods provides huge amounts of data
- Data reduction using splines
- Locally refined splines allow modelling of local details
- Resolve conflicts between different data sets by deconfliction
- To be continued

IQmulus - A High-volume Fusion and Analysis Platform for Geospatial Point Clouds, Coverages and Volumetric Data Sets
www.iqmulus.eu