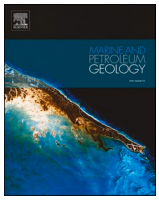
Marine and Petroleum Geology 153 (2023) 106283

Contents lists available at ScienceDirect

Marine and Petroleum Geology

journal homepage: www.elsevier.com/locate/marpetgeo



Using python libraries and k-Nearest neighbors algorithms to delineate syn-sedimentary faults in sedimentary porous media

Manuel Martín-Martín [a,*], Manuel Bullejos [b], David Cabezas [c], Francisco Javier Alcalá [d,e]

[a] Departamento de Ciencias de la Tierra y Medio Ambiente, University of Alicante, AP 99, 03080, Alicante, Spain
[b] Departamento de Álgebra, University of Granada, 18010, Granada, Spain
[c] Departamento de Análisis Matemático, University of Granada, 18010, Granada, Spain
[d] Departamento de Desertificación y Geo-Ecología, Estación Experimental de Zonas Áridas (EEZA–CSIC), 04120, Almería, Spain
[e] Instituto de Ciencias Químicas Aplicadas, Facultad de Ingeniería, Universidad Autónoma de Chile, 7500138, Santiago, Chile

ARTICLE INFO

Keywords:
Python libraries
KNN algorithm
Syn-sedimentary faults
3D stratigraphic architecture
Llobregat river delta
NE Spain

ABSTRACT

This paper introduces a methodology based on Python libraries and machine learning k-Nearest Neighbors (KNN) algorithms to create an interactive 3D HTML model (3D_Vertical_Sections_Faults_LRD.html) that combines 2D grain-size KNN-prediction vertical maps (vertical sections) from which syn-sedimentary faults and other features in sedimentary porous media can be delineated. The model can be visualized and handled with conventional web browsers. The grain-size physical parameter is measurable, constant over instrumental time, handleable mathematically, and its range can be associated to lithological classes. Grain-size input data comes from a public database of 433 boreholes in the Llobregat River Delta (LRD) in NE Spain. Four lithological classes were defined: Pre-Quaternary basement, and Quaternary gravel, sand, and clay–silt. Using a new KNN-prediction algorithm, seven NW–SE (transversal) and three SW–NE (longitudinal) vertical sections were created following the orientation of faults identified in surface and detected in reflection seismic geophysical surveys. For exploratory K values in the 1–75 range were used. K around 25 provides the general and smoothy shape of the basement top surface, whereas $K = 1$ is a optimal value to represent the heterogeneity of the LRD at short distance. Using a new KNN-prediction confidence algorithm inspired in the Similarity Ratio algorithm for machine-learning KNN, the vertical sections overall confidence was evaluated as satisfactory. A general decreasing confidence trend according to the decreasing data density with depth and from inland to seaward was found. The vertical sections created with $K = 1$ show horizontal interruptions (displacements or vertical steps) in the basement continuity and in the Quaternary coarse bodies (gravel and sand) attributable to the action of Quaternary active faults. These faults have been linked or correlated with well-known active faults in the area related in much cases with the Valencia Trough opening. Moreover, several faults detected in surface and other identified in this paper by the first time have been revealed as fault zones made of fault branches with different steps in an echelon-like arrangement. Faulting seems to be more evident in the Pleistocene Lower Detrital Complex and much less active or inactive in the Holocene Upper Detrital Complex. Syn-tectonic gravel channels faulty controlled, progradation of gravel lobes, and lateral migration of channel bars were also observed. At its current development stage, this methodology could also be applied to other geological environments, making the due minor modifications of the code, and is especially suitable to reduce the high (usually unmeasurable) uncertainty associated to the qualitative geological data used in more complex numerical tools aimed at modelling a lot of geological resources (groundwater, minerals, geothermal, petroleum) or different Earth phenomena.

1. Introduction

Numerical modelling is increasingly replacing the classic qualitative geological ways of data representation and mapping allowing for interactive 3D visualizations while provide a measure of the confidence of the mapped parameters and variables (Jessell, 2001; Wycisk et al.,

* Corresponding author. Departamento de Ciencias de la Tierra y del Medio Ambiente, Universidad de Alicante, Campus San Vicente, San Vicente del Raspeig, 03080, Alicante, Spain.
E-mail address: manuel.martin@ua.es (M. Martín-Martín).

https://doi.org/10.1016/j.marpetgeo.2023.106283
Received 14 March 2023; Received in revised form 18 April 2023; Accepted 20 April 2023
Available online 22 April 2023

2009; Ford et al., 2010; Rohmer et al., 2020). Currently, a wide number of modelling tools for 3D visualizations exist (GemPy, OSGeo), of which open source Python libraries (GeoPandas) and posts listing libraries (Albion, Gisgeography, Parpoil) devoted to Geographic Information System (GIS) and geological mapping (Semmo et al., 2015; Hobona et al., 2006; Vázquez-Suñé et al., 2016; Evangelidis et al., 2018; Miao et al., 2017; Husillos, 2022; Pyrcz, 2023) are of special interest in different applied geology fields, such as those intended to groundwater, minerals, geothermal and petroleum exploration.

The experience gained by the researchers responsible of this work with Python libraries for 3D visualization and geological data handling was essential to develop new applications devoted to data classification and 3D visualization of the stratigraphic architecture of sedimentary bodies, essentially porous media (Bullejos et al., 2022a, 2022b, 2023). In one of these, Bullejos et al. (2022b) used (i) the grain size of sediments as measurable physical parameter to make quantitative evaluations, (ii) the same regular grid to draw comparable sections and models over space, (iii) a machine learning k-Nearest Neighbors (KNN) algorithm to produce an interactive 3D model based on horizontal sections of the grain-size classes, and (iv) some Python libraries to produce interactive 3D models of the stratigraphic arrangement of the Quaternary Llobregat River Delta (LRD) near to Barcelona city in northeastern Spain. To design and run these models, the public grain-size database prepared by the Water Authority of Catalonia (*Agència Catalana de l'Aigua*) for groundwater purposes was utilized. The application of Bullejos et al. (2022b) included Jupyter notebooks describing the methodology and a version of the Python code, which is downloadable from the GitHub repository (https://github.com/dcabezas98/knn-stratigraphic-visualization). Moreover, Bullejos et al. (2023) performed an additional algorithm to evaluate the confidence of the KNN predictions as a variant of the Similarity Ratio algorithm for machine-learning KNN (Delany et al., 2005; Hu et al., 2009; Murphy and Redfern, 2015). The implemented metric assigned a confidence value based on weights inversely proportional to the distance of each grain-size class. The obtained average-weighted confidence of the 2D grain-size KNN-prediction horizontal maps was moderate to high attending to a logical decreasing data density with depth. The KNN-algorithm confidence proved its suitability in the supposedly well-depurated grain-size database used in these experiences.

The present work refines the above described KNN methodology — Python libraries, and KNN-prediction and KNN-prediction confidence algorithms — to create an interactive 3D HTML model

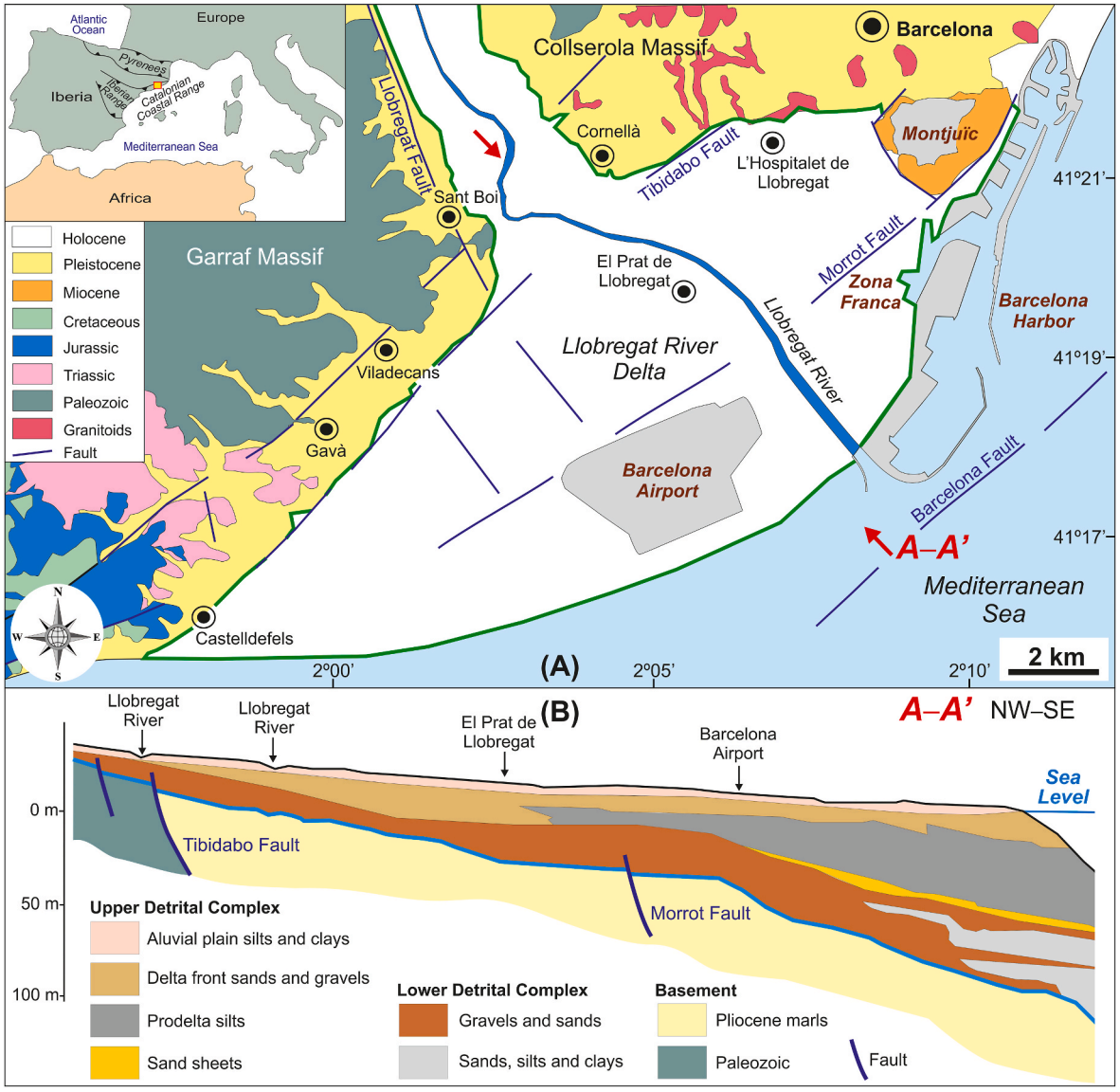


Fig. 1. (A) Simplified geological map of the onshore LRD area (green line) updated and simplified from Almera (1891), Medialdea and Solé-Sabarís (1973, 1991), and Alonso et al. (1974). (B) Simplified geological cross-section A–A' (with NW–SE orientation) of the onshore LRD; see location in (A).

(3D_Vertical_Sections_Faults_LRD.html) that combines 2D grain-size KNN-prediction vertical maps (vertical sections) from which syn-sedimentary faults and other features in sedimentary porous media can be delineated. The Quaternary LRD was again the sedimentary porous media selected to show the application. Aimed at producing comparable (and integrable) results to the horizontal 3D models developed previously by Bullejos et al. (2022a, 2022b, 2023), this paper uses the same grain-size database and grid-size for spatial mapping.

2. Geological framework for application

The LRD is a densely populated coastal plain of about 100 km^2, in the southeast sector of the metropolitan area of Barcelona city in the Catalonia region in northeastern Spain (Fig. 1A). The historical water abundance and its strategic location near to Barcelona city have favored the development of an important industrial activity. The degradation of groundwater quantity and quality due to high exploitation rates to supply to the increasing population and industrial activity produced negative consequences and leaded to many hydrogeological studies over the last seven decades (Solé-Sabarís, 1963; Marqués, 1984; Custodio, 1987; Font et al., 1987; Iribar et al., 1997; Abarca et al., 2006; Vázquez-Suñé et al., 2006; Postigo et al., 2021). Other modern development milestones occurred in the region, such as the Olympic Games in 1992 and the Llobregat Delta Infrastructure and Environment Plan (*Pla d'Infraestructures i Medi Ambient del Delta del Llobregat*, LRD Infrastructure Plan) started from the nineteens (Resolution 12956/1994) modified the LRD land use. As response to these damages and changes in the LRD, in 2004 the Water Authority of Catalonia created the Technical Unit of the Llobregat Aquifers (*Mesa Tècnica dels Aqüífers del Llobregat*, METALL) to assess the cumulative impact of civil works and other human activities on the groundwater resource (Official Statement, 2004). One of the first tasks of METALL was to compile and homogenize the huge geological and hydrogeological information generated in the area. This public geological and hydrogeological database, which is available on request, has been used in this paper.

The geological studies in the area started at the end of the XIX century by Almera (1891). In the XX century, several studies (Llopis, 1942; Solé-Sabarís, 1963; Medialdea and Solé-Sabarís, 1973, 1991; Alonso et al., 1974) allowed proposing geological maps and 2D cross-sections aimed to support hydrogeological studies. Sedimentological studies performed in the 1970s and 1980s (Medialdea and Solé-Sabarís, 1973; Serra and Verdaguer, 1983; Marqués, 1984) allowed clarifying the geology of the LRD. In the 1980s, the prodeltaic bodies of the emerged delta, dated as Holocene, were studied in detail (Manzano, 1986) and the geological characterization of the continental shelf with support of 2D marine reflection seismic geophysical surveys promoted by the Geological Survey of Spain (IGME, 1986, 1989). These studies, with a strong sedimentological component, allowed the sequential division of the LRD and the arrangement of the Quaternary materials. In the 1990s, this huge geological background was combined with the former geological information compiled from dozens of boreholes to make modern groundwater evaluations (Iribar et al., 1997). Contemporary with the LRD Infrastructure Plan development at the beginning of the XXI century, new geological data allowed fine research (Alcalá et al., 2003a, 2003b, 2003c; Parcerisa et al., 2008; Gámez et al., 2009; Salvany and Aguirre, 2022) aimed to detail the 2D stratigraphic architecture of the LRD. The main findings were (i) delineation of the Pliocene–Quaternary boundary, (ii) definition of aquifer levels, and (iii) identification of the aquifer levels affected by seawater intrusion and mobilization of pollutants (Vázquez-Suñé et al., 2006; Abarca et al., 2006; Postigo et al., 2021).

From a geological point of view, the LRD is regionally fed by the Llobregat River and its tributaries arriving from the Pre-Pyrenean Range and locally from the Llobregat River lower valley reliefs (Garraf and Collserola massifs) belonging to the Catalonian Coastal Range (Medialdea and Solé-Sabarís, 1973, 1991). This range is a NE–SW-oriented mountain chain that gives pass downward to the Mediterranean coast (Fig. 1A). The LRD is also bounded toward the NE by the Montjuïc relief. This area represents a Miocene-Quaternary rifted margin associated with the opening of the Valencia Trough. Nowadays, this opening in the Catalonian area is related to several well-known active faults (Perea et al., 2012; Sanz de Galdeano, 2019): the NE-SW Montseny and Pla de Barcelona Faults and the NW-SE Gulf de Roses and Amer Faults. This tectonics is similar to other Mediterranean opening as the case of the Tyrrhenian sea spreading in the Calabria area (Brutto et al., 2016; Tripodi et al., 2018; Mangano et al., 2022). In detail, the LRD area is affected by two main fault families detected in surface in the basement reliefs, and also in the LRD by reflection seismic geophysical surveys, coincident in orientation with the above mentioned active faults: the NE–SW oriented Tibidabo, Morrot, and Barcelona fault families and the NW–SE oriented Llobregat fault family, that conditioned the reliefs and the location of the Llobregat River outlet (Llopis, 1942; Solé-Sabarís, 1963; Medialdea and Solé-Sabarís, 1973, 1991; Alonso et al., 1974; Serra and Verdaguer, 1983; Marqués, 1984; Manzano, 1986; IGME, 1986, 1989; Iribar et al., 1997). From a hydrogeological point of view, Pliocene and older rocks are considered the basement in the LRD, which is separated from the Quaternary formations by an important unconformity surface (Alcalá et al., 2003a, 2003b, 2003c; Simó et al., 2005; Gámez et al., 2009) (Fig. 1B). The Pliocene basement is made of estuarine marls, silts, and clays (Simó et al., 2005; Gámez et al., 2009). The Quaternary record was divided into two depositional sequences Pleistocene and Holocene in age (IGME, 1986, 1989; Alcalá et al., 2003a, 2003b, 2003c; Simó et al., 2005; Gámez et al., 2009). The terms Lower Detrital Complex (LDC) and Upper Detrital Complex (UDC) have also been adopted in the scientific literature (Gámez et al., 2009) for the same depositional sequences (Fig. 1B). According to reflection seismic geophysical surveys performed in the offshore delta (IGME, 1986, 1989), the LDC can be divided into three parasequences (Alonso et al., 1974). In general terms, the LDC is made of conglomerate bodies (locally with sand) with intercalated silt- and clay-rich intervals (Serra and Verdaguer, 1983; Marqués, 1984; Manzano, 1986; IGME, 1986, 1989). Four main paleochannels of gravels were defined in the LDC (Simó et al., 2005; Gámez et al., 2009): one at the SE end of the LRD in the Castelldefels area, two coincident with the current Llobregat River course, and the last in the Zona Franca area at the NE end of the LRD. The UDC, from bottom to top, is made of a sand layer, a silt bed, a gravel bed (locally with sand), and upper silt and clay cover forming the current alluvial plain and the associated coastal wetlands and marshes (Manzano, 1986; IGME, 1986, 1989). The LRD coastal plain is also modeled by different streams coming from the neighbor reliefs (Alcalá et al., 2003a, 2003b, 2003c; Simó et al., 2005; Gámez et al., 2009) and the regional littoral drift which distributes the shoreline sedimentation towards the SW (Font et al., 1987; Chiocci et al., 1997; Zecchin et al., 2011).

3. Methodology

3.1. Assimilation of our previous methodology

This section describes succinctly the methodology implemented by Bullejos et al. (2022a, 2022b, 2023) to create interactive 3D HTML models based on 2D horizontal sections in the Quaternary onshore LRD. This paper refines this methodology, which consist in the same or lightly depurated (i) grain-size database prepared by the METALL (Official Statement, 2004), (ii) regular grid for comparable spatial mapping, (iii) Python libraries for data clasification, interpolation and rendering prior 3D visualization of models, (iv) KNN algorithm for data classification and prediction, and (v) KNN-algorithm for prediction confidence. The rationale is to produce comparable (and integrable) results to the obtained by Bullejos et al. (2022a, 2022b, 2023), but in this case adapted to create an interactive 3D HTML model from the mapped KNN-predictions (vertical sections) and their confidences (vertical

confidence sections).

The grain-size database was additionally revised to detect possible outliers unnoticed in previous works. This database included 433 boreholes in the LRD with their location (coordinates x and y), prospecting depth (coordinate z), basement position, and Quaternary infilling grain-size record (Alcalá et al., 2020) in an XLS (Excel) file with meter-by-meter georeferenced values. These data were classified into four grain-size classes attending to the relevance of their lithological equivalences for groundwater exploration purposes as in Bullejos et al. (2022a): (i) pre-Quaternary basement, (ii) Quaternary clay–silt (fine size) sediments (<1 mm), (iii) Quaternary coarse sand (medium size) sediments (1–5 mm), and (iv) Quaternary gravel (coarse size) sediments (>5 mm).

Python (https://www.python.org) is a widely-used friendly programming language with many open-source libraries (packages and modules) (Yukun et al., 2019). Several of these libraries were employed by Bullejos et al. (2022a, 2022b, 2023) to create 3D HTML models from horizontal sections of the mapped four defined grain-size classes. This work uses these same libraries and involved procedures. The details can be consulted in Bullejos et al. (2022a, 2022b, 2023).

The KNN-algorithm is a robust machine learning classifier (Gou et al., 2019; Pratama, 2018; Wang et al., 2018; Huang et al., 2020), which takes the proximity and similarity of the data into account to make predictions on the objective variable. This operation assumed that the inferred grain-size class at a given point is mostly dependent on the nearest data. Therefore, the KNN-algorithm is used to estimate data from a set of measurements (input dataset) in a wider space of individuals. The KNN-algorithm was used to obtain the spatial distribution of the four defined grain-size classes in the nodal positions of a 300 m × 300 m regular grid. The results were horizonal sections and 3D models. The details can be consulted in Bullejos et al. (2022b, 2023).

A measure of the confidence is needed to ensure the KNN-prediction reliability. Inspired by the Similarity Ratio algorithm described in the machine-learning KNN literature (Delany et al., 2005; Hu et al., 2009), a metric to quantify the confidence of the KNN-predictions was introduced. This metric assigned a confidence value based on weights inversely proportional to the distance between each grain-size class. The higher difference between the highest weight (the one assigned to the neighbor whose class is taken) and the weights of the nearest neighbors belonging to the other classes (all but the inferred class), the higher the confidence value. The basement confidence mapping was a particular case since below its top surface, no other grain-size classes associated to Quaternary materials can be found. The basement class deserved special treatment and the confidence formulation was modified to favor the basement when several basement data close to each other and apart from other non-basement data points existed. The details can be consulted in Bullejos et al. (2023).

3.2. Construction of 2D vertical grain-size sections

A main part of this work is to create vertical sections by using the above described KNN methodology and the four defined grain-size classes. As defined above, grain-size classes are equivalent to generic lithologies from which to deduce the essential stratigraphic elements defining the internal architecture of the Quaternary LRD, i.e., the basement top surface, and lithological distributions and boundaries.

For the construction of each vertical section, the KNN-algorithm must examine the K-nearest classes and select the most frequent one. The routine assigns a weight inversely proportional to the distance between each neighbor in order to favor the closest ones. The choice of a suitable value for the parameter K was a crucial issue. Smaller values of K may generate unrealistic polygonal regions, whereas larger values favor more realistic, smother regions although they may disregard isolated data. Thus, different values of K (1, 2, 3, 5, 7, 15, 25, and 75) were used to create the vertical sections. The geological logic was applied to conduct this process under feasible realizations, hence avoiding problems with edge/boundary effects typically viewed in KNN applications.

3.3. Construction of 3D models as HTML files

Different Python libraries are used to create an interactive 3D HTML model, such as pandas, numpy, plotly, matplotlib, shapely.geometry, utm, and sklearn.neighbors. A detailed description can be found in Bullejos et al. (2022a, 2022b). In the Python development environment, the selected figure to be the 3D model is defined and saved as an HTML file (Leifeld, 2013; Li et al., 2019; Gur et al., 2022). The main advantage of displaying the Python model as a HTML file is that users only need a conventional web browser to visualize it. In the 3D model, three created features stand out: the modelling contour, the vertical sections, and the faults. The Llobregat Delta contour has been created using updated satellite imagery from Google Earth. First, a KML file with the latitudes and longitudes of the points that determine the modelling contour is created. Then, the KML file with the UTM coordinates of those points is converted into a CSV file. Finally, the matplotlib scatter3d command is used to add the contour to the figure that will determine the model. The vertical sections are drawn from point clouds using the KNN prediction algorithm. The point clouds are obtained from the grain-size coordinates included in the borehole database (Alcalá et al., 2020). The process to produce these predictions is detailed in Bullejos et al. (2022a, 2022b, 2023). A specific vertical section will be defined by the segment delimited by two end-points that the user must fix. The obtained KNN-predictions from the data points computed near to this segment are selected. Once the data in the point clouds are selected, the matplotlib Scatter3d command is employed to add the vertical sections to the 3D HTML model.

After the vertical sections are obtained, the geological logic is required to decide which points in those sections are attributable to faults. A general criterion is when a same drastic horizontal interruption (displacement or vertical step) in the basement and/or Quaternary infilling continuity appears in the successive aligned vertical sections. This criterion can be corroborated when the subsurface interruption is correlatable with the trace of faults observed in surface or by means of geophysical surveys. Each fault determined by a number of points will be added to the model by using the matplotlib command Mesh3d. This command allows to draw geometric figures such as faults, for instance, providing the two end-points for a linear fault or a set of points for more complex fault geometries. The user can view different perspectives of the model, hide some elements to focus on others, zoom in or out a particular area, and take a picture from a specific perspective. The model can also be rotated around to provide a more precise view of it. Moreover, the proper positioning of the vertical geological cross-sections in the 3D HTML model serves to define the correct positioning of the deduced syn-sedimentary faults.

4. Results

4.1. KNN-predictions: vertical grain-size sections

The location of the 433 boreholes with grain-size data in the LRD is shown in Fig. 2A. An HTML figure with the borehole distribution in the LRD is supplied as Supplementary Material (3D_Boreholes_LRD.html). Ten vertical sections are created following intentionally the main density of boreholes and the orientation of the NW–SE (transversal) and SW–NE (longitudinal) faults families identified in surface (Fig. 1A). The seven transversal vertical sections are labeled A–B, C–D, E–F, G–H, I–J, K–L, and M–N. The three longitudinal vertical sections are labeled O–P, Q–R, and S–T. The trace of each vertical section is shown in Fig. 2B.

The same vertical:horizontal ratio of 1:50 was used to plot all the vertical sections. The reason was to deduce better vertical geological details (Fig. 3). Due to the progressive decreasing of the borehole number (data density) with depth, 160 m below sea level (m b.s.l.) was

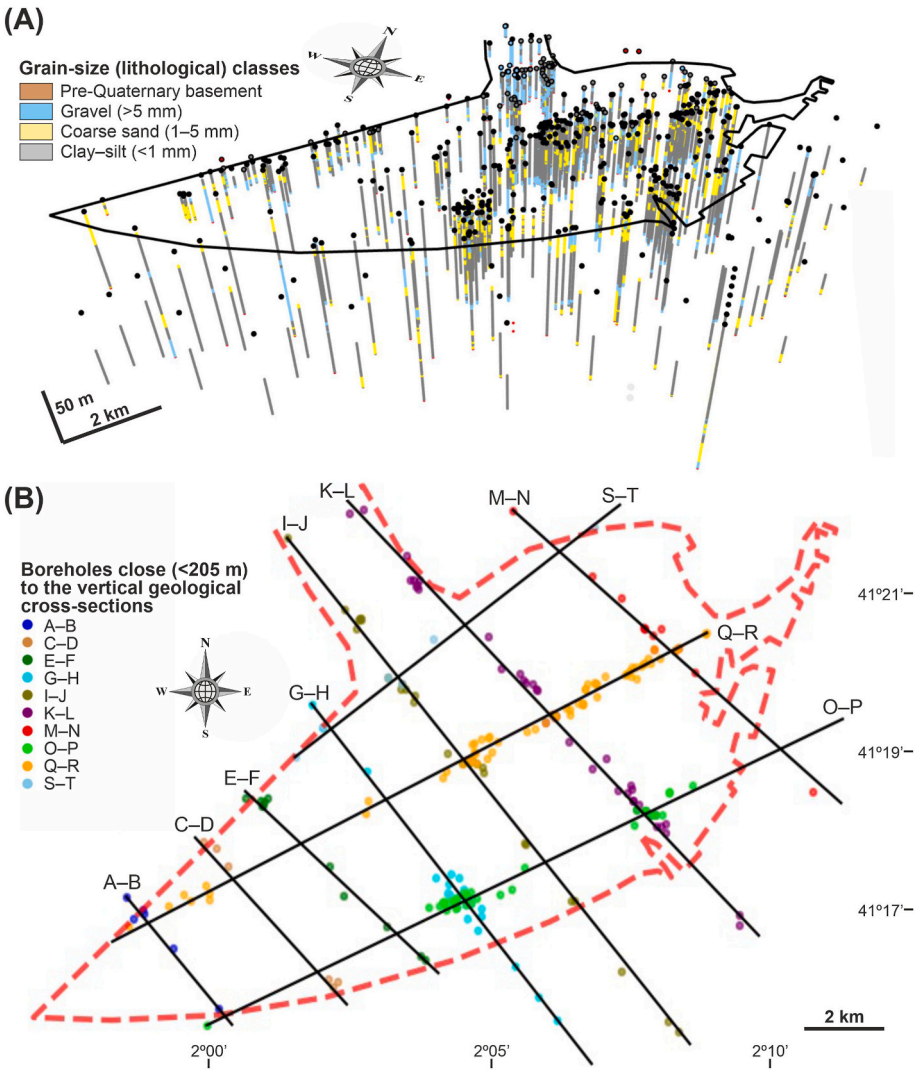


Fig. 2. (A) 3D distribution of the grain-size classes along the axis *z* in each of the compiled boreholes in the LRD (Alcalá et al., 2020). For a better display, the adopted plotting coordinates are $x = 2$, $y = 2$, and $z = 0.5$. An interactive 3D HTML version of this model is provided as supplementary material (3D_Boreholes_LRD.html). (B) 2D sketch map showing the positioning of the vertical geological cross-sections and the closest boreholes involved in each section.

the maximum depth adopted for KNN-predictions. This maximum depth was not decided arbitrarily, as it was based on the decreasing confidence of the KNN-predictions of the horizontal sections found by Bullejos et al. (2023). Below 160-m depth, the KNN-prediction confidences became low to very low. In each vertical section, the four defined grain-size classes appear identified with different colors: pre-Quaternary basement in brown, and the clastic (Wentworth, 1922) Quaternary infilling as gravel (coarse size) in cyan, sand (medium size) in yellow, and clay–silt (fine size) in grey. The boreholes located less than 250 m from the trace of each the vertical section have been drawn with high opacity, while the boreholes located in the range of 250–1000 m have been faintly drawn (Fig. 2B).

In an initial phase, different values of K (1, 2, 5, 7, 15, 25, and 75) were used to create the of vertical sections. The transversal G–H and longitudinal S–T vertical sections are two examples selected in Fig. 3). In both cases, only the KNN predictions for $K = 1$, $K = 5$, $K = 25$, and $K = 75$ are displayed. All the vertical sections (KNN-prediction maps) obtained for the entire range of trialed values of K (1, 2, 5, 7, 15, 25, and 75) are included as Supplementary Material.

As expected, $K = 1$ provides more realistic KNN-predictions of the Quaternary infilling since the KNN-algorithm gives more importance for the nearest neighbor to every point in the grid. The progressive increasing of K means increasing searching distances, thus increasing the number of input data integrated in the grain-size outputs. The resulting smoothing effect is typical to many other interpolation techniques and, despite that the KNN-predictions capture the essential features of the mapped variable (grain-size equivalent to lithologies) and produce easier to interpret results, it also glimpses less confident results since omits important details at short distance. Thus, the results confirm a positive relationship between increasing K values and KNN-prediction smoothing degree, and glimpse an expected negative relationship between increasing K values and KNN-prediction confidences, as described in next Section 4.2.

Regarding the direct geological implications of using lower and higher K values, the pre-Quaternary basement appears and disappears in depth for lower K values (e.g., $K = 1$ and $K = 5$ in Fig. 3), which is geologically impossible since once it is reached it should continuous in depth. This condition is the result of using a lot of data points with very different grain-size magnitudes at short distance, and then applying the same weight to the pre-Quaternary basement and Quaternary grain-size classes. As described above, the smoothing of the KNN-predictions for higher K values (e.g., $K = 25$ and $K = 75$ in Fig. 3) is nevertheless of assistance to capture the shape of the basement top surface, which is an essential stratigraphic element to be mapped. In this case, K values around 25 seem to be optimal to capture the general and smoothy shape of the basement surface with a deepening structure toward the sea

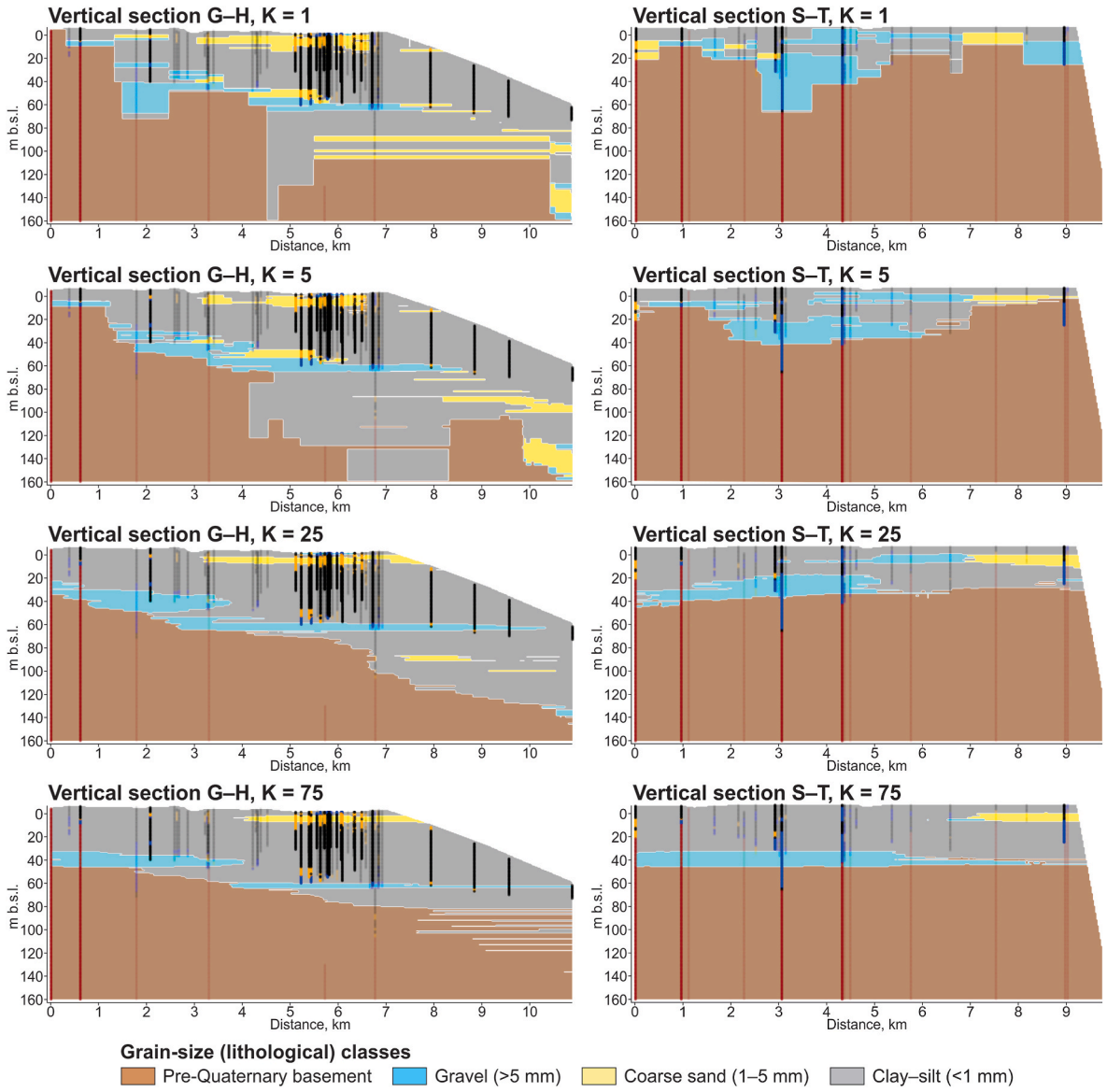


Fig. 3. KNN predictions of the grain-size classes for $K = 1$, $K = 5$, $K = 25$, and $K = 75$. The plots represent the vertical geological cross-sections G–H (left column plots) and S–T (right column plots); see location of sections in Fig. 2B.

demonstrated by reflection seismic geophysical surveys (IGME, 1986, 1989) (Fig. 3).

In what concerns to the Quaternary grain-size classes, the small bodies of minority classes (gravel, sand, or even the basement in some cases) are not reflected and substituted by the majority clay–silt class for higher K values (e.g., $K = 25$ and $K = 75$ in Fig. 3). However, the lower K values (e.g., $K = 1$ and $K = 5$ in Fig. 3) maintain well the stratigraphic heterogeneity demonstrated in detailed stratigraphic studies (Simó et al., 2005; Gámez et al., 2009; Parcerisa et al., 2008) (Fig. 3). However, every grain-size data point is wanted to be considered in order to represent the heterogeneity at short distance. The value $K = 1$ is the only one that can determine this heterogeneity. As can be seen at approximately the third kilometer of the vertical section S–T at depth 50–60 m b.s.l. (Fig. 3), the KNN-predictions for $K = 25$ and $K = 75$ ignore a set of gravel levels reported in the borehole lithological records. The same happens at km 9 and 20 m b.s.l. in this section. Hereafter, only the KNN predictions for $K = 1$ are considered to construct the 3D model.

4.2. *Confidence of the KNN-predictions*

As displayed in Fig. 2A, the borehole distribution and the subsequent grain-size data density vary over space and depth in the LRD. At a given depth, the variable data density lead to a heterogeneous spatial distribution of the KNN-prediction confidences, as was reported by Bullejos et al. (2023). This particular heterogeneity at a given depth is additional to the general decreasing trend of the KNN-prediction confidence with depth, which is due to the progressive decreasing of the borehole number with depth, as was also reported by Bullejos et al. (2023).

The spatial distribution of the KNN-prediction confidence determines the reliability of the vertical sections described in Section 4.1, and therefore the deductions that can be inferred from them. For a proper spatial integration of the KNN-predictions, confidence maps of those same vertical sections are created to identify possible divergences that can spoil the definition of the essential stratigraphic elements, including faults. The confidence maps creation followed the procedures described by Bullejos et al. (2023); see also Delany et al. (2005) and Hu et al. (2009) for further details about the KNN-confidence algorithms. Thereby, a KNN-prediction confidence map was created for each

KNN-prediction map, i.e., for each vertical section. After demonstrate that $K = 1$ is a necessary optimal value to represent the heterogeneity of the different Quaternary grain-size classes at short distance, Fig. 4 shows the KNN-prediction confidence maps for the vertical sections G–H and S–T. The KNN-prediction confidence is expressed as percentage from 25% (minimum value, purple color) to 100% (maximum value, red color). Several intermediate confidence levels as low (25–35%, dark bluish colors), satisfactory (35–50%, bluish-greenish colors), high (50–75%, greenish-orangish colors), and very high (75–100%, reddish colors) exist (Fig. 4).

Regarding the direct geological implications, the higher KNN-prediction confidences coincide with the borehole locations, as expected (Fig. 4). Boreholes are vertical lines of grain-size data with 100% confidence, whereas the confidence degree between boreholes is a function of the distance among them, as has been programmed in the developed KNN-algorithm confidence. When the borehole record ends, the confidence below this depth decreases again to satisfactory (35–50%) or even low (25–35%) levels (Fig. 4). Of special interest is the pre-Quaternary basement confidence mapping, since once the basement top surface is reached, no other Quaternary grain-size classes can be found deeper. Therefore, the basement KNN-prediction confidence must be 100% once this class is firstly reached. However, the KNN-prediction confidence obtained with $K = 1$ is low to satisfactory at most, thus deteriorating the spatial basement confidence, as was also identified above for the KNN-predictions.In relation with the confidence evolution with depth and from inland to seaward, the shallowest tracts of the vertical sections over 50 m b.s.l. usually show high to very high confidences in the 60–90% range since they used the higher data densities (Fig. 4). In the vertical section G–H, the existence of boreholes reaching the basement at about 50 m b.s.l. determines high to very high confidences at the inland end and low confidences to seaward where the number of boreholes is clearly lower. Since the basement is usually deepest to seaward and the offshore LRD is characterized by the lowest data density, the confidence associated to the first basement appearance is typically low in this area, as shown in the vertical section G–H (Fig. 4). Even in this circumstance of low data density inferred by the few existing boreholes, the KNN-prediction confidence in the offshore LRD reaches satisfactory (greenish color) or high (orangish color) level for $K = 1$. The KNN-prediction confidence of the other vertical sections shows similar patterns. All these confidence maps are included as Supplementary Material.

4.3. Interactive 3D HTML model

The Python libraries described in Section 3.3 (pandas, numpy, plotly, matplotlib, shapely.geometry, utm, and sklearn.neighbors) were operated sequentially to create an interactive 3D model of the Quaternary grain-size (lithologies) classes in the LRD. This model is based on the serial ordination of the seven transversal vertical sections A–B, C–D, E–F, G–H, I–J, K–L and M–N, and the three longitudinal vertical sections O–P, Q–R and S–T, as shown in Fig. 2B. The 3D model has been saved as an HTML file so that users can visualize it simply with a conventional web browser. In this way, the HTML file facilities different views, zooming, rotating, screenshots and moving around, as well as hiding elements by toggling in the legend to focus on details. The model is provided in Supplementary Material (3D_Vertical_Sections_Faults_LRD.html).

An example proposed by the authors (the users can choose others) shows some screenshots of the interactive 3D HTML model (Fig. 5). Fig. 5A shows a zenithal view of the LRD with the transversal and longitudinal vertical sections. Other zenithal oblique views of the LRD show the transversal (Fig. 5B) and longitudinal (Fig. 5C) vertical sections, separately, while Fig. 5D shows a zenithal oblique view of the LRD with both the transversal and longitudinal vertical sections.

The 3D HTML model allows the fine correlation of the above described horizontal interruptions (displacements or vertical steps) in the pre-Quaternary basement continuity, and in the main gravel bodies in the Quaternary infilling (Fig. 5). It is thought that most of these displacements or vertical steps are linked to the action of Quaternary active faults. The displacements or vertical steps can be attributable to faults since they are identified in the serial vertical sections. The orientation of these faults in space correlates very well with the fault families identified in surface (Fig. 1) and with those fault families detected by reflection seismic geophysical surveys. They are assimilable to syn-sedimentary faults.

5. Discussion

5.1. Syn-sedimentary faults

The study area is framed into a rift zone affected by two main fault families, as corroborated by the fault structures detected in the emerged basement reliefs (Llopis, 1942; Solé-Sabarís, 1963; Medialdea and Solé-Sabarís, 1973, 1991; Alonso et al., 1974; Serra and Verdaguer, 1983; Marqués, 1984; Gámez et al., 2009) and by reflection seismic geophysical surveys in the LRD itself (IGME, 1986, 1989). The main orientations of these fault families are NE–SW (Tibidabo, Morrot, and Barcelona faults) and NW–SE (Llobregat Fault). These two main fault families are roughly perpendicular and can be considered a system. The perpendicular positioning of the two sets of vertical sections was intentional to detect these and other related faults.

As described in Section 4.1, K values around 25 are optimal to capture the general and smoothy shape of the basement top surface, whereas $K = 1$ is a necessary optimal value to represent the heterogeneity of the Quaternary grain-size (lithologies) classes and basement at short distance. In the KNN-prediction maps created with $K = 1$ can be recognized horizontal interruptions (displacements or vertical steps) in

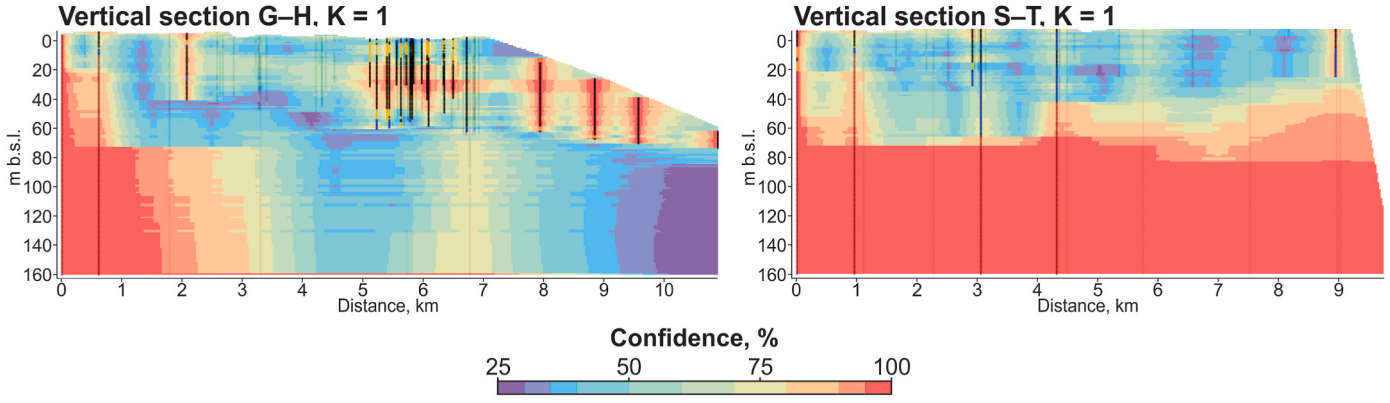


Fig. 4. KNN-prediction confidences of the grain-size classes for $K = 1$. The plots represent the vertical geological cross-sections G–H (left plot) and S–T (right plot); see location of sections in Fig. 2B.

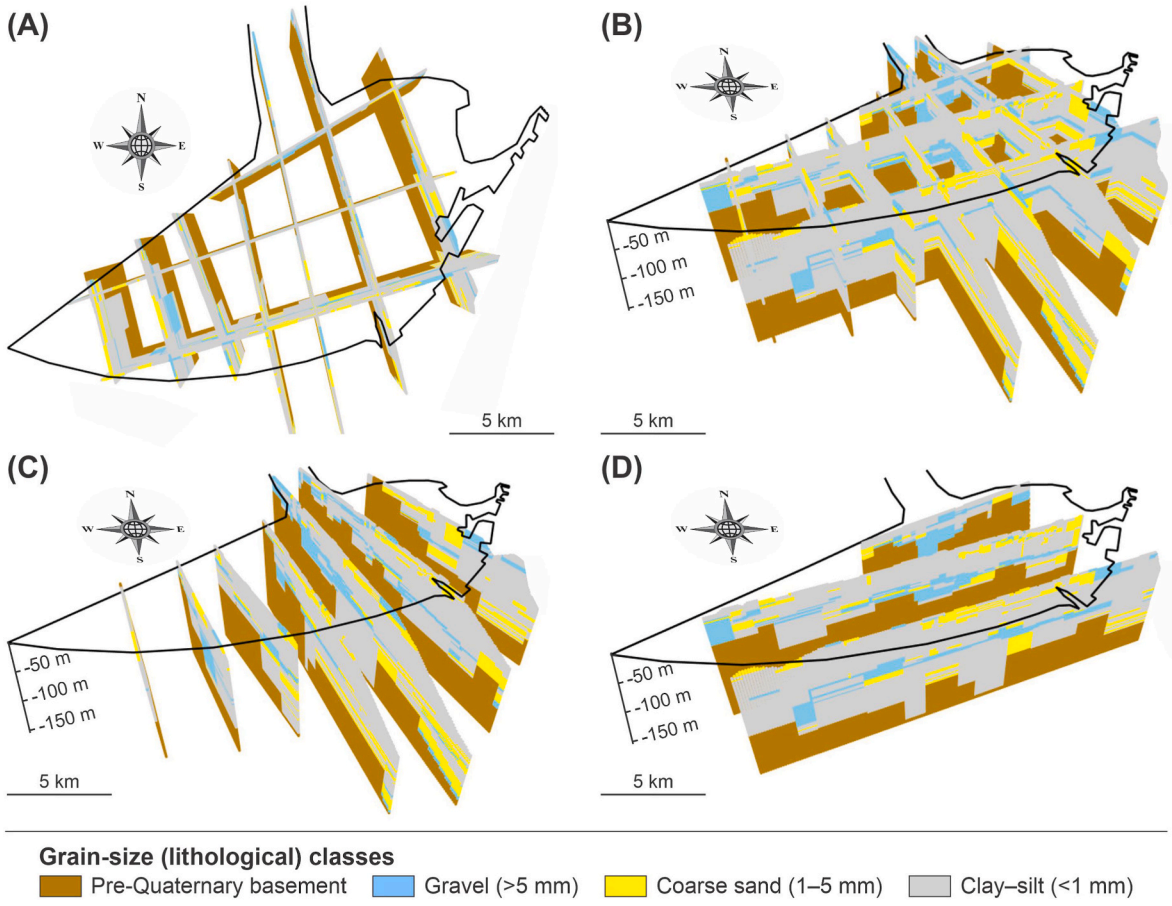


Fig. 5. Screenshots of the interactive 3D HTML model provided in Supplementary Materials (3D_Vertical_Sections_Faults_LRD.html). (A) Zenithal view of the LRD with the transversal (NW–SE) and longitudinal (SW–NE) vertical sections. (B) Zenithal oblique view of the LRD with the transversal and longitudinal vertical sections. (C) Zenithal oblique view of the LRD with the transversal vertical sections only. (D) Zenithal oblique view of the LRD with the longitudinal vertical sections only.

the basement continuity, but also in the coarse bodies (gravel and sand) in the Quaternary filling itself. In a number of cases, these displacements or steps in the basement produce a space infilled predominantly by coarse deposits like gravel and sand. The rationale used in this paper to explain this fact is that most of these displacements or steps in the pre-Quaternary basement and in the Quaternary infilling are linked to the action of Quaternary active faults. Moreover, since in much cases these displacements in the basement continuity include a filling dominated by coarse deposits, a tectonic control of coarse-clastic sedimentation can be proposed. Those displacements or steps whose lateral continuity enables its correlation in the serial vertical sections have been signaled and named in both the transversal (Fig. 6) and longitudinal (Fig. 7) vertical sections. Thus, Figs. 6 and 7 can be also used as transversal and longitudinal correlation panels.

Moreover, during the performing of the interactive 3D HTML model described in Section 3.3, most of these faults were tentatively labeled accordingly with their alienation to the fault families well recognized in surface (Tibidabo, Morrot, and Llobregat faults). In other cases, fault families detected by reflection seismic geophysical surveys were also found (Barcelona Fault). Also, other faults were tentatively proposed for the first time (Castelldefels, Viladecans, Airport, Zona Franca, and Harbor faults). Even, several simple faults detected in surface, or detected here by the first time, have been revealed as fault zones made of several fault branches with different displacements or steps (Tibidabo, Morrot, Llobregat, and El Prat faults). In detail, Tibidabo, Morrot and Barcelona faults are synthetic with the Valencia Trough opening and show sunken blocks seaward. Therefore, these faults could be related to the Montseny and Pla de Barcelona Active Faults (Perea et al., 2012; Sanz de Galdeano, 2019). On the contrary, the Airport and Harbor faults are antithetic to the previous ones showing the sunken blocks landward. On another hand, Castelldefels, Viladecans, Llobregat, and Zona Franca Faults show sunken blocks northeastward in a similar way to the Amer Active Fault (Perea et al., 2012; Sanz de Galdeano, 2019), while El Prat Fault has a counter regime to the formers as the Gulf de Roses Active Fault (Perea et al., 2012; Sanz de Galdeano, 2019). The fault actuation seems to be more evident in the Pleistocene LDC causing displacements in the Pre-Quaternary basement and in the Quaternary infilling. These faults seem to be inactive or much less active in the Holocene UDC, where displacements are much less evident (Figs. 6 and 7).

5.2. *Syn-tectonics deposition of gravel bodies*

Along with his interest for groundwater exploration purposes, the definition of the depositional typologies of the gravel bodies provides excellent criteria about the syn-tectonic sedimentary processes occurred in the study area. In addition to the fault families recognized and deduced here by the first time, Figs. 6 and 7 show other syn-tectonic sedimentary features interpreted as gravel channels induced by fault actuation, progradation of channel lobes, and lateral migration of channel bars of mappable entity. Asymmetric gravel channels associated to faults actuation are usually attached to a fault showing a higher gravel thickness in the closer part to the fault that decreases as it moves away from the fault. Good examples are signaled in the transversal C–D and E–F (Fig. 6) and longitudinal S–T, Q–R and O–P (Fig. 7) vertical sections. In the longitudinal vertical sections (Fig. 7) at least three main gravels channels are recognized, in agreement with findings reported by Simó

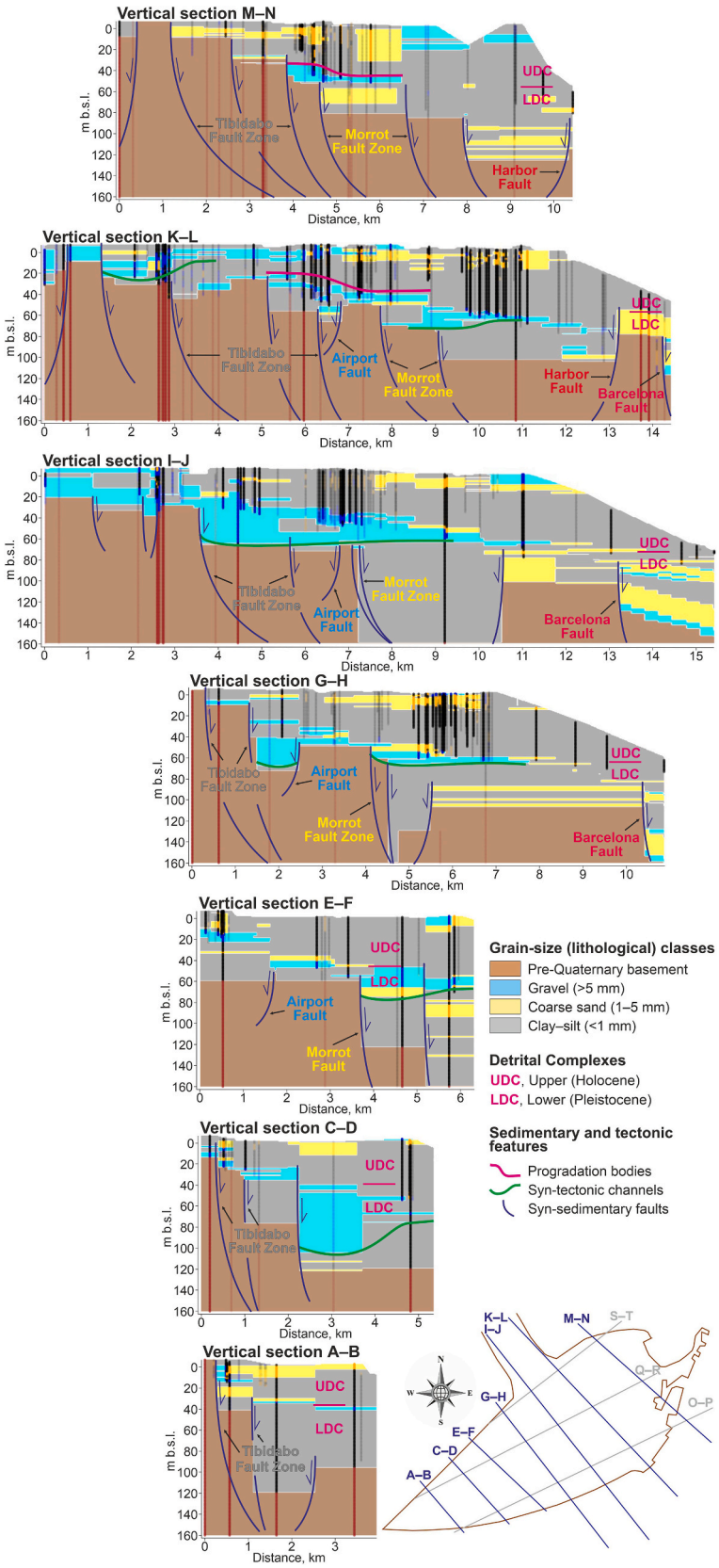


Fig. 6. Transversal (NW–SE) vertical sections A–B to M–N with the grain-size (lithological) classes, showing the location of correlated syn-sedimentary faults, and progradation of syn-tectonic gravel lobes and channels. The boundary of the Lower and Upper Detrital Complexes and the location of the boreholes (highlighted accordingly to their proximity) are also indicated.

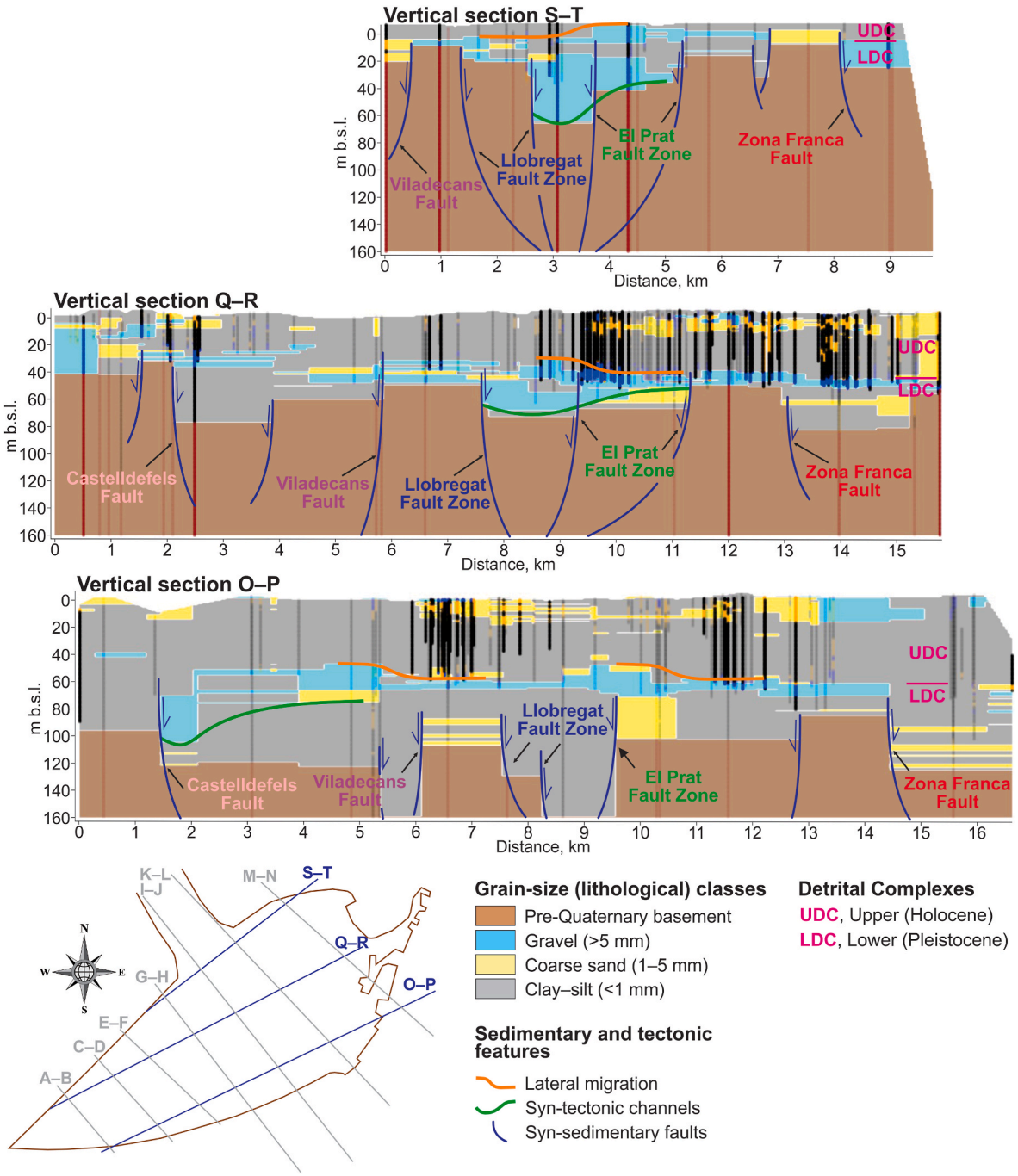


Fig. 7. Longitudinal (SW–NE) vertical sections A–B to M–N with the grain-size (lithological) classes, showing the location of correlated syn-sedimentary faults, and lateral migration of syn-tectonic gravel bodies and channels. The boundary of the Lower and Upper Detrital Complexes and the location of the boreholes (highlighted accordingly to their proximity) are also indicated.

et al. (2005) and Gámez et al. (2009), who proposed four paleo-channels in the LDC. The general evolution from bottom to top of these gravels channel deposits is transgressive. A GIF model (GIF Channel deposits transgression) is provided in Supplementary Material.

In the transversal vertical sections (Fig. 6) the progradation of channel lobes is also recognized, which also appears as asymmetrical bodies. They are signposted in the vertical sections M–N and K–L, where the thickness of the gravel bodies, which probably coincide with paleo-river channels, decreases to seaward. On the contrary, in the longitudinal vertical sections (Fig. 7) the recognized gravel bodies show transversal migration to the paleo-river channels flow and probably they are linked to lateral channel bars migration. Good examples of the lateral migration of gravel channels appear in the vertical sections S–T (left oriented) and Q–R and O–P (right oriented).

5.3. Interactive 3D HTML model

The KNN-prediction confidence of the Quaternary grain-size (lithologies) classes decreases with depth and from inland to seaward according to the decreasing data densities in these ways. The first appearance determines high confidences for the pre-Quaternary basement. The confidence in deep areas is low if the basement is not reached or there are many data points from the Quaternary clay–silt class, which is the majority class. An interactive 3D HTML model

(3D_Vertical_Sections_Faults_LRD.html) is provided in Supplementary Material.

However, the process of identification, naming, and correlation of syn-sedimentary faults would have been impossible without a 3D visualization after correct positioning of the performed vertical sections. As an example, Fig. 8 shows some screenshots of the created 3D model. In the vertical sections shown in Fig. 8, the Quaternary grain-size classes (gravel, sand, and clay–silt) have been hidden, and only the pre-Quaternary basement and the deduced syn-sedimentary faults have kept visible.

Fig. 8A shows two zenithal views of the Pre-Quaternary basement, showing the transversal vertical sections A–B to M–N and the location of the deduced and correlated longitudinal Tibidabo, Airport, Morrot, Puerto, and Barcelona faults. In turn, Fig. 8B shows other two zenithal views of the Pre-Quaternary basement, showing the longitudinal sections O–P to S–T and the location of the deduced and correlated transversal Castelldefels, Viladecans, Llobregat, El Prat, and Zona Franca faults. These zenithal views enable also proposing tentatively the lateral evolution of some of the deduced faults. In a few cases one fault passes from one unique trace to several branches, thus constituting a fault zone and revealing an echelon structure. This is the case of the Tibidabo Fault (Fig. 8A), which passes from two traces in the SW part of the LRD to one in the Viladecans area, and to three branches in the NE one. Similar features are observed for the Llobregat and El Prat faults, which have one single trace seaward and two and three branches landward, respectively (Fig. 8B). The Morrot Fault shows a single trace in the SW and NE ends of the LRD but two branches in the central part (Fig. 8A).

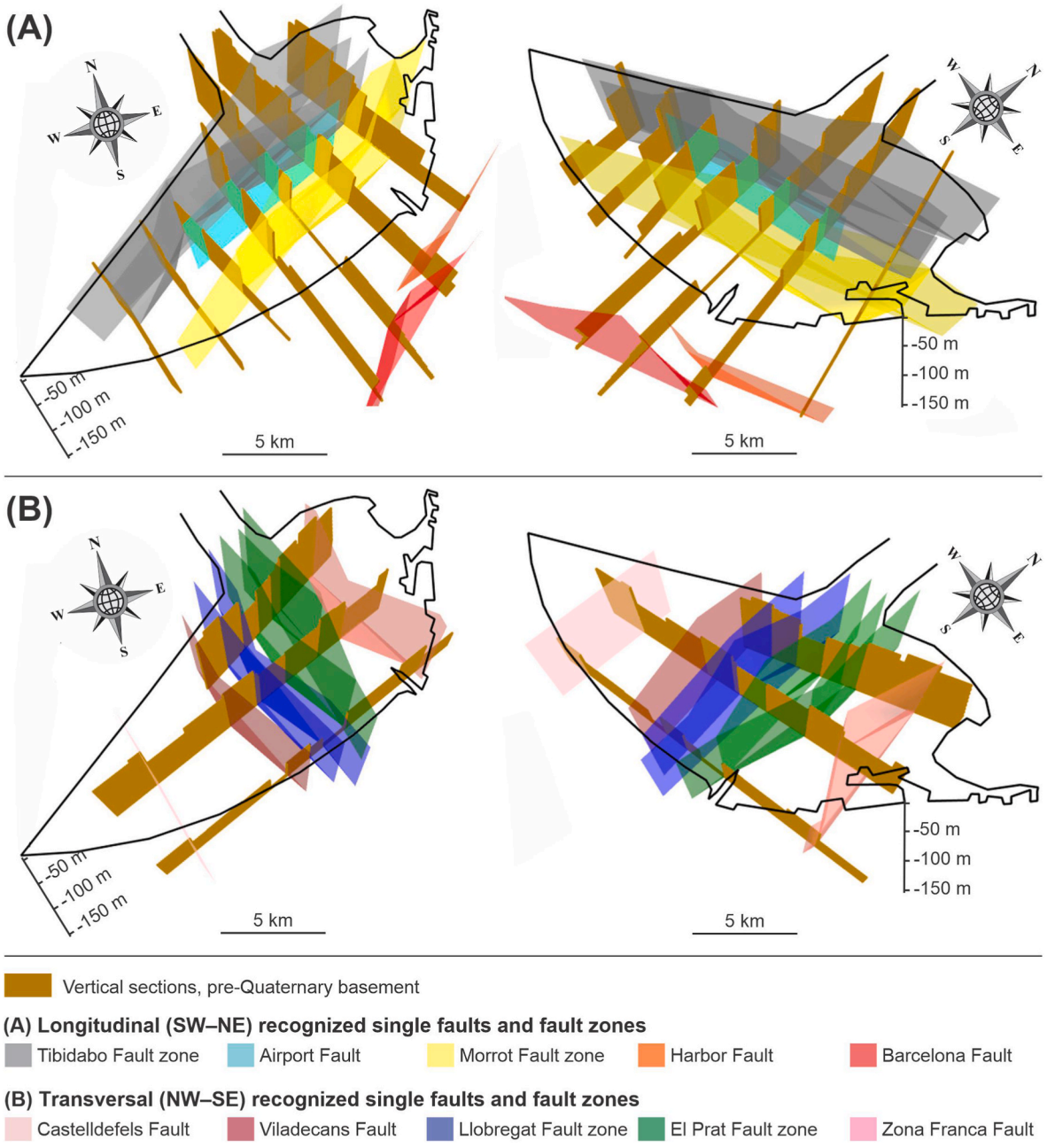


Fig. 8. Screenshots of the interactive 3D HTML model provided in Supplementary Materials (3D_Vertical_Sections_Faults_LRD.html). (A) Two zenithal views of the Pre-Quaternary basement from the transversal (NW–SE) vertical sections A–B to M–N, with location of the deduced and correlated longitudinal Tibidabo, Airport, Morrot, Harbor, and Barcelona faults. (B) Two zenithal views of the Pre-Quaternary basement from the longitudinal (SW–NE) vertical sections O–P to S–T, with location of the deduced and correlated transversal Castelldefels, Viladecans, Llobregat, El Prat, and Zona Franca faults.

6. Conclusion

This paper introduces a new KNN methodology to create grain-size vertical sections from which to delineate syn-sedimentary faults in sedimentary porous media. Four lithological (including grain-size) classes are defined. This enables to deduce essential stratigraphic elements, faults and other stratigraphic features.

Seven transversal and three longitudinal vertical sections were created, whose orientations followed intentionally the fault families identified in surface and/or detected by reflection seismic geophysical surveys.

For exploratory K values in the 1–75 range were used. K around 25 was optimal to capture the general and smoothy shape of the basement top surface, whereas $K = 1$ was a necessary optimal value to represent the heterogeneity of the Quaternary grain-size classes and Pre-Quaternary basement at short distance.

In general, the KNN-prediction confidence tends to decrease with depth and from inland to seaward according to the decreasing data density in these ways. The KNN-prediction maps created with $K = 1$ show horizontal interruptions (displacements or vertical steps) in the basement continuity and in the coarse bodies (gravel and sand) in the Quaternary filling itself attributable to the action of Quaternary active faults.

Several faults detected in surface or in this paper by the first time, have been revealed as fault zones made of several fault branches with different steps in an echelon-like arrangement. These faults have been linked or correlated with well-known active faults in the area in relation, in much cases, with the Valencia Trough opening. The faulting action seems to be more evident in the Pleistocene LDC causing displacements in the Pre-Quaternary basement and Quaternary infilling, and inactive or much less active in the Holocene UDC, where displacements are much less evident.

Other syn-tectonic sedimentary features as gravel channels faulty controlled, progradation of gravel lobes, and lateral migration of channel bars have also been identified.

Social and environmental interests underlying the groundwater resource and the consequent high density of geological data produced over the last seven decades motivated the LRD choosing to implement the new KNN methodology introduced in this paper. However, this methodology may have similar or even more interest in other applied geology fields, such as mineral, geothermal and petroleum exploration, or for other environmental, archeological, and civil engineering aims. The introduced methodology has used the grain-size physical parameter, which is easily measurable, constant (or very little variable) over instrumental time, and handleable mathematically. Other parameters and variables with a similar physical behavior can be also used, such as seismic, geotechnical, geoelectrical or mechanical ones. At its current development stage, this methodology could also be applied in other geological environments, preferably sedimentary porous media, making the due minor modifications of the code.

This paper proves that the combined use of KNN-predictions and KNN-prediction confidences is suitable to create interactive 3D HTML geological models in areas with a high-density of borehole data, such as heavily exploited Quaternary detrital aquifers, prospecting zones, or urban areas. Finally, this KNN methodology could specially be indicated as numerical geological support to reduce the typical high (usually unmeasurable) uncertainty associated to the use of qualitative geological data in other more complex numerical tools aimed at modelling a lot of geological resources or Earth phenomena.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Research supported by the Spanish Ministry of Education and Science research project PID2020-114381 GB-I00, and the Research Groups and projects of the Generalitat Valenciana from Alicante University (CTMA-IGA). Revision performed by Prof. S. Critelli is acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpetgeo.2023.106283.

References

Abarca, E., Vázquez-Suñé, E., Carrera, J., Capino, B., Gámez, D., Batlle, F., 2006. Optimal design of measures to correct seawater intrusion. Water Resour. Res. 42, W09415 https://doi.org/10.1029/2005WR004524.

Alcalá, F.J., Miró, J., García-Ruz, A., 2003a. Sobre la intrusión marina en el sector oriental del acuífero profundo del delta del Llobregat (Barcelona, España). Breve descripción histórica y evolución actual. Boletín Real Soc. Española Hist. Nat. 98, 41–50.

Alcalá, F.J., Miró, J., Rodríguez, P., Rojas-Martín, I., Martín-Martín, M., 2003b. Actualización geológica del delta del Llobregat (Barcelona, España). Implicaciones geológicas e hidrogeológicas. In: A., J., de la Orden, J.A., Gómez, J.D., Ramos, G., Mejías, M., Rodríguez, L. (Eds.), Tecnología de la Intrusión de Agua de Mar en Acuíferos Costeros: Países Mediterráneos; López-Geta, vol. 1. Geological Survey of Spain, Madrid, Spain, pp. 45–52.

Alcalá, F.J., Miró, J., Rodríguez, P., Rojas-Martín, I., Martín-Martín, M., 2003c. Características estructurales y estratigráficas del substrato Plioceno del Delta de Llobregat (Barcelona, España). Aplicación a los estudios hidrogeológicos. Geotemas 5, 23–26.

Alcalá, F.J., Martín-Martín, M., García-Ruz, A., 2020. A Lithology Database from Historical 457 Boreholes in the Llobregat River Delta Aquifers in Northeastern Spain. Figshare Dataset. https://doi.org/10.6084/m9.figshare.11865495.v1.

Almera, J., 1891. Mapa Geológico y Topográfico De La Provincia De Barcelona: Región Primera o De Contornos de la Capital Detallada, Scale 1:40,000, Memory and Maps. Diputación de Barcelona, Barcelona. Available online: https://cartotecadigital.icgc.cat/digital/collection/catalunya/id/2174. (Accessed 18 April 2022). accessed on.

Alonso, F., Peón, A., Rosell, J., Arrufat, J., Obrador, A., 1974. Geological Map of Spain, Scale 1:50,000, Sheet Nº 421; Barcelona, Memory and Maps. Geological Survey of Spain, Madrid, Spain. Available online: http://info.igme.es/cartografiadigital/geologica/Magna50Hoja.aspx?language=es&id=421. (Accessed 18 April 2022). accessed on.

Brutto, F., Muto, F., Loreto, M.F., De Paola, N., Tripodi, V., Critelli, S., Facchin, L., 2016. The neogene-quaternary geodynamic evolution of the central calabrian arc: a case study from the western catanzaro trough basin. J. Geodyn. 102, 95–114.

Bullejos, M., Cabezas, D., Martín-Martín, M., Alcalá, F.J., 2022a. A Python application for visualizing the 3D stratigraphic architecture of the onshore Llobregat River Delta in NE Spain. Water 14, 1882. https://www.mdpi.com/2073-4441/14/12/1882.

Bullejos, M., Cabezas, D., Martín-Martín, M., Alcalá, F.J., 2022b. A K-nearest neighbors algorithm in Python for visualizing the 3D stratigraphic architecture of the Llobregat River Delta in NE Spain. J. Mar. Sci. Eng. 10, 986. https://doi.org/10.3390/jmse10070986.

Bullejos, M., Cabezas, D., Martín-Martín, M., Alcalá, F.J., 2023. Confidence of a k-nearest neighbors Python algorithm for the 3D visualization of sedimentary porous media. J. Mar. Sci. Eng. 11, 60. https://doi.org/10.3390/jmse11010060.

Chiocci, F.L., Ercilla, G., Torres, J., 1997. Stratal architecture of Western Mediterranean Margins as the result of the stacking of Quaternary lowstand deposits below 'glacio-eustatic fluctuation base-level. Sediment. Geol. 112, 195–217.

Custodio, E., 1987. Seawater intrusion in the Llobregat delta near Barcelona (Catalonia, Spain). In: Groundwater Problems in the Coastal Areas, Studies and Reports in Hydrology, vol. 45. UNESCO, Paris, France, pp. 436–463.

Delany, S.J., Cunningham, P., Doyle, D., 2005. Generating estimates of classification confidence for a case-based spam filter. Proc. of ICCBR '05 3620, 170–190 of LNAI.

Evangelidis, K., Papadopoulos, T., Papatheodorou, K., Mastorokostas, P., Hilas, C., 2018. 3D geospatial visualizations: animation and motion effects on spatial objects. Comput. Geosci. 111, 200–212. https://doi.org/10.1016/j.cageo.2017.11.007.

Font, J., Julia, A., Rovira, J., Salat, J., Sanchez-Pardo, J., 1987. Circulación marina en la plataforma continental del Ebro determinada a partir de la distribución de masas de agua y los microcontaminantes orgánicos en el sedimento. Acta Geol. Hisp. 21–22, 483–489.

Ford, J., Mathers, S., Royse, K., Aldiss, D., Morgan, D.J.R., 2010. Geological 3D modelling: scientific discovery and enhanced understanding of the subsurface, with examples from the UK. Z. Der Dtsch. Ges. Fur Geowiss. 161, 205–218.

Gámez, D., Simó, J.A., Lobo, F.J., Barnolas, A., Carrera, J., Vázquez-Suñé, E., 2009. Onshore–offshore correlation of the Llobregat deltaic system, Spain: development of

deltaic geometries under different relative sea-level and growth fault influences. Sediment. Geol. 217, 65–84.
Gou, J., Ma, H., Ou, W., Zeng, S., Rao, Y., Yang, H., 2019. A generalized mean distance-based k-nearest neighbor classifier. Expert Syst. Appl. 115, 356–372. https://doi.org/10.1016/j.eswa.2018.08.021.
Gur, I., Nachum, O., Miao, Y., Safdari, M., Huang, A., Chowdhery, A., Narang, S., Fiedel, N., Faust, A., 2022. Understanding HTML with Large Language Models. https://doi.org/10.48550/arXiv.2210.03945 arXiv.
Hobona, G., James, P., Fairbairn, D., 2006. Web-based visualization of 3D geospatial data using Java3D. IEEE Comput. Graph. Appl. 26, 28–33. https://ieeexplore.ieee.org/document/1652923.
Hu, R., Delany, S.J., Mac Namee, B., 2009. Sampling with confidence: using k-nn confidence measures in active learning. Proc. of ICCBR 9, 181–192.
Huang, S., Huang, M., Lyu, Y., 2020. An improved KNN-based slope stability prediction model. Adv. Civ. Eng., 8894109 https://doi.org/10.1155/2020/8894109.
Husillos, C., 2022. cesarhusrod/sarai_piezo_precip: Mejora de la documentación (v1.0.1). https://doi.org/10.5281/zenodo.7197288. Zenodo.
IGME, 1986. Geological Map of the Spanish Continental Shelf and Adjacent Areas, Scale 1:200,000, Sheet Nº 42; Tarragona, Memory and Maps. Geological Survey of Spain, Madrid, Spain. Available online: https://info.igme.es/cartografiadigital/tematica/Fomar200Hoja.aspx?language=es&id=42. (Accessed 18 April 2022). accessed on.
IGME, 1989. Geological Map of the Spanish Continental Shelf and Adjacent Areas, Scale 1:200,000, Sheet Nº 42E; Barcelona, Memory and Maps. Geological Survey of Spain, Madrid, Spain. Available online: https://info.igme.es/cartografiadigital/tematica/Fomar200Hoja.aspx?language=es&id=42E. (Accessed 18 April 2022). accessed on.
Iribar, V., Carrera, J., Custodio, E., Medina, A., 1997. Inverse modelling of seawater intrusion in the Llobregat delta deep aquifer. J. Hydrol. 198, 226–247.
Jessell, M., 2001. Three-dimensional geological modelling of potential-field data. Comput. Geosci. 27, 455–465.
Leifeld, P., 2013. Texreg: conversion of statistical model output in R to LATEX and HTML tables. J. Stat. Software 55, 1–24.
Li, Y., Yang, Z., Chen, X., Yuan, H., Liu, W., 2019. A stacking model using URL and HTML features for phishing webpage detection. Future Generat. Comput. Syst. 94, 27–39.
Llopis, N., 1942. Tectomorfología del Macizo del Tibidabo y valle inferior del Llobregat. Estud. Geográficos 3, 321–383.
Mangano, G., Zecchin, M., Civile, M., Ceramicola, S., Donato, A., Muto, F., Tripodi, V., Critelli, S., 2022. Mid-miocene to recent tectonic evolution of the punta stilo swell (calabrian arc, southern Italy): an effect of calabrian arc migration. Mar. Geol. 448, 1–11. https://doi.org/10.1016/j.margeo.2022.106810, 106810.
Manzano, M., 1986. Estudio Sedimentológico del Prodelta Holoceno del Llobregat. Master's Thesis, University of Barcelona, Barcelona, Spain, p. 82.
Marqués, M.A., 1984. Les Formacions Quaternàries del Delta del Llobregat. Institut d'Estudis Catalans, Barcelona, Spain, p. 295.
Medialdea, J., Solé-Sabarís, L., 1973. Geological Map of Spain, Scale 1:50,000, Sheet nº 420; Hospitalet de Llobregat, Memory and Maps. Geological Survey of Spain, Madrid, Spain. Available online: http://info.igme.es/cartografiadigital/geologica/Magna50Hoja.aspx?language=es&id=420. (Accessed 18 April 2022). accessed on.
Medialdea, J., Solé-Sabarís, L., 1991. Geological Map of Spain, Scale 1:50,000, Sheet nº 448; El Prat de Llobregat, Memory and Maps. Geological Survey of Spain, Madrid, Spain. Available online: http://info.igme.es/cartografiadigital/geologica/Magna50Hoja.aspx?language=es&id=448. (Accessed 18 April 2022). accessed on.
Miao, R., Song, J., Zhu, Y., 2017. 3D geographic scenes visualization based on WebGL, 1. In: 6th International Conference on Agro-Geoinformatics. IEEE: Fairfax VA, USA, pp. 1–6. Available online: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8046999. (Accessed 9 June 2022). accessed on.
Murphy, A., Redfern, S., 2015. Confidence measures in multiclass speech emotion recognition using ensemble learning to catch blunders. Int. J. Sci. Technol. Eng. 2, 118–122. https://ijste.org/Article.php?manuscript=IJSTEV2I3013.
Official Statement, 2004. The water authority of Catalonia creates the technical unit of the Llobregat Aquifers. In: Official Journal of Catalonia. Department of the Environment and Housing, Government of Catalonia, Barcelona, Spain. Available online. https://govern.cat/salapremsa/notes-premsa/68710/agencia-catalana-aigua-crea-mesa-tecnica-dels-aqueifers-del-llobregat. (Accessed 18 April 2022) (accessed on.
Parcerisa, D., Gámez, D., Gómez-Gras, D., Usera, J., Simó, J.A., Carrera, J., 2008. Estratigrafía y petrología del subsuelo precuaternario del sector SW de la depresión de Barcelona (Cadenas Costeras Catalanas, NE de Iberia). Rev. Soc. Geológica España 21, 93–109.
Perea, H., Massana, E., Santanach, P., 2012. An active zone characterized by slow normal faults, the northwestern margin of the Valencia Trough (NE Iberia): a review. J. Iber. Geol. 38, 31–52.
Postigo, C., Ginebreda, A., Barbieri, M.V., Barceló, D., Martín-Alonso, J., de la Cal, A., Boleda, M.R., Otero, N., Carrey, R., Solà, V., Queralt, E., Isla, E., Casanovas, A., Frances, G., López de Alda, M., 2021. Investigative monitoring of pesticide and nitrogen pollution sources in a complex multi-stressed catchment: the lower Llobregat River basin case study (Barcelona, Spain). Sci. Total Environ. 755, 142377 https://doi.org/10.1016/j.scitotenv.2020.142377.
Pratama, H., 2018. Machine learning: using optimized KNN (K-nearest neighbors) to predict the facies classifications, 1. In: Proceedings of the 13th SEGJ International Symposium. Society of Exploration Geophysicists of Japan, Tokyo, Japan, pp. 538–541. https://doi.org/10.1190/SEGJ2018-139.1.
Pyrcz, M., 2023. GeostatsGuy lectures. Available online: https://www.youtube.com/c/GeostatsGuyLectures. (Accessed 22 April 2023).
Resolution 12956/1994, 1994. Cooperation agreement on infrastructure and environment in the Llobregat Delta. In: Official Journal of Spain. Ministry of Public Works, Transports and Environment; Government of Spain, Madrid, Spain. Available online: https://www.boe.es/diario_boe/txt.php?id=BOE-A-1994-12956. (Accessed 18 April 2022). accessed on.
Rohmer, O., Bertrand, E., Mercerat, E.D., Régnier, J., Pernoud, M., Langlaude, P., Alvarez, M., 2020. Combining borehole log-stratigraphies and ambient vibration data to build a 3D Model of the Lower Var Valley, Nice (France). Eng. Geol. 270, 105588.
Salvany, J.M., Aguirre, J., 2022. The Neogene and Quaternary deposits of the Barcelona city through the high-speed train line. Geol. Acta 18, 1–19. https://doi.org/10.1344/GeologicaActa2020.18.10.
Sanz de Galdeano, C., 2019. Active faults in iberia. In: Quesada, C., Oliveira, J.T. (Eds.), The Geology of Iberia: a Geodynamic Approach, Regional Geology Reviews". Springer Nature, Switzerland, pp. 33–75.
Semmo, A., Trapp, M., Jobst, M., Doellner, J., 2015. Cartography-oriented design of 3D geospatial information visualization–overview and techniques. Cartogr. J. 52, 95–106. https://doi.org/10.1080/00087041.2015.1119462.
Serra, J., Verdaguer, A., 1983. La plataforma holocena en el prodelta del Llobregat. In: Obrador, A. (Ed.), X Congreso Nacional de Sedimentología, vol. 2. University of Barcelona, Barcelona, Spain, pp. 49–51.
Simó, J.A., Gámez, D., Salvany, J.M., Vázquez-Suñé, E., Carrera, J., Barnolas, A., Alcalá, F.J., 2005. Arquitectura de facies de los deltas cuaternarios del río Llobregat, Barcelona, España. Geogaceta 38, 171–174.
Solé-Sabarís, L., 1963. Ensayo de interpretación del Cuaternario Barcelonés. Misc. Barcinonensia 2, 7–54.
Tripodi, V., Muto, F., Brutto, F., Perri, F., Critelli, S., 2018. Neogene-Quaternary evolution of the forearc and backarc regions between the Serre and Aspromonte Massifs, Calabria (southern Italy). Mar. Petrol. Geol. 95, 328–343. https://doi.org/10.1016/j.marpetgeo.2018.03.028.
Vázquez-Suñé, E., Abarca, E., Carrera, J., Capino, B., Gámez, D., Pool, M., Simó, T., Batlle, F., Niñerola, J.M., Ibáñez, X., 2006. Groundwater modelling as a tool for the European water framework directive (WFD) application. The Llobregat case. Phys. Chem. Earth 31, 1015–1029. https://doi.org/10.1016/j.pce.2006.07.008.
Vázquez-Suñé, E., Marazuela, M.A., Velasco, V., Diviu, M., Pérez-Estaún, A., Álvarez-Marrón, J., 2016. A geological model for the management of subsurface data in the urban environment of Barcelona and surrounding area. Solid Earth 7, 1317–1329.
Wang, X., Yang, S., Zhao, Y., Wang, Y., 2018. Lithology identification using an optimized KNN clustering method based on entropy-weighed co-sine distance in Mesozoic strata of Gaoqing field, Jiyang depression. J. Petrol. Sci. Eng. 166, 157–174. https://doi.org/10.1016/j.petrol.2018.03.034.
Wentworth, C.K., 1922. A scale of grade and class terms for clastic sediments. J. Geol. 30, 377–392. https://doi.org/10.1086/622910.
Wycisk, P., Hubert, T., Gossel, W., Neumann, C., 2009. High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites. Comput. Geosci. 35, 165–182.
Yukun, L., Zhenguo, Y., Xu, C., Huaping, Y., Wenyin, L., 2019. A stacking model using URL and HTML features for phishing webpage detection. Future Generat. Comput. Syst. 94 (27–39) https://doi.org/10.1016/j.future.2018.11.004. ISSN 0167-739X.
Zecchin, M., Ceramicola, S., Gordini, E., Deponte, M., Critelli, S., 2011. Cliff overstep model and variability in the geometry of transgressive erosional surfaces in high-gradient shelves: the case of the Ionian Calabrian margin (southern Italy). Mar. Geol. 281, 43–58.

Websites

Albion: 3D geological models in QGIS. Available online: https://gitlab.com/Oslandia/albion.
GemPy: Open-source 3D geological modeling. Available online: https://www.gempy.org.
GeoPandas. Available online. https://geopandas.org/en/stable.
GISgeography. 15 Python Libraries for GIS and Mapping. Available online: https://gisgeography.com/python-libraries-gis-mapping.
OSGeo: The Open Source Geospatial Foundation. Available online: https://www.osgeo.org/.
Parpoil, B. Open source and geology. Available online: https://oslandia.com/en/2020/07/09/geologie-open-source.