

3D modeling and integration of heterogeneous geo-data

M. Miola^{†1} and D. Cabiddu¹ and S. Pittaluga¹ and M. Mortara¹ and M. Vetuschi Zuccolini² and G. Imitazione³

¹CNR-IMATI, Genova, Italy

²DISTAV, University of Genova, Italy

³RINA Consulting, Genova, Italy

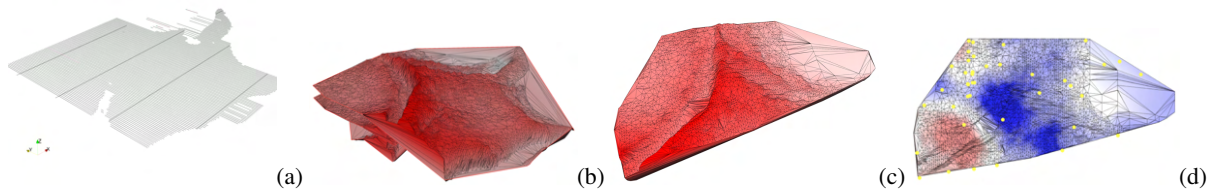


Figure 1: Representation of the developed procedure to generate a 3D model of a submarine deposit, starting from (a) available data analysis, to (b) the definition of triangular meshes, to (c) the generation of tetrahedral volume, and to (d) the integration of geotechnical parameters.

Abstract

This paper tackles the volumetric representation of geophysical and geotechnical data, gathered during exploration surveys of the subsoil. The creation of a 3D model as support to geological interpretation has to take into account the specificity of the diverse input data, that are heterogeneous. Some data are massive, but cover the domain unevenly, e.g., structured along dense differently spaced lines, while others are very sparse, e.g., borehole locations with soil sampling and CPTU (Piezocone Penetration Test) locations. In this work, we focus on the exploration and analysis of underwater deposits.

After a discussion about the data typically acquired in an offshore campaign, we present an automatic process to generate the subsurfaces and volume defining an underground deposit, starting from the identification of relevant morphological features in seismic data. In particular, data simplification and refinement based on geostatistics have been applied to generate regular 2D meshes from strongly anisotropic data, in order to improve the quality of the final 3D tetrahedral mesh. Furthermore, we also use geostatistics to predict geotechnical parameters from local surveys and estimate their distribution on the whole domain: in this way the 3D model will include relevant geological features of the deposit and allow extrapolating different geotechnical information with associated uncertainty. The volume characterization and its 3D inspection will improve the structural and stratigraphic interpretation of deposits, to support geological analysis and planning of future engineering activities.

Keywords: Representation of subsoil, 3D modeling, Geostatistical analysis

1. Introduction

Thanks to the technological developments of the last decades, there are new approaches to describe and comprehend environmental phenomena. Through faithful digital representations reconstructed from heterogeneous real data, simulations on the digital model can potentially be considered reliable experiments on the real phenomenon. This is also the concept behind the so-called digital twin,

applied to various complex systems, from factories to entire cities [FFDB20].

In the sector of geological investigations, knowledge of the properties of the subsoil is an essential requirement for the planning and construction of infrastructure building, for the assessment of seismic and hydrogeological risks and for the identification and investigation of resources deposits.

In order to characterise a portion of submarine deposit, offshore surveys are carried out to understand the lithostratigraphy of the

[†] Corresponding Author

site of interest. These campaigns typically combine geotechnical and geophysical investigations, carrying out soil sampling, in-situ testing and geophysical data acquisitions (i.e. seismic survey, *Multi Beam Echo Sounder* MBES, *Side Scan Sonar* SSS, Magnetometer, ...). The results of these campaigns consist of different types of data in terms of investigated properties, format and areal extent. The characterisation of the site, i.e. the interpretation of the obtained information, to which other pre-existing data are often added, requires a high degree of processing and transversal skills.

At present, available data is managed and processed by a variety of software programs with specific purposes; achieving a global and comprehensive view of the data set, which is fundamental for an appropriate understanding of the site, is hard-working, non-automatic and still based on two-dimensional paradigms. In this scenario, 3D modelling and integrated visualisation of soil and subsoil morphology represents today a concrete possibility of innovation, allowing to interact and analyse data in a much more effective and direct way. In addition to rendering, a volumetric model of the lithostratigraphy can also support specific simulations for the characterization and interpretation of the environment under study.

Our goal is then to develop a 3D volumetric modelling system for the representation analysis and visualisation of heterogeneous sub-seabed data. Our approach integrates computer graphics methodologies related to 3D geometric modelling with geophysical and geostatistical techniques.

The approach comprehends two main stages: firstly, the input geophysical data must be properly processed in order to correctly represent the reference surfaces and the consistent subsoil volumetric layers. Secondly, the geotechnical data available at sparse, but well-defined, locations are estimated throughout the whole deposit volume using geostatistical techniques, producing a 3D map of expected values, along with a quantitative assessment of the uncertainty related to each estimate.

The modeling phase offered the biggest challenges, related to the treatment seismic data. Indeed, the seismic data must be firstly processed to achieve a 3D point cloud, which may contain noisy, redundant data, and outliers. Moreover, the point density is likely to be strongly anisotropic, depending on acquisition settings. Generally, the acquisition lines are spaced at pre-defined distance (i.e. 75 m, 100 m, 150 m, ...). In addition, seismic data are expressed in time domain; therefore a time to depth conversion has to be applied. We provide a brief background on seismic surveys and related data in Section 2.

We treat data anisotropy by applying a simplification along lines where points accumulate. We then reconstruct the reference interfaces among soil layers, namely, the seabed and the top and bottom surfaces bounding the deposit, as separate triangle meshes. Finally, we resolve intersections among interfaces, which have no geological meaning, and produce a tetrahedral mesh constrained to the surfaces. We leveraged on available libraries or implement variants of state of the art methods to achieve the final 3D model.

For the estimate of geotechnical measures in unsampled locations, or rather where data is not available, we apply a former method based on Gaussian Simulations, developed for implementing an *adaptive sampling* of environmental variables [BCP*18].

The paper is organised as follows. We first introduce the background on data acquisition for (submarine) soil surveying. Then, we briefly describe relevant approaches to the modeling and visualization of subsoil, and proprietary software in use for corporate geological investigations. The pipeline of our approach is described in Section 4, comprehending data analysis and simplification (Section 4.1), meshing (Section 4.2), and volume reconstruction (Section 4.3). Section 4.4 summarises the approach, based on geostatistics, to estimate geotechnical parameters at unsampled locations over the volume of interest. We tested our approach on a real case study, an offshore resource, for which geophysical (seismic profiles) and geotechnical (parameters from CPTU and laboratory tests on samples) data were available (Section 5). Results of our approach are presented and discussed in Section 6. Future directions conclude the paper.

2. Background on survey data

The investigation of a submarine deposit is carried out by geophysical surveys, mainly based on the ship-based seismic reflection method. The general principle of seismic reflection is based on the emission of seismic pulses by a source and the propagation of the seismic waves through the deposit. Assuming a stratified deposit, some waves are reflected at geological interfaces, according to the change of physical properties of layers. The reflected waves are detected by a transducer (receiver system), which converts the ground motion into electrical signals. Then, the seismic waveforms are recorded and displayed on a seismograph [KBH02].

Different technologies can be used for seismic reflection surveys. For this work, we refer to seismic profiles data (SBP, *Sub Bottom Profile*) acquired by AUV *Autonomous Underwater Vehicle*; this is a compact and easy to handle instrument that allows measurements to be taken autonomously. The device exploits the reflection principle described above: it emits and acquires acoustic waves, moving autonomously in the water and maintaining a uniform flight altitude with respect to the seabed.

Seismic data are acquired along survey lines, which follow a well-established and a priori defined scheme of investigation. The format in which data are stored is known as SEG-Y [seg], one of the standards developed by *SEG Society of Exploration Geophysicists*. SEG-Y is a textual and binary format. It mainly contains a summary of the features of seismic data (i.e.: number of traces, number of trace samples, sample format, sample interval, time length), the trace header (detailed information about the single recorded trace, organized into specific bytes), and the relative trace data (recorded signals amplitude versus time). A pre-processing step is typically required to correct seismic traces and to remove noise with specific geophysical tools. At the end of geophysical processing, SEG-Y files are ready to be analysed. These data are individual 2D representations of the deposit, as vertical sections of the soil. The analysis permits us to recognize shallow structures of the deposit, mapping the first reflectors and also sub-surfaces (see Figure 3). The accuracy in detecting lithological horizons decreases as depth increase.

Seismic surveys are typically integrated by geotechnical investigations, that are used to determine the geotechnical properties of

soils. In our case study, we refer to *Piezicone Penetration Test* (CPTU) [AST20]. In detail, CPTU is an in-situ test that consists of inserting a conical tip into the soil at a constant rate. The test measures the effort required for the tip penetration, the friction that develops on the sleeve, and excess pore pressure generated during the penetration. Typically, these measurements are presented graphically with depth variation. Strength parameters, as friction angle or undrained shear strength (depending on the mechanical behavior of the soil), are estimated by specified correlations. In this case, we refer to geotechnical parameters of undrained shear strength derived from CPTU by correlations, proposed by [RL89]. Laboratory vane shear tests [LAL*11, KHI88], known as *Miniature Vane* [AST16], *Torvane*, *Triaxial* [AST15] and *Direct Simple Shear* [AST17] are also used for the determination of additional measurements of undrained shear strength on terrain samples.

3. Related Works

Traditionally, the trend in geological modeling has been to work mainly with one-dimensional and bi-dimensional data. The geological data processing produces lines, maps and sections (i.e. lines, thematic maps, surfaces, cross-sections, ...), so that, in general, 2D models are able to represent all the information of interest. The visualization and the management of these data are supported by commercial software, organized as *Geographical Information System* GIS, like QGIS [qgi], Global Mapper [glo] or Surfer [sur]. Typically, the team geologists, geophysicists and engineers collaborate to understand the areal extension of a material and the volume conformation of the deposit from diverse, not integrated representations. This way of working is due to the fact that the creation of an integrated and explicit three-dimensional model certainly requires a greater effort and further specific expertise.

The challenge in 3D geological modeling is the reconstruction of a valid model, representing consistent geological configurations and fitting as far as possible the real entities. To this aim, it is often necessary to impose geometric and topological constraints and to apply local editing operations [CCLCdV*09, API*19]. Thus, the transition from 2D to 3D modeling requires understanding geological underground conditions through the analysis and correlation of diverse data, with the appropriate software, to perform consistency checks and achieve an accurate interpretation. From a visualization point of view, a three-dimensional result is immediate, as a layering of identified sub-surfaces during data analysis [PGA*18].

The research in 3D geological modeling is driven by different methodologies and purposes, as described by [ZHZ*12], involving mathematical techniques and visualisation methods. [XL14] implements a method for geological reservoir models, using stratigraphic interfaces as surfaces constraints for the generation of tetrahedral mesh by the construction of geodesic isolines. [WSXM20] proposes a clustering-based bubble method for generating high-quality tetrahedral meshes of geological models. [WMX20] assumes that the input geological model consists of triangulated surfaces and analyses possible geometric defects (self-intersections, isolated points, inconsistent orientations), that are not supported.

Very relevant to our research the *Project GeoMol* focused on developing an open infrastructure to store and exchange multi-dimensional and heterogeneous geo-data at transnational level

[Geo15]. The 3D modelling was supported by a-priori thickness grids of geological units, and the representation itself was an object-relational data base storing geometric entities (points, lines, face and tetrahedron sets) that originate from modelling software like GOCAD [Mal92], Petrel [pet], Move [mov] or AutoCAD [aut].

The interest in volumetric representation firstly gave rise to the idea of extending the functionalities of GIS to the three-dimensional space. [TG07] illustrates methods for domain discretisation, typically used for two-dimensional geological cross section (e.g. orthogonal cellular mesh, quadtree mesh, 2D unstructured triangular mesh) and how they can be extended in 3D case. GIS systems (e.g., QGIS [qgi], GRASS GIS [gra], ArcGIS [arc]) indeed offer functionalities for the management of massive heterogeneous data, and also permit to derive additional information through data cross-correlation. However, we aim to investigate the potential of 3D graphics structures and algorithms to the representation and integration of geological data.

Concerning commercial solutions, beside the traditional three-dimensional graphic modellers such as AutoCAD [aut, GC17] or Rhinoceros [rhi], there are products specifically developed for the geological sector. GEOREKA [geo] targets mineral explorations, and provides tools for 3D digitisation and surface manipulation, according to data-driven modelling approach. Referring to the oil and gas sector, solutions include Petrel [pet] and GOCAD [Mal92]; thanks to their multiple functionalities, they guarantee a complete level of analysis, from data interpretation to the generation of a 3D geological model for simulation.

Kingdom [kin], developed by *IHS Markit*, is the commercial geophysical interpretation software, commonly used in corporate practise. It consists of several packages, each one with specific functionalities and applications for exploration and geo-modelling. It offers tools for basic and advanced interpretation of seismic data, regarding tracking of horizons, identification and interpretation of faults and fractures, localisation of bright spots and generation of grids with different interpolation techniques for the creation of surface maps. The identification of the structures of the deposit is done through the creation of seismic attributes [CM05]: the package contains a library of more than 50 attributes, which can be displayed in 3D space.

The open-source alternative to *Kingdom* is *OpendTect* [ope18]. It is released under a triple licensing scheme (free, commercial, and academic). To the best of our knowledge, the free license provides all the necessary functionalities for processing, interpreting and visualizing seismic reflection data in a 3D working space [AJG17]. It allows the editing and the display of seismic data; concerning the interpretation step, it provides tools for the automatic and manual tracking of horizons and for faults picking. Additional supported functionalities permit to grid points with implemented algorithms, as inverse distance, curvature and triangulation. Also free 3rd party plugins can be loaded and used for advanced applications.

Both software provide similar tools for interpreting seismic profiles. On the one hand, *Kingdom* provides more functionalities, especially for advanced analysis, and is easier to use for data management. On the other hand, *OpendTect* (free version) is less intuitive in the use of various applications and makes some requirements on importing data. In particular, the import tool of *OpendTect* do not

support seismic sections with variable acquisition delay trace by trace, but it requires that all traces have the same start time.

In our work, we adopted the open-source solution to pre-process seismic data and obtain a 3D point cloud, which represents the input to our algorithm.

4. Algorithm Overview

Our methodology consists of a general pipeline for the generation of the 3D model of the subsoil, according to typical procedures conducted in practice for the characterization of any deposit and as an alternative solution of existing, mostly commercial, tools. Mainly, our procedure is divided into three sequential steps, as follows:

1. *Data analysis*: pre-processing and point cloud extraction (see Section 4.1). The input set of seismic data is cleaned (if necessary, depending on the acquisition hardware settings) and horizons corresponding to interesting layer of the deposit are mapped, either manually or automatically.
2. *2D modeling*: point cloud triangulation and surface optimization (see Section 4.2). The point clouds obtained above are triangulated to represent the reference surfaces of the deposit; both geometric and geostatistical techniques are considered to improve the mesh quality.
3. *3D modeling*: volume generation by tetrahedralization (see Section 4.3). The reference meshes are finally used as constraints to generate a tetrahedral mesh where the mapped horizons are explicitly encoded.

Then, we present a geostatistical method to estimate data from other sources, measured at local position, over the whole volume (see Section 4.4).

4.1. Data analysis

Input seismic data refers to the geophysical surveys, acquired for the investigation of a submarine area. In particular, the analysis and the interpretation activities have focused on the recognition of sub-seabed horizons (see Section 5).

As anticipated, all the seismic lines given by the AUV-SBP acquisition are imported into the *OpendTect* software. The import tool specifically supports seismic lines with traces having the same signal start (bytes 105-106) and sample rate; thus, it reads the start and sample rate of the first trace and automatically applies the same values to the remaining traces. This is not always the case. To avoid the loss of information related to erroneous start times in the acquisition and keep the data coherent with reality, we need to modify each SEG-Y file to account for a start time shift, related to the content of bytes 105-106 (*Lag time a*) of the file header. The above processing is carried out in *Seismic Unix*, an open-source seismic utilities package, that provides a series of functionalities suitable for SEG-Y processing [CS19]. Specifically, *su-shift* function is used to translate individual signals, taking into account the variable acquisition delay (*Lag time a*) for each trace. Figure 2 represents an example of shift traces, performed on a generic seismic section.

Once seismic data are imported consistently, we proceed with the interpretation in *OpendTect*. Interpretation consists in tracking

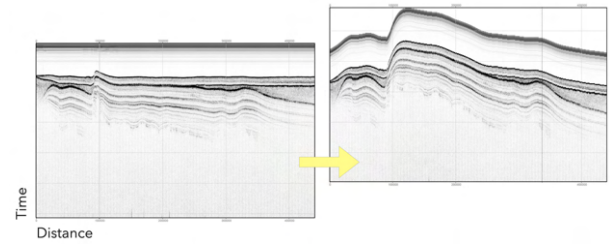


Figure 2: Visualization of a seismic section into 2D viewer of *OpendTect* with default setting import tool (at left). Visualization of the same seismic section, after the shift on single traces (at right).

horizons, i.e., selecting points on acoustic interfaces of interest in a seismic section, for three reference interfaces (in detail seabed, top MTD, and bottom MTD). These are intuitively recognized in almost all seismic sections of AUV-SBP; however, an automatic recognition is not always correct. *OpendTect* provides both an automatic and a manual option. The automatic process selects points that describe the trend of a horizon. The manual option ensures the correct acquisition of the horizon in complex geological zones. An example of horizons tracking on a generic seismic section is shown in Figure 3.

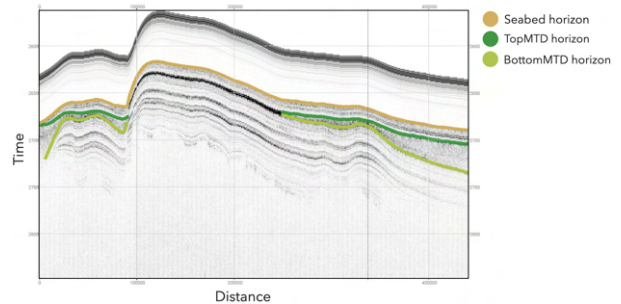


Figure 3: An example of seismic section with horizons tracking; in this section, seabed, top MTD and bottom MTD are recognized.

For each seismic section and selected layer (seabed, top, bottom), points along a horizon are exported in a separate ASCII file. The output files contain triplets of coordinates, such as x,y in meters and z in seconds, for each shot point.

4.2. 2D modeling

Now we create a 2D triangle mesh for each recognized acoustic interface of the deposit, using a well-known state-of-the-art algorithm and geometric processing tools [She96, Liv19]. In this case, the triangulation is chosen because of the geometrical configuration of point data, that are not placed on regular grid. The technique guarantees a more accurate modeling of the complexity of terrain surfaces and also it is more adaptable for representing zones with different density of points.

We load the ASCII files for all the horizons corresponding to the same level and, since we don't expect undercuts, we compute

the Delaunay triangulation on the horizontal plane (x,y). Then the elevation of each point is restored by dropping z value. Lacking an explicit boundary including all points, the convex hull is used as external constraint for triangulation.

Triangulating the data as they are, however, is likely to generate a high amount of almost degenerate triangles (see Figure 4), which are notoriously prone to numerical errors. Typically for the submarine acquisition surveys, data show a strong spatial anisotropy, with densely packed points along the seismic acquisition lines. The seismic lines, in turn, are unevenly spaced, e.g., along two principal directions of investigation.

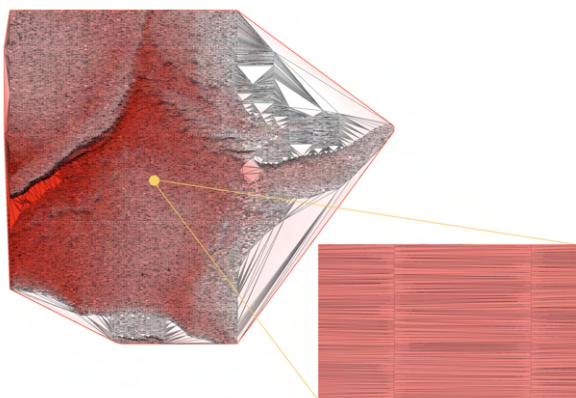


Figure 4: A detail about almost degenerate triangles on seabed mesh. The low resolution of the image is the maximum obtainable, to demonstrate the configuration of triangles prone to numerical error.

Two optimization strategies, described in Section 4.2.1 and 4.2.2 respectively, are proposed in order to obtain a result that is as isotropic as possible, and to limit volume generation problems at a later stage.

4.2.1. Data Simplification

Since the point cloud is structured into (nearly straight) lines of very close points, we decided to first apply a simplification method specific for this configuration to remove redundant point along each line, i.e., the *Douglas-Peucker* algorithm [DP73]. The algorithm produces an approximation of an original planar curve, based on points decimation.

Perpendicular distance is calculated between the straight line, that connects the first and last points of the set and each internal point, and it is compared with a threshold (tolerance), as the criterion for point removal. If the perpendicular distance is lower than the tolerance, the point is removed from the list; otherwise, the line is split in two at that point, and the algorithm repeats recursively on the two subsets. The result of the algorithm is a simplified curve that preserves relevant original information.

To simplify each line, we first project points on their average vertical plane. In the case of acquisition lines almost parallel to the N-S and to the E-W principal directions, as in our case study (see

Section 5), that corresponds to perform the simplification on the (x,z) and the (y,z) plane, respectively.

The tolerance is chosen according to the variability of z coordinate of the points. For giving a quantification on z coordinate, that is an important approach in geological applications, we modify the simplification algorithm and replace the point to line distance along the perpendicular direction with the vertical distance. Vertical distance is calculated as the difference in z coordinates of a point and its vertical projection on the straight-line between start and end points. This version of the algorithm facilitates the quantification of the committed error in z during the simplification; the maximum error is represented by the tolerance. A scheme of the *Douglas-Peucker* algorithm is shown in Figure 5, distinguishing the two implemented distances.

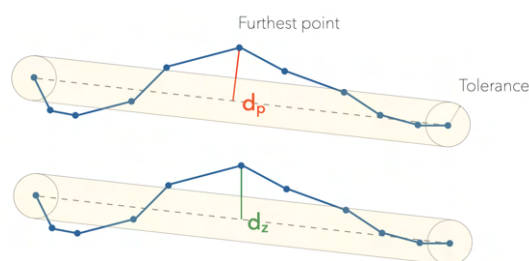


Figure 5: A scheme of Douglas-Peucker algorithm with the representation of both perpendicular (d_p) and vertical (d_z) distance.

From a geometrical point of view, the meshes obtained from simplified data maintain the most relevant points and are much more regular, but still have potential for improvement. Long and thin triangles still exist between acquisition lines. However, refinement in-between adjacent acquisition lines requires adding depth information where no samples are available.

4.2.2. Mesh Refinement

The triangle library [She96] already provides the functionality to refine a mesh, including options to control the final quality through geometric constraints. Therefore, it is used to refine meshes, setting the quality constraint of a minimum angle of 20° . Again, we perform the mesh refinement in 2D, inserting new points on the (x,y plane). Now, adequate interpolation techniques are necessary to estimate the z coordinate of each added point consistently with the nearby sampled points. The choice made in this context concerned the use of geostatistical techniques.

Indeed, geostatistical analysis allows the evaluation of the spatial correlation of regionalized variables. The degree of spatial dependence of observations is quantified by the *experimental variogram*. It is a graphical representation, generally in a scatter plot, of the relative distance between each pair of points (on the x axis) and the corresponding value of semi-variance (on the y axis). From a quantitative point of view, the semi-variance is obtained by the followed formula, for a given h value:

$$\gamma(h) = \frac{1}{2} \text{var}[Z(x) - Z(x + \mathbf{h})] = \frac{1}{2} \sum [((Z(x) - Z(x + \mathbf{h}))^2]$$

where \mathbf{h} is the vector representing a directional distance between points (lag).

The variogram depends only on the relative euclidean distance \mathbf{h} between points. Experimental variogram can be best-fitted by a mathematical function (variogram model), expressed in terms of three parameters, namely range, sill value and nugget effect [Cre85]. In detail, the nugget effect is the variance of the spatial dataset at short distance below the minimal point distance (for $h \rightarrow 0$); the sill represents the dataset variance due to the point configuration up to the asymptotic value reached by the variogram model; consequently the sum of nugget and sill is equal to the variance of the entire dataset regardless the spatial positions; the range is the distance at which the sill is reached, and beyond that distance the data are no longer spatially correlated.

After variography, we proceed to the interpolation of known points to obtain a prediction of the z value of points introduced in the refinement phase. In this context, interpolation by *kriging* [WO07] guarantees a reasonable estimate with associated uncertainty of the variable under study. *Kriging* is a regression method, which calculates the value to be assigned $z^*(x_0)$ to the point x_0 by a linear combination of the sampled values $z(x_\alpha)$, nearby to the point to be estimated. It has the following mathematical formulation:

$$z^*(x_0) = \sum_{\alpha=1}^n \lambda_\alpha z(x_\alpha)$$

where λ_α are the coefficients of linear interpolation, n is the number of nearby observations; * symbol indicates the linear estimator considered. The λ_α coefficients are determined minimizing the average quadratic error; they depend on reciprocal positions between sampled points and on the variogram model.

Assuming to consider a Z random function at each point, defined as the sum of deterministic trend component (or *drift*) and stochastic residual [Wac03], sampled data are decomposed into the two mentioned parts. The trend represents the behaviour at a large scale and is responsible for the non-stationary condition; while the residual represents the local short-scale variability. To satisfy the assumption of stationarity, the common practice is to remove drift from random variable.

$$Z(u) = m(u) + R(u), \forall (x) \in R$$

$$E(R(u)) = 0$$

$$E(Z(u)) = m(u)$$

The first step concerns the removal of the drift from the z coordinate of points, using a polynomial function of one degree, that averages out the behaviour of the deterministic component of the points. The following formula represents the equation of the plane that best fits sampled points, solving the least-squares problem.

$$m(x, y) = a_0x + a_1y + b$$

where a_0 , a_1 , b are the coefficients determined by least square minimization. Then, residuals are determined as the difference between the z coordinate of the generic point and its position on the plane approximating sampled points.

Geostatistical techniques used for this application require gaussianity and normality in the data distribution; to ensure this condition, the *Normal Score* function is used to transform original data into a normal distribution. After analysing the behaviour of the residuals from a statistical point of view, the steps previously described (from analysis of spatial correlation to interpolation by kriging) are applied on the normalised residuals. The final result is an estimate of the normal residual values and the relative variance for the new points added. Then, a *Back Transforming* function restores the original data distribution. Knowing the residual estimated and the value assumed by the polynomial function in the generic point added, it is possible to calculate the z value (magnitude in seconds) and the relative estimation error at the new point.

The refined meshes show a more regular configuration, eliminating the majority of long and thin triangles. The geostatistical approach allows to insert new points with a known estimation error with respect to the depth coordinate (see Figure 10).

4.3. 3D modeling

The volume generation exploits the state-of-the-art *TetGen* library [Si15]. Starting from the surfaces produced for each interface of interest of the deposit as described in Section 4.2, the goal is to fill the space between triangular meshes with tetrahedral elements.

Firstly, additional pre-processing is required to prepare the data for the tetrahedralization, and further geometrical checks are performed to guarantee geological consistency and avoid errors in the following volume generation. The developed procedure consists of two steps, as follows.

1. *Domain closure* definition of a common external boundary.
2. *Check for geological consistency* solving mesh intersections to avoid stratigraphic incoherency.

The need to define a common external boundary derives from the features of the obtained triangular meshes. The surfaces may present a different coverage area, due to the existence or not of all the layers in the seismic sections. Also, we do not have a well-defined bounding box, which includes all surveys. Thus, we define a common working domain, in order to facilitate the subsequent process of lateral closure of the meshes.

Typically, the seabed surface exhibits the widest area coverage, while deeper surfaces are smaller. We could choose to define the domain either as the widest or the smallest covered area across the three layers. On the one hand, the smallest area excludes a portion of available data; on the other hand, the widest includes unsampled areas. We decide to go for the smallest; of course the whole reconstructed surfaces will contribute to the visualization of the deposit. We work in the 2D plane (x - y). Firstly, we extrapolate new sub-datasets for the other surfaces (seabed and top MTD), considering only points that are included in the 2D polygon representing the chosen external constraint. Then, the common boundary is defined as the union of the convex hulls of the subsets. Figure 6 shows a scheme of the described procedure. The elevation values of the boundary points for each sub-dataset are defined by associating the average of z values of adjacent points.

Geological consistency checks, instead, concern the recognition

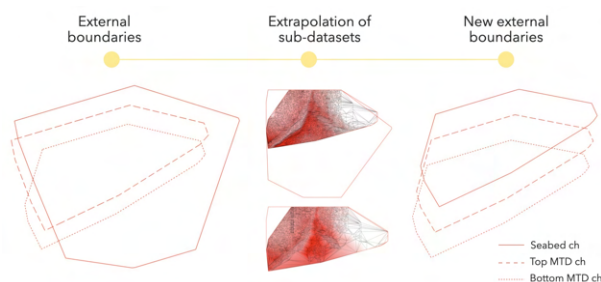


Figure 6: A schematic representation of the steps to define the common external boundary for the meshes.

and the identification of local errors, and their treatment. Typical errors are related to intersections between surfaces that are inconsistent with the geological context. The motivations are linked to the propagation of unintentional errors in the interpretation of geophysical data (e.g., manual errors during the tracking of horizons, noise and outliers in the data) and to particular configurations of surfaces, that may be very close in the z direction. However, layers cannot physically interpenetrate, but rather collide with one layer disappearing. In this case, two horizons bounding a unit tend to coincide into one interface. In this work, we assume horizons are separate, to avoid non-manifold configurations.

Cleaning is done by measuring the vertical distance (with sign) between each vertex of a mesh and the closest one on the adjacent mesh below. The sign of the distance indicates if, locally, the correct layers sequence is preserved; otherwise, the vertex is removed because its position represents an inconsistent geological condition. Figure 11 shows an example of the results obtained by the definition of the common external boundary and the application of consistence checks on the reference meshes.

Finally, since TetGen requires a closed input mesh, it is necessary to close laterally the survey area. So, we create a correspondence map between boundary points of adjacent layers: the closure consists in connecting pairs of corresponding vertices and adding the corresponding triangles to create the lateral surface. Beside misplaced vertices, triangles from different layers can still intersect when two surfaces are locally very close. The tool developed and proposed in [CLSA20] permits to solve the intersections between surfaces. Firstly, the method identifies intersecting elements, and detects intersection points and segments. Then, local geometric editing operations are performed (split triangles, split edges) to solve the intersections on closed mesh.

4.4. Integration of geotechnical data

The volumetric model can be used to visually inspect the morphology of the resources of interest and to extract quantitative features characterizing the deposit, e.g., volume, thickness, maximum slope. However, so far, geophysical data are only considered. Now we want to integrate further heterogeneous data coming from other investigations of the same region, possibly given as local and sparse measurements performed on site. Similarly to the z estimates in the mesh refinement phase (Section 4.2), we leverage on geostatistical

techniques to predict distribution of geotechnical parameters (e.g. estimation of undrained shear strength from CPTU and vane tests).

To do so, we adopt a method developed for the prediction of the distribution of environmental parameters or pollutants in water volumes [BCP*18]. In that work, the estimation and related uncertainty was used to drive the sampling sequence in an adaptive fashion. In our case, we simply provide the input measurements at a set of sparse 3D points obtaining, for each tetrahedron of the domain, an estimate and a related uncertainty value. By colour-coding tetrahedra we can communicate the distribution of geotechnical dimensions in unsampled locations. We refer the interested reader to [BCP*18] for further details. An example of the estimation is provided in Section 6 for a real case study.

5. Case Study

Our case study concerns the analysis and the interpretation of heterogeneous data, acquired for the investigation of a submarine area, located 200 km from shore. Water depth is variable from 1940 m to 2140 m. All available data are provided by *RINA Consulting*, external partner for this research project. Because this application is related to a real survey campaign, specific information about area localization and geographic coordinates are classified and cannot be disclosed.

5.1. Input Dataset

The seismic profile dataset consists of AUV-SBP data; in particular, it includes 160 N-S lines with a spacing of 150 m, and 4 almost perpendicular E-W lines, spaced about 5 km (see Figure 7). Data are supplied in SEG-Y format.

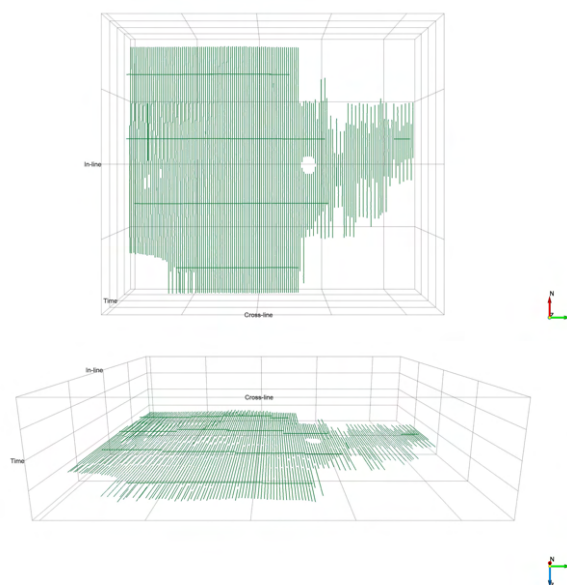


Figure 7: Disposition of seismic acquisition lines in survey area on (N-E) and (E-Z) plane in the visualization environment of OpenTect.

Files contain information about survey features and requirements. Each SEG-Y file describes a vertical section of the subsoil, with the z coordinate expressed as times (milliseconds); the horizon mapping is then performed in the time domain; a time-depth conversion would be necessary for metric evaluation.

Beside the geophysical dataset the results of geotechnical surveys by means of CPTU are available; the investigation area contains 9 CPTU and 5 boreholes. For each CPTU location, we have information about tip resistance (q_c , kPa), sleeve friction (f_s , kPa), and pore pressure response (u_2) measured during the test. For boreholes, we have a collection of undrained shear strength (S_u , kPa) at different depths, measured by vane tests.

6. Results

The input seismic dataset (AUV-SBP data) has been interpreted in *OpenTect*, as discussed in Section 4.1; the recognized interfaces of the deposit are distinguished as seabed, top MTD, and bottom MTD. In detail, MTD or *Mass Transport Deposit* indicates a buried geological body, composed of a heterogeneous material, which presents a chaotic behaviour from an acoustic point of view. The tracking of horizons (with both automatic and manual options, depending on the complexity of mapping horizons), generates a point cloud consisting, respectively, of about 3.2 millions points for the seabed, 1 million points for the top MTD, and 700000 points for the bottom MTD. The points are consequently processed, to remove possible duplicates, due to issues of numerical precision.

Surface meshing is performed with [Liv19] [She96] using the *Delaunay* criterion and choosing the convex hull as the external boundary. The variability of the z coordinate in seconds is limited with respect to the metric coordinates, and therefore scarcely perceptible in a 3D visualization environment. Associating a color scale to z values of the vertices makes it possible to communicate the variability more intuitively. The color scale is defined from red (minimum z values) to white (maximum z values). Figure 8 shows meshes, obtained by data triangulation for each dataset.

Line simplification (see Section 4.2.1) is performed for each geological level and varying tolerance value. The highest value of tolerance corresponds to the highest reduction of points. Figure 9 compares the original mesh with the result obtained with simplified data configuration with a tolerance of 0.001 s. In this case, the simplification, with a tolerance of 0.001 s, reduces the number of points down to about 9000 for the seabed, 7000 for the top MTD and 3000 for the bottom MTD.

As described in Section 4.2.2, mesh refinement is carried out according to a geometric criterion, included in the command line switches for triangular mesh optimisation in *Triangle* library. In our case, refinement is applied on the simplified meshes, setting the constraint of a minimum angle of 20° . Figure 10 represents triangular meshes obtained respectively with simplification (tolerance of 0.001 s) and both simplification (tolerance of 0.001 s) and refinement strategies for the reference layers.

We proceed with volume generation (see Figure 11) as described in Section 4.3 using either the simplified or the simplified and refined meshes. Figure 12 shows the volume generated by using simplified meshes with the tolerance of 0.001 s.

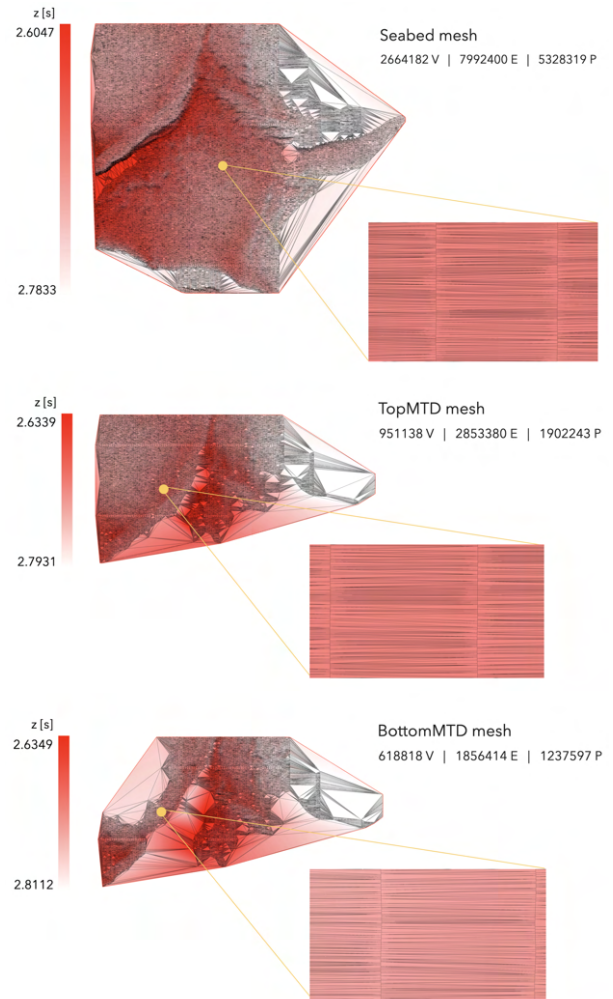


Figure 8: 2D meshes and relative details for each horizon.

Finally, to test the potential for integrating geotechnical data, we apply the geostatistical analysis to a set of samples. We built a synthetic field representing the distribution of a geotechnical parameter like the undrained shear strength (S_u). The collected samples were used to estimate the value of S_u over the whole mesh. The choice of samples locations was adaptively driven to minimize the uncertainty of the parameter estimate [BCP*18]. The map in Figure 13 was created by applying sequential Gaussian simulations [WO07] using 40 samples (20 randomly chosen and 20 adaptively chosen). This procedure shows the possibility of combining volumes obtained from geometric reconstruction from seismic records with data from geotechnical investigations to obtain an integrated model of the study area.

7. Limitations and ongoing work

At this stage of our research project, we have developed the full pipeline, from the interpretation of the seismic data to the volume generation. We also performed initial tests for the computation of

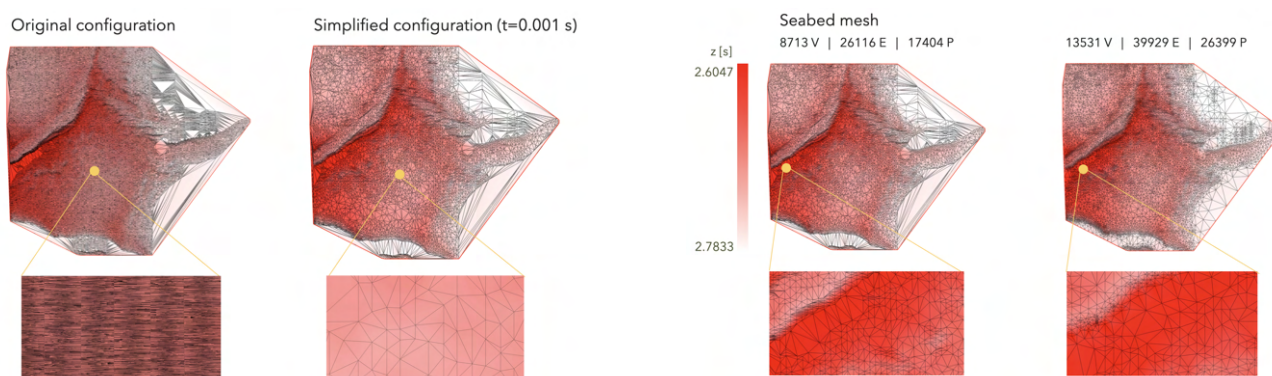


Figure 9: Example of the application of simplification strategy: comparison between original and simplified meshes for seabed horizon at the same level of detail.

the distribution of geotechnical data on the whole volume, even if, for the time being, we couldn't experiment with the real CPT data. These are given either as numeric values in tabular form for a given x-y position at varying z, or as graphic plots. Therefore, before applying the geostatistical approach, we need to extract numerical values from the graphic plots, but most importantly, we observe that depth is expressed in meters in the geotechnical surveys, whereas we have been treating times so far.

In order to correlate geotechnical data (z in meters) and seismic data (z in seconds), it is necessary to apply a time-depth conversion, by means of an appropriate velocity model. The velocity profile is generally not known a-priori [KBH02]; it varies with depth and depends on properties of individual layers, as in particular the density and the elastic moduli. Generally, a seismic inversion process is used to estimate velocity profile, starting from the assumption of the quantitative properties of the stratified deposit. Therefore, to simplify the conversion procedure, which requires an in-depth study of the propagation velocities of the subsurface, an average shear wave velocity of approximately 1550 m/s can be applied. This is a normal practice in the case of expedite studies to obtain an order of magnitude of absolute depths below sea bottom. Using the approximate velocity model, we are confident to achieve the estimation of undrained shear strength (S_u , kPa) on the whole volume of investigation smoothly.

The time / depth discrepancy also affects the visualization, as is evident in many figures throughout the paper. So far, we scaled z by a factor to perceive the correct surface morphology; applying the conversion to meters will also improve visualisation.

Another open issue is the correct segmentation of the volumetric layers. Indeed, as surfaces intersect, we currently solve the issue from the geometric point of view, by splitting triangles, not from the structural point of view. We are going to detect volumetric clusters where the correct layering order is not maintained and prune them away, locally collapsing the surfaces.

Finally, we encountered unexpected errors with Tetgen when generating the tet mesh from refined meshes. Apparently, the re-

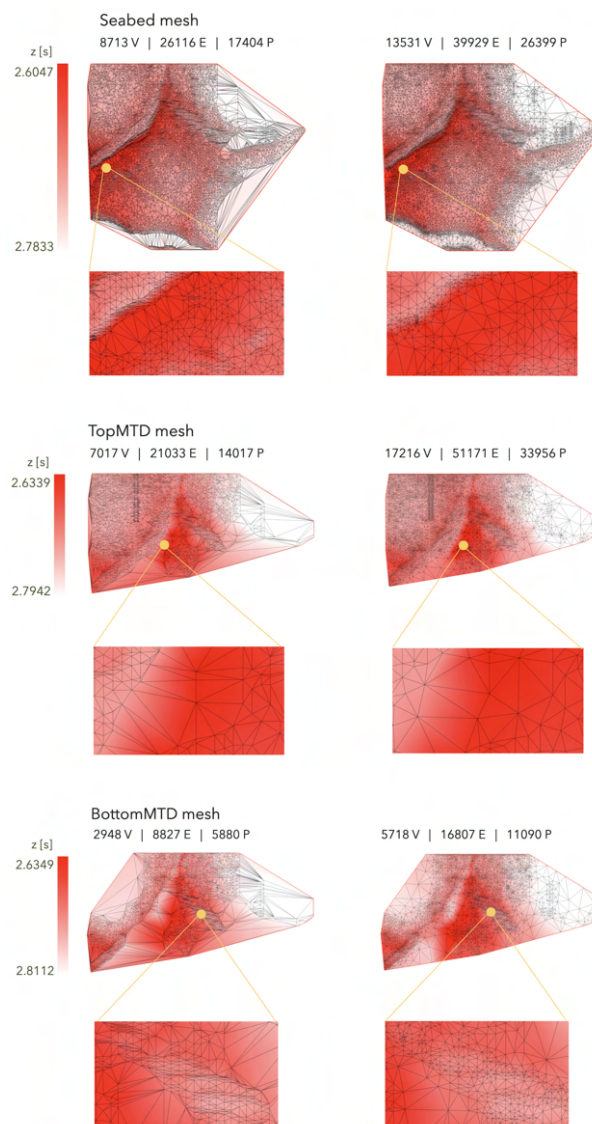


Figure 10: Comparison between simplified meshes, obtained with a tolerance of 0.001 s (left), and simplified (with the same tolerance) and refined meshes (right).

finement produces nearly flat tetrahedra to appear, especially near the borders. We are still investigating this issue.

8. Conclusions

In this work, we presented a pipeline for the representation of sub-soil layers and in particular for the reconstruction of a volumetric model of a mass deposit of interest. We described the modeling issues proper of the geological domain and data, and addressed them by combining or customising available libraries and state-of-the-art methods. Our focus was on the integration of heterogeneous data, proper of this context; specifically, we used geophysical (seismic) data to build the 3D representation of the domain, and describe our

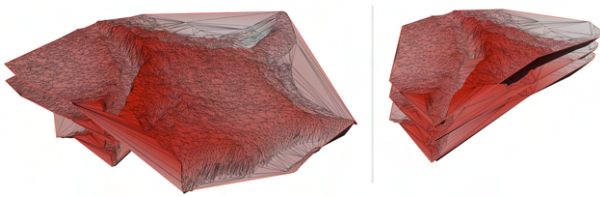


Figure 11: Visualization of the three reference meshes with their coverage area (left) and after definition on common external boundary and application of consistence checks (right). A scale factor and also a translation on z coordinate are applied only for a better visualization of each singular surfaces.

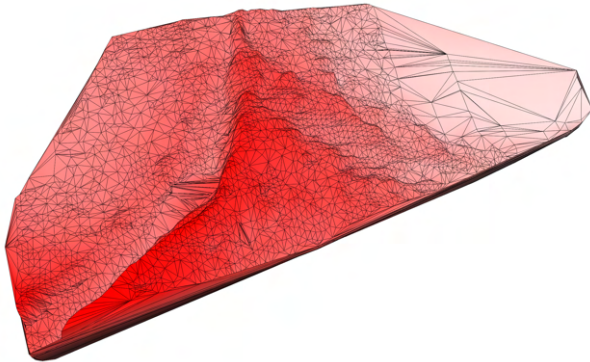


Figure 12: Tetrahedral volume. A scale factor is applied on z coordinate for a better visualization of the morphology.

approach to integrate it with geotechnical data, through geostatistical analysis. Finally, we discuss results obtained on a real case study of an underwater deposit.

As future improvements, a clean segmentation of the volumetric layers will allow a quantitative characterisation of stratigraphic units (e.g., volume, local thickness, slope, approximate weight if information about the material is known). A graphic interface with several visualisation facilities will be developed, in order to query the local properties of the deposit and access geotechnical estimations and related uncertainty in an intuitive manner.

We need to test our approach on further data sets, especially on seismic data acquired following different patterns, to prove the

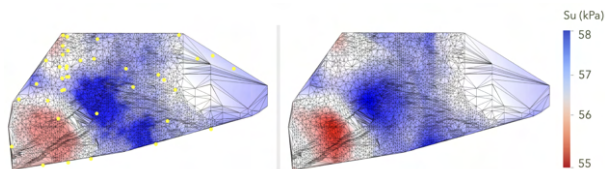


Figure 13: Estimation of a geotechnical parameter: the synthetic field (right) and the estimated value of the variable sampled at the positions displayed by yellow dots (left).

method useful in the general case. Possibly the variogram could give indications about whether to simplify and/or refine data and along which directions, to make the process fully automatic.

Acknowledgements

This work has been partially supported by the GEO3D project ("Metodi e modelli 3D volumetrici per la rappresentazione, analisi e visualizzazione di dati eterogenei risultanti da campagne geognostiche atte alla caratterizzazione litostratigrafica di siti di interesse, con applicazioni alla progettazione di opere civili ed impianti energetici"), POR FSE Liguria 2014-2020 - Asse 3 "Istruzione e Formazione" Programma Operativo Regione Liguria Fondo Sociale Europeo 2014-2020, Obiettivo Specifico 10.5 "Innalzamento dei livelli di competenze".

The authors also thank Serena Berretta for her contribution in implementing geostatistical techniques, related to the project "Monitoraggio Adattivo in Tempo reale con Automatizzazione del Campionamento - Aree Costiere Portuali - MATRAC-ACP". This project is funded by the EU Interreg V-A Italia Francia Marittimo 2014-2020 - Asse prioritario del Programma 2 - Protezione e valorizzazione delle risorse naturali e culturali e gestione dei rischi - Obiettivo specifico della Priorità d'Investimento 6C2-Accrescere la protezione delle acque marine nei porti.

References

- [AJG17] AZIZ I. A., JAAFAR J., GILAL A. R.: The Study of OpenDetect Seismic Data Interpretation and Visualization Package in Relation to Seismic Interpretation and Visualization Models. *IJCSNS* 17, 10 (2017), 124. 3
- [API*19] ANQUEZ P., PELLERIN J., IRAKARAMA M., CUPILLARD P., LÉVY B., CAUMON G.: Automatic correction and simplification of geological maps and cross-sections for numerical simulations. *Comptes Rendus Geoscience* 351, 1 (2019), 48–58. doi:<https://doi.org/10.1016/j.crte.2018.12.001>. 3
- [arc] GIS Mapping Software, Location Intelligence & Spatial Analytics | Esri. <https://www.esri.com/en-us/home>. Accessed: 17/09/2021. 3
- [AST15] ASTM D2850-15. Standard Test Method for Unconsolidated-Undrained Triaxial Compression Test on Cohesive Soils. ASTM International, West Conshohocken, PA, 2015. doi:[10.1520/D2850-15](https://doi.org/10.1520/D2850-15). 3
- [AST16] ASTM D4648 / D4648M-16. Standard Test Methods for Laboratory Miniature Vane Shear Test for Saturated Fine-Grained Clayey Soil. ASTM International, West Conshohocken, PA, 2016. doi:[10.1520/D4648_D4648M-16](https://doi.org/10.1520/D4648_D4648M-16). 3
- [AST17] ASTM D6528-17. Standard Test Method for Consolidated Undrained Direct Simple Shear Testing of Fine Grain Soils. ASTM International, West Conshohocken, PA, 2017. doi:[10.1520/D6528-17](https://doi.org/10.1520/D6528-17). 3
- [AST20] ASTM D5778-20. Standard Test Method for Electronic Friction Cone and Piezocone Penetration Testing of Soils. ASTM International, West Conshohocken, PA, 2020. doi:[10.1520/D5778-20](https://doi.org/10.1520/D5778-20). 3
- [aut] Autodesk | 3D Design, Engineering & Construction Software. <https://www.autodesk.com>. Accessed: 17/09/2021. 3
- [BCP*18] BERRETTA S., CABIDDU D., PITTALUGA S., MORTARA M., SPAGNUOLO M., ZUCCOLINI M. V.: Adaptive Environmental Sampling: The Interplay Between Geostatistics and Geometry. In *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference* (2018), Livesu M., Pintore G., Signoroni A., (Eds.), The Eurographics Association. doi:[10.2312/stag.20181306](https://doi.org/10.2312/stag.20181306). 2, 7, 8

- [CCLCdV*09] CAUMON G., COLLON P., LE CARLIER DE VESLUD C., VISEUR S., SAUSSE J.: Surface-Based 3D Modeling of Geological Structures. *Mathematical geosciences* 41 (11 2009), 927–945. doi:10.1007/s11004-009-9244-2. 3
- [CLSA20] CHERCHI G., LIVESU M., SCATENI R., ATTENE M.: Fast and Robust Mesh Arrangements Using Floating-Point Arithmetic. *ACM Trans. Graph.* 39, 6 (Nov. 2020). doi:10.1145/3414685.3417818. 7
- [CM05] CHOPRA S., MARFURT K. J.: Seismic attributes — a historical perspective. *Geophysics* 70, 5 (09 2005), 3SO–28SO. doi:10.1190/1.2098670. 3
- [Cre85] CRESSIE N.: Fitting variogram models by weighted least squares. *Journal of the International Association for Mathematical Geology* 17 (1985), 563–586. 6
- [CS19] COHEN J. K., STOCKWELL J. J. W.: CWP/SU: Seismic Un*x Release No. 4.6: an open source software package for seismic research and processing. <https://nextcloud.seismic-unix.org/index.php/s/wZNoFCfJdXj5iKB>, 2019. 4
- [DP73] DOUGLAS D. H., PEUCKER T. K.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization* 10, 2 (1973), 112–122. doi:10.3138/FM57-6770-U75U-7727. 5
- [FFDB20] FULLER A., FAN Z., DAY C., BARLOW C.: Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access* 8 (2020), 108952–108971. doi:10.1109/ACCESS.2020.2998358. 1
- [GC17] GONZÁLEZ CHIOZZA S.: Computer aided geological design: AutoCAD® applications in structural geology. *Terrae Didactica* 13 (08 2017), 113. doi:10.20396/td.v13i2.8650087. 3
- [geo] 3D Geological Modelling Software | Affordable | Georeka Software. <https://www.georeka.com/>. Accessed: 17/09/2021. 3
- [Geo15] GEOMOL TEAM: GeoMol – Assessing subsurface potentials of the Alpine Foreland Basins for sustainable planning and use of natural resources. Project Report, 188 pp. https://www.geomol.eu/geomol/report/GeoMol_Report_web_reduced.pdf, 2015. 3
- [glo] Global Mapper - Blue Marble Geographics. <https://www.bluemarblegeo.com/global-mapper/>. Accessed: 17/09/2021. 3
- [gra] GRASS GIS - Bringing advanced geospatial technologies to the world. <https://grass.osgeo.org>. Accessed: 17/09/2021. 3
- [KBH02] KEAREY P., BROOKS M., HILL I.: *An introduction to geophysical exploration*, vol. 4. John Wiley & Sons, 2002. 2, 9
- [KHI88] KOLK H. J., HOOPE J., IMS B. W.: Evaluation of Offshore In-Situ Vane Test Results. In *Vane Shear Strength Testing in Soils: Field and Laboratory Studies*. ASTM International, West Conshohocken, PA, 1988. doi:10.1520/STP10341S. 3
- [kin] Kingdom: Seismic and Geological Interpretation Software. <https://ihsmarkit.com/products/kingdom-seismic-geological-interpretation-software.html>. Accessed: 17/09/2021. 3
- [LAL*11] LUNNE T., ANDERSEN K., LOW H. E., RANDOLPH M., SJURSEN M.: Guidelines for offshore in situ testing and interpretation in deepwater soft clays. *Canadian Geotechnical Journal* 48 (2011), 543–556. doi:10.1139/t10-088. 3
- [Liv19] LIVESU M.: Cinolib: A Generic Programming Header Only C++ Library for Processing Polygonal and Polyhedral Meshes. *Transactions on Computational Science XXXIV* (2019). <https://github.com/mlivesu/cinolib/>. doi:10.1007/978-3-662-59958-7_4. 4, 8
- [Mal92] MALLETT J.: Gocad: A computer aided design program for geological applications. In *Three-dimensional modeling with geoscientific information systems*. Springer, 1992, pp. 123–141. 3
- [mov] MOVE, Structural Geology Modelling Software. <https://www.petex.com/products/move-suite/>. Accessed: 17/09/2021. 3
- [ope18] OpendTect User Documentation - 6.2. https://doc.opendtect.org/6.2.0/doc/od_userdoc/Default.htm, 2018. 3
- [pet] Petrel E&P Software Platform. <https://www.software.slb.com/products/petrel>. Accessed: 17/09/2021. 3
- [PGA*18] PAN X., GUO W., AUNG Z., NYO A., CHIAM K., WU D., CHU J.: *Procedure for Establishing a 3D Geological Model for Singapore*. 05 2018, pp. 81–89. doi:10.1007/978-981-13-0011-0_9. 3
- [qgi] QGIS. <https://www.qgis.org/it/site/>. Accessed: 17/09/2021. 3
- [rhi] Rhino - Rhinoceros 3D. <https://www.rhino3d.com/it/>. Accessed: 17/09/2021. 3
- [RL89] RAD N., LUNNE T.: Direct correlations between piezocone test results and undrained shear strength of clay. *Publikasjon-Norges Geotekniske Institutt*, 177 (1989), 1–7. 3
- [seg] SEG Technical Standards. <https://library.seg.org/seg-technical-standards>. Accessed: 17/09/2021. 2
- [She96] SHEWCHUK J. R.: Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In *Applied Computational Geometry: Towards Geometric Engineering*, Lin M. C., Manocha D., (Eds.), vol. 1148 of *Lecture Notes in Computer Science*. Springer-Verlag, May 1996, pp. 203–222. From the First ACM Workshop on Applied Computational Geometry. 4, 5, 8
- [Si15] S1 H.: Tetgen, a delaunay-based quality tetrahedral mesh generator. *ACM Transactions on Mathematical Software (TOMS)* 41, 2 (2015), 1–36. 6
- [sur] Surfer | 2D & 3D mapping, modeling & analysis software for scientists and engineers. <https://www.goldensoftware.com/products/surfer>. Accessed: 17/09/2021. 3
- [TG07] TURNER A. K., GABLE C. W.: A review of geological modeling. *Three-dimensional geologic mapping for groundwater applications* (2007), 81. 3
- [Wac03] WACKERNAGEL H.: Geostatistical models and kriging. *IFAC Proceedings Volumes* 36, 16 (2003), 543–548. 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands, 27-29 August, 2003. doi:https://doi.org/10.1016/S1474-6670(17)34818-8. 6
- [WMX20] WANG B., MEI G., XU N.: Method for generating high-quality tetrahedral meshes of geological models by utilizing cgal. *MethodsX* 7 (2020), 101061. doi:https://doi.org/10.1016/j.mex.2020.101061. 3
- [WO07] WEBSTER R., OLIVER M. A.: *Geostatistics for environmental scientists*. John Wiley & Sons, 2007. 6, 8
- [WSXM20] WANG B., SUN Y., XU N., MEI G.: A Clustering-Based Bubble Method for Generating High-Quality Tetrahedral Meshes of Geological Models. *Applied Sciences* 10, 15 (2020). doi:10.3390/app10155292. 3
- [XL14] XING H., LIU Y.: Mesh Generation for 3D Geological Reservoirs with Arbitrary Stratigraphic Surface Constraints. *Procedia Computer Science* 29 (2014), 897–909. 2014 International Conference on Computational Science. doi:https://doi.org/10.1016/j.procs.2014.05.081. 3
- [ZH*12] ZU X. F., HOU W. S., ZHANG B. Y., HUA W. H., LUO J.: Overview of Three-dimensional Geological Modeling Technology. *IERI Procedia* 2 (2012), 921–927. International Conference on Future Computer Supported Education, August 22-23, 2012, Fraser Place Central - Seoul. doi:https://doi.org/10.1016/j.ieri.2012.06.192. 3