



UNICA

UNIVERSITÀ
DEGLI STUDI
DI CAGLIARI



Università di Cagliari

UNICA IRIS Institutional Research Information System

This is the Author's [*accepted*] manuscript version of the following contribution:

[Paolo Ciampi, Carlo Esposito, Giorgio Cassiani, Gian Piero Deidda, Adrian Flores-Orozco, Paolo Rizzetto, Andrea Chiappa, Manuele Bernabei, Andrea Gardon, Marco Petrangeli Papini, Contamination presence and dynamics at a polluted site: spatial analysis of integrated data and joint conceptual modeling approach, *Journal of Contaminant Hydrology*, 248, 2022, pages 104026]

The publisher's version is available at:

<https://doi.org/10.1016/j.jconhyd.2022.104026>

When citing, please refer to the published version.

1 **Title**

2 Contamination presence and dynamics at a polluted site: spatial analysis of integrated data and joint conceptual
3 modeling approach

4 **Author names and affiliations**

5 Paolo Ciampi^{1,8*}, Carlo Esposito^{1,8}, Giorgio Cassiani², Gian Piero Deidda³, Adrian Flores-Orozco⁴, Paolo Rizzetto⁵,
6 Andrea Chiappa⁶, Manuele Bernabei⁶, Andrea Gardon⁶, Marco Petrangeli Papini^{7,8}

7 1. Department of Earth Sciences, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy;
8 paolo.ciampi@uniroma1.it, carlo.esposito@uniroma1.it

9 2. Department of Geosciences, University of Padua, Via Gradenigo 6, 35131, Padua, Italy;
10 giorgio.cassiani@unipd.it

11 3. Department of Civil, Environmental Engineering and Architecture, University of Cagliari, via Marengo, 2,
12 09123, Cagliari, Italy; gpdeidda@unica.it

13 4. Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria; adrian.flores-
14 orozco@geo.tuwien.ac.at

15 5. Logistic Headquarter of Italian Air Force, Viale dell'Università, 4, 00185, Rome, Italy;
16 paolo.rizzetto@am.difesa.it

17 6. Department of Technological Aerospace Materials–Flight Test Center of Italian Air Force, Pratica di Mare,
18 00071, Pomezia, Rome Italy; andrea.chiappa@am.difesa.it, manuele.bernabei@am.difesa.it,
19 andrea.gardon@am.difesa.it

20 7. Department of Chemistry, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy;
21 marco.petrangelipapini@uniroma1.it

22 8. CERI Research Center, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy.

23 * **Corresponding author:** paolo.ciampi@uniroma1.it; Tel.: (+39 3489194607)

24

25

26

27 **Abstract**

28 Contaminated sites are complex systems posing challenges for their characterization as both contaminant distribution
29 and hydrogeological properties vary markedly at the metric scale, yet may extend over broad areas, with serious issues
30 of spatial under-sampling in the space. Characterization with sufficient spatial resolution is thus, one of the main
31 concerns and still open areas of research. To this end, the joint use of direct and indirect (i.e., geophysical) investigation
32 methods is a very promising approach. This paper presents a case study aspiring to demonstrate the benefit of a
33 multidisciplinary approach in the characterization of a hydrocarbon-contaminated site. Detailed multi-source data,
34 collected via stratigraphic boreholes, laser-induced fluorescence (LIF) surveys, electrical resistivity tomography (ERT)
35 prospecting, groundwater hydrochemical monitoring, and gas chromatography-mass spectrometry (GC-MS) analyses
36 were compiled into an interactive big-data package for modeling activities. The final product is a comprehensive
37 conceptual hydro-geophysical model overlapping multi-modality data and capturing hydrogeological and geophysical
38 structures, as well as contamination distribution in space and dynamics in time. The convergence of knowledge in the
39 joint model verifies the possibility of discriminating geophysical findings based on lithological features and
40 contamination effects, unmasking the real characteristics of the pollutant, the contamination mechanisms, and the
41 residual phase hydrocarbon sequestration linked to the hydrogeological dynamics and adopted remediation actions. The
42 emerging conceptual site model (CSM), emphasizing the necessity of a large amount of multi-source data for its
43 reliable, high-resolution reconstruction, appears as the necessary tool for the design of remedial actions, as well as for
44 the monitoring of remediation performance.

45 **Keywords**

46 Hydrocarbon contamination; 3D modeling; multi-source geodatabase; laser-induced fluorescence; electrical resistivity
47 tomography.

48

49

50

51

52

53

54 **1. Introduction**

55 Contamination and hazards related to leaking underground fuel storage tanks represent an open environmental problem
56 that needs to be addressed through the investigation and remediation of petroleum hydrocarbon sites (Ghosh et al. 2019;
57 McCall et al. 2018). Petroleum hydrocarbons are part of the contaminant class known as light non-aqueous phase
58 liquids (LNAPLs) (Vasudevan et al. 2016). These widespread and persistent pollutants are typically released into the
59 environment as mixtures of various chemical compounds (Åslund et al. 2013). Furthermore, LNAPL aging and
60 weathering induce mutations in the composition of the mixture, resulting in an impoverishment of both volatile and
61 soluble chemicals (Totsche et al. 2003), thus accumulating toxic, semi-, and non-volatile, insoluble constituents of
62 heavier molecular weight (Lari et al. 2019). This last aspect tends to affect the selection of a reasonable characterization
63 method and the choice of an appropriate approach for remediation (Brusseau, 2019; Suthersan et al. 2016).

64 An adequate characterization of the nature, chemical transformation, and spatial distribution of LNAPLs in the
65 subsurface is one of the main open research questions (Lari et al. 2018; Totsche et al. 2003). Detailed local data may be
66 obtained through core and borehole surveys, but such evidence is inherently 1D and unevenly distributed (e.g., Deiana
67 et al. 2007). Field studies highlight the limitations and substantial errors that result from the use of traditional
68 prospecting techniques (i.e., soil coring and groundwater monitoring) in estimating the amount and spatial extent of
69 LNAPLs in the subsurface (Algreen et al. 2015). Investigations with direct methods are affected by the limited number
70 of samples across a 3D potentially contaminated space (the subsoil) inevitably leading to spatial aliasing and inaccurate
71 reconstruction of the pollution spatial extent (Binley et al. 2015; Cassiani et al. 2014; Ciampi et al. 2021b; Crook et al.
72 2008; Deiana et al. 2007; McCall et al. 2018). Aliasing occurs when the sampling frequency is inadequately low
73 compared with the frequency of signal variation (Shannon, 1949). As a result of spatial aliasing, the sampled variable
74 assumes smooth variations in space, with a spatial frequency that is much lower than the true one, thus appearing
75 different from what reality is (i.e., an alias). The impact of such aliasing on the assessment of the contamination extent
76 and total pollutant mass is dramatic, resulting in the overestimation of contaminated volumes and pollutant masses.
77 Correspondingly, geophysical methods capture the subsurface with high spatial resolution, permitting to depict
78 hydrogeological heterogeneities (Ruggeri et al. 2014), and define confining geological structures which control
79 groundwater flow and contaminant migration. Hence, geophysical methods are potentially able to bridge the gap
80 between resolution and coverage associated with conventional hydrological investigations (e.g., Crook et al. 2008).
81 Geophysical investigations may characterize the distribution of a plume with high LNAPL concentration (Bücker et al.
82 2017; Caterina et al. 2017; Flores Orozco et al. 2012, 2015, 2019a, 2019b, 2021; Xia et al. 2021), monitor LNAPL leaks
83 and the evolution of the pollution source (Shao et al. 2019) and thus avoid the interpolation of ground truth data. In the

84 context of environmental science, geophysical techniques have become an effective instrument to assist the study of the
85 shallow subsurface and to control hydrological dynamics and hydrochemical processes (e.g., Binley et al. 2010, 2015).
86 A few of the distinctive advantages of geophysical exploration tools include minimizing the requirement for direct
87 intrusive surveys (Chambers et al. 2010) and delivering spatially continuous records of subsurface geology (Samouelian
88 et al. 2005). Some geophysical methods may emphasize potential relationships between the meaningful measured
89 physical parameters and the hydrological and environmental crucial aspects concerning the contaminated site
90 characterization (Cassiani et al. 2014). Site investigation using different survey techniques (boreholes and geophysical
91 methods) in combination with an integrated approach for data interpretation can reduce the collection of redundant
92 information (Abbaspour et al. 2000). Also, merging linear prospection methods, which assist spatialization of data, with
93 traditional, non-substitutable point survey techniques (Binley et al. 2015; Crook et al. 2008) can go beyond the
94 limitations arising from their distinct implementation. On the one hand, conventional investigations are expensive,
95 invasive, one-dimensional, and usually characterized by limited densities and irregular distributions. On the other hand,
96 geophysical techniques cannot replace in situ sampling, do not directly record lithological or contaminant properties,
97 and require interpretation by conventional methods to avoid potential misinterpretation of findings (Arato et al. 2014).
98 Note that the valid hydrological understanding of geophysical data is influenced by the constitutive links (e.g.,
99 petrophysical relationships) that translate recorded geophysical parameters (e.g., electrical conductivity) into
100 hydrological properties (e.g., water and clay content) (Binley et al. 2015).

101 While electrical properties may be linked to lithological structure, it is extremely challenging to link geophysical
102 signals to contamination (Binley et al. 2015). A considerable amount of experimental research on the geophysical
103 response of contaminants has revealed a gap in understanding a comprehensive physicochemical framework capable of
104 relating geophysical signatures and directly measured pollutant characteristics (Arato et al. 2014; Binley et al. 2015;
105 Cassiani et al. 2014; Flores Orozco et al. 2021; Prasanna et al. 2008). Due to the complexity and the large number of
106 variables involved in physicochemical processes within polluted porous media, the pursuit for common and versatile
107 models that couple geophysical records with contaminant features is not effective (Binley et al. 2015), and some degree
108 of site-specific relationships have generally to be sought (Cassiani et al. 2014). In this respect, ancillary direct
109 information about contaminant presence and state is essential. In this regard, the development of laser-induced
110 fluorescence (LIF) technology helps to deliver direct knowledge on LNAPL migration and distribution with high
111 resolution (Teramoto et al. 2019). LIF is a direct, real-time, and in-situ detection system for screening non-aqueous free-
112 phase pollutants in the subsurface. LIFs measure a percentage fluorescence intensity relative to the standard calibration,
113 known as the reference emitter (RE), which reflects the amount of oil in the pores (Teramoto et al. 2019). The LIF
114 technology utilizes ultraviolet (UV) laser light provided by direct push boring instruments to excite polycyclic aromatic

115 hydrocarbons (PAHs) molecules present in LNAPLs and simultaneously records the resulting fluorescence as a function
116 of depth, enabling the semi-quantitative characterization of LNAPL distribution within the subsurface at least in terms
117 of quasi-continuous 1D profiles with depth (Pepper et al. 2002). The LIF measurements coupled with cone
118 penetrometer testing (CPT) have been extensively used for time- and cost-effective in situ detection of fuels and
119 petroleum products and demonstrated their effectiveness in obtaining geophysical and geotechnical properties (via
120 specific sondes placed on direct-push devices) of subsurface environments (Einarson et al. 2018; Gruiz et al. 2017;
121 Pepper et al. 2002). Note that currently available LIF equipment is not designed to detect dissolved-phase contaminants
122 (Fedotov et al. 2019). In the investigation and management of contaminated sites, the challenge is to integrate the
123 information coming from different data sources to provide a consistent, realistic, and accurate conceptual model (Harris
124 et al. 2004). Usually, multi-modality data analyze different parameters, in different configurations, with various
125 investigation depths. Instead of handling each data set individually, a single, coherent image (model) should be
126 generated. (Pollard et al. 2004). Coupled hydrogeophysical techniques aim to bridge this gap, but a knowledge
127 harmonization procedure is still an open area of research (Binley et al. 2015). The synthesis of a huge volume of
128 information and diverse sources of experience typically found at most polluted sites into a convergent, hybrid, and
129 multi-source geodatabase may simply develop and enhance a model by collecting and incorporating new evidence or
130 reinterpreting and validating available data (Binley et al. 2015; Chiabrande et al. 2019).

131 In this study, we suggest a stepwise refinement methodology to develop a comprehensive 3D conceptual site model
132 including multi-source data gained from direct and indirect methods. The expected data-driven model contributes to the
133 convergence of different types of spatial subsurface information (i.e., lithological, hydrogeological, geophysical,
134 chemical, geotechnical), establishing a connection between the environmental variables to overcome both the spatial
135 sampling limitations of direct methods and the interpretation of geophysical investigations. We aim at investigating the
136 effectiveness of a single big-data package and multi-source hydrogeophysical model, capturing hydrogeological and
137 geophysical evidence, as well as contamination dynamics over time. The application of data fusion has the goals to (i)
138 reduce the uncertainty associated with subsurface interpretation, (ii) decipher geophysical findings based on geological,
139 chemical, and physical information, and (iii) provide compelling insights into LNAPL behavior in the saturated and
140 unsaturated domains. The presented case study concerns contamination caused by jet fuel in a military airbase in Italy.
141 From the approach investigated here, integration of heterogeneous data in nature and resolution demonstrates to provide
142 additional information without the requirement for additional investigations, differentiating geophysical results based on
143 lithologic characteristics and contamination effects as well as revealing the actual distribution and mechanisms of
144 contamination, pollutant aging, and residual phase hydrocarbon sequestration related to hydrogeologic dynamics and
145 adopted remediation measures.

146 2. Materials and Methods

147 2.1 *The case study: remediation history and available data*

148 The study site is the military airport of Decimomannu (Cagliari, Italy), affected by jet fuel-JP8 spills due to a leaking in
149 a fuel transfer line (Ciampi et al. 2021b). The spills have occurred in 2007 (40 m³), in 2009 (5 m³), and 2010 (5 m³).
150 The remediation/safety measure adopted at the site consists of pumping wells and a hydraulic barrier for groundwater
151 extraction (Brusseau, 2019). During the characterization and remediation of the site, a suite of investigations has taken
152 place (Tab. 1 of Supplementary Material), such as grain size analysis of cores, hydrogeological tests, geophysical
153 surveys, groundwater samplings and analyses. In total, 85 stratigraphic boreholes were realized from 2007 to 2016 to
154 deliver an overview of the geological sequence found at the site. They reach depths ranging from 10 m to 26 m below
155 ground and cover an investigation area of about 26.5 hectares. The deposit permeability was first estimated from the
156 grain distributions reported in Flores Orozco et al. (2021). Additionally, two pumping tests and ten slug tests were
157 performed as part of this work to provide a measurement of the aquifer permeability coefficient. For the monitoring of
158 groundwater levels and contamination, 62 piezometers have been installed on-site. Periodic hydrochemical
159 measurements were made on the piezometric network between 2011 and 2018, providing the necessary information to
160 deduce the evolution of the hydrocarbon contaminant plume. Some comparative analyses, performed through gas
161 chromatography-mass spectrometry (GC-MS) on "fresh" products (original jet fuel) and supernatant sporadically
162 recovered in the piezometric network, enriched the collected data. Such analyses aimed at delivering the speciation of
163 the hydrocarbon mixture components to research any evidence of aging or weathering (Vozka et al. 2019). Additionally,
164 to gain direct information about the distribution of the NAPL, 30 points were surveyed using the LIF-ultraviolet optical
165 screening tool (UVOST) technology in combination with CPT measurements.

166 2.2 *The geophysical dataset*

167 The geophysical dataset consists of several surface Electrical Resistivity Tomography (ERT) - (e.g., Binley and Kemna,
168 2005; Cassiani et al. 2014; Crook et al. 2008) - lines located both inside and outside of the base perimeter. The detailed
169 map of the ERT investigations realized as part of this study is shown in Figure 1 of the Supplementary Material. Two
170 slightly different acquisition strategies were adopted:

- 171 • Line 15, which runs along the Southern side of the airbase, is a 330 m line with 1 m electrode spacing,
172 composed of 7 individual ERT lines made of 72 electrodes each, with a partial superposition of neighboring
173 lines of 24 electrodes;
- 174 • Lines 1-14 are single ERT lines made of 48 electrodes each, with electrode spacing equal to 1 m.

175 In both cases, a dipole-dipole skip-4 acquisition scheme was adopted (skipping 4 electrodes in each dipole means that
176 the dipole lengths, for both current injection and voltage difference measurement are 5 m long). The full reciprocal
177 acquisition was performed to assess measurement errors, as good practice for high-quality data surveys (e.g. Cassiani et
178 al. 2006). A complete acquisition of all reciprocals (swapping potential with current electrodes) is essential for
179 estimating the errors in the acquisition and permits the elimination of outliers before data inversion (Binley et al. 1995).
180 The location of the ERT profiles (see Figure 1 of the Supplementary Material for details) was decided based on the need
181 to investigate geological constraints on contamination dynamics in areas hydrogeologically down gradient of spills and
182 the hydraulic barrier. In particular, line 15 covers the entire Southern border of the base in the area of interest, while the
183 other short lines sample with fine detail the areas where contamination, and thus biodegradation, is expected to be
184 maximal, with a few lines also placed outside the expected contaminant plume to provide background (uncontaminated
185 information). The ERT lines to the West cover for the most part, inside and outside the base fence, the region where
186 clays are expected to be very shallow or emerge at the surface.

187 In all cases, inversion of ERT was conducted using the Profiler-R2 suite of programs provided by Lancaster University
188 (<http://www.es.lancs.ac.uk/people/amb/Freeware/R2/R2.htm>) now incorporated in the ResIPy package (Blanchy et al.
189 2020). The inversion strategy is based on Occam's approach, thus obtaining the smoothest model compatible with the
190 error in the data, in this case, equal to 5% reciprocal error.

191 *2.3 The big-data package for multi-source geomodeling*

192 This huge volume of different source data was then georeferenced and added into an interactive big-data package,
193 which is structured as a multiple excel worksheet and relational database (Ciampi et al. 2019b). The developed multi-
194 thematic, four-dimensional (4D) big-data package considers time as the fourth dimension and should permit the
195 management, integration, and release of data during the knowledge acquisition phase (Ciampi et al. 2019a, 2021a),
196 behaving as a decision support tool (DST) during the remediation period (see e.g., Huysegoms and Cappuyns, 2017).
197 The big data package aided in the planning of investigations as it permitted to follow field findings, and in time-lapse
198 monitoring of remediation actions at the pilot-scale described by Ciampi et al. 2021b. Grid and block spatial modeling
199 of multidisciplinary data contained in the geodatabase has the purpose of generating a multi-source conceptual model in
200 2D or 3D, containing geological, hydrochemical, and geophysical information (Ciampi et al. 2021b; Wang and Huang,
201 2012). The 3D conceptualization of spatial and physical parameters was built using the RockWorks 17 software
202 (Ciampi et al. 2019b). The reconstruction of a solid model that overlays different types of knowledge arises from the
203 spatial interpolation and joint processing of the geological, geophysical, and hydrochemical parameters (Kaliraj et al.
204 2015, Safarbeiranvand et al. 2018). The parameters include stratigraphic borehole data (depth and lithological types),

205 groundwater level elevations, LIF data (percentage fluorescence intensity), geophysical information (resistivity),
206 geotechnical records (cone resistance), and the chemical analysis of water sampled (contaminant concentration). The
207 interpolation of all the above point data was performed using the algorithm of inverse distance weighting (Mirzaei and
208 Sakizadeh, 2016; Safarbeiranvand et al. 2018) to generate quasi-continuous 3D models (and 2D canvas), in which the
209 spatial distribution of the parameters obtained from all investigations is easily accessible. The inverse squared distance
210 (weighting exponent of 2) was used and the search neighborhood was limited to 4 points, so that the extrapolated value
211 gradually approaches the value of the nearest sample point, honoring the data value (Liu et al. 2020). Additional options
212 included a high fidelity filter to preserve control point values and low smoothing (Falivene et al. 2010). Joint grid and
213 block modeling, which employs elementary volumes (voxels) in a three-dimensional mesh, aims to store, overlay, and
214 represent multi-source information related to stratigraphic, piezometric, resistivity, fluorescence, and cone resistance
215 data in a geo-referenced space (Høyer et al. 2015). The 3D mesh covers the area of the airbase and extends vertically
216 from 8.8 m below sea level to 23.4 m above sea level, representing the maximum depth of investigation and the
217 maximum elevation of the ground surface respectively. The solid and block geological model was built by interpolating
218 the top and base grid surfaces of each unit listed in the database to isolate the volumes of the different strata. Spatial
219 interpolation of water table elevations and contaminant concentrations was performed to generate the piezometric
220 surface and contamination state contour maps. Voxel modeling was used to interpolate data acquired via ERT, LIF, and
221 CPT. A voxel stores a single numerical value for each physical parameter assigned by interpolation to explain potential
222 spatial relationships among aggregated complex data-driven structures. The voxel grid discretization is 0.5 m x 0.5 m x
223 0.2 m in the x, y, z directions. The multi-source block model has a size of 780 x 1379 x 162 voxels. These dimensions
224 were chosen to achieve a high resolution of the mapped geological structures, consistent with the acquisition resolution
225 of ERT investigations. Signal acquisition from LIF-CPT images was averaged at the set voxel resolution to combine
226 multi-modality data within a unique 3D mesh domain. A distance clipping filter was employed to limit the resistivity
227 model based on a node's distance of 5 m from the ERT lines. The integrated extraction of geological-physical attributes
228 from each voxel of the 3D mesh and their coupled analysis was intended to geostatistically discriminate lithological
229 structures based on electrical properties. Such a joint-modeling approach has the purpose of developing a CSM that
230 considers extension and degree of contamination, characteristics, and chemical-physical parameters that condition the
231 mobility and the pollutant partition among aqueous, non-aqueous, solid, and gas phases. The adopted holistic approach
232 aims to demonstrate how the joint integration of the different investigations overcomes not only the limitations related
233 to their single applicability but also the indirect and thus uncertain nature of non-invasive investigations.

234 **3. Results**

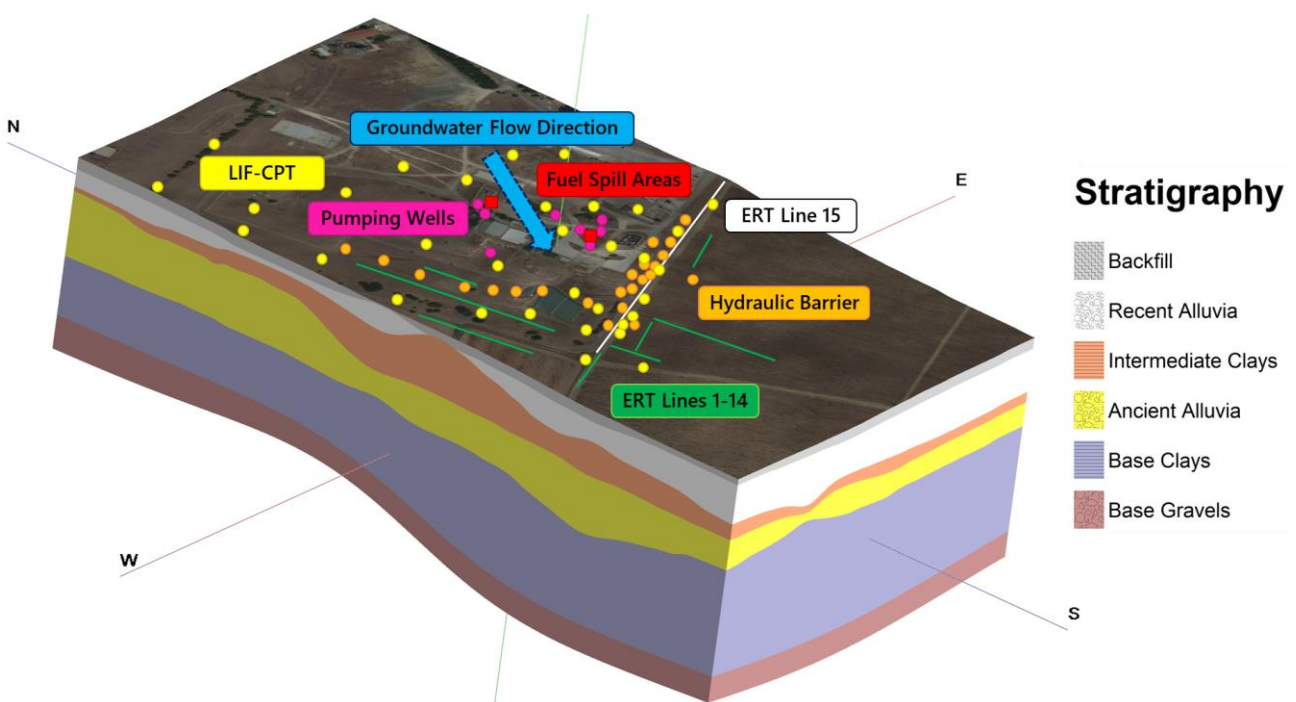
235

3.1 Geological and Hydrogeological Settings

236 In the Decimomannu airbase area, the most recent deposits are related to a Plio-Quaternary depositional sequence of
237 alluvial sediments (Bini, 2013; Reuter et al. 2017). Building on the information collected through the execution of
238 stratigraphic surveys and as illustrated in Fig. 1, the geological structure of the subsoil is subdivided as follows:

- 239 1. Backfill (anthropogenic) materials to a depth of 1-1.5 m;
- 240 2. Recent alluvia extending to depths between 1 m and 5 m and characterized by gravels and sands with a
241 presence of fine fraction;
- 242 3. Intermediate clays forming a horizon of sandy-gravelly clays having hazelnut color characterized by an
243 average thickness of 1.5 m;
- 244 4. Ancient alluvia defining a layer around 3.5 m thick (on average) comprised of gravel and sand in a silty-clay
245 matrix;
- 246 5. Base clays found in a thick level of clays and silty clays are located at depths between 10 m and 24 m;
- 247 6. Base gravels at a depth of about 24 m from the ground surface, a horizon made of gravels and sands immersed
248 in a silty-clayey matrix.

249 The reconstructed three-dimensional geological model reveals both the irregularity of the stratigraphic contacts and the
250 geometric structures that characterize the different horizons. A vertical exaggeration factor is used to mark the
251 lithological steps. The set of all information acquired during the phases of characterization and remediation converges
252 within the solid geo-referenced model (Figure 1).



253

254 **Fig. 1** Three-dimensional geological model of the Decimomannu military airbase depicting the stratigraphic
255 relationships. Position of the fuel spill areas, pumping wells, hydraulic barrier, LIF-CPT investigations, and ERT lines
256 inside the military domain

257 The recent and ancient alluvia have a highly variable thickness and are separated by intermediate clays. The alluvial
258 sequence hosts the shallow aquifer and overlies the base clays. Based on the particle size distribution of soil samples,
259 the hydraulic conductivity of the coarse-grained deposits ranges from 1.8×10^{-4} and 6.5×10^{-6} m/s for the recent alluvia
260 and between 1.7×10^{-6} to 1.9×10^{-8} m/s for the ancient alluvia as reported by Flores Orozco et al. (2021). A permeability
261 coefficient of about 4.2×10^{-9} m/s has been attributed to the intermediate clays, while for the base clays it approximates
262 2.7×10^{-10} m/s. Aquifer permeability obtained by slug tests varies between 9.96×10^{-4} and 2.54×10^{-6} m/s. The aquifer
263 permeability coefficient estimated by pumping tests ranges from 1.48×10^{-3} to 3.15×10^{-4} m/s. The base and the
264 intermediate clays have a relevant hydrogeological role as aquiclude and aquitard, respectively. The base gravels
265 constitute the confined aquifer, while the shallow aquifer is the most sensitive to fuel spills. In undisturbed conditions,
266 the piezometric surface stands at 4.5 m below ground level. Groundwater flows from NE to SW and is hosted in an
267 aquifer that exhibits variable conditions from locally phreatic to partially confined elsewhere. In such a geological
268 context, intensive extraction by pumping wells and hydraulic barrier may potentially trigger local modifications of
269 groundwater head distribution and drawdowns of the water table.

270 *3.2 Geophysical Model*

271 The acquisition and the incorporation of ERT profiles within the voxel-based and multi-source model provide the
272 necessary data to refine and strengthen the conceptual geological model, which arises from the interpolation of point
273 measurements. Hence, ERT data reinforces and validates stratigraphic data, and avoids potentially serious spatial
274 aliasing effects from interpolation of borehole data above ERT resolution threshold (Binley et al. 2015; Crook et al.
275 2008). In particular, the overlay of voxel-based geophysical data and geological information portrays a clear correlation
276 between the low resistivity layers and the clays. This correlation is expected due to the high surface area and surface
277 charge of clays, which in turn contribute to surface conductivity in addition to the electrolytic conductivity (e.g., Revil
278 et al. 2017; Flores Orozco et al. 2021). The intermediate layer of clays is of course the main structural feature affecting
279 groundwater flow, and thus contaminant distribution, because of their very low hydraulic conductivity. The geophysical
280 surveys concluded that the intermediate clays reach the maximum thickness in the western sector, while locally they
281 disappear to the east, where communication of groundwaters hosted in the alluvial sediments can occur, as illustrated in
282 Fig. 2.

283

284 **Fig. 2** Integrated three-dimensional model illustrating the results of the geophysical surveys, the layer of intermediate
285 clays, and the stratigraphic logs

286 The filling material reveals values of resistivity generally between 80 and 250 $\Omega\cdot\text{m}$ (log10 resistivity range: 1.9-2.4).
287 Coarse deposits constituting recent and ancient alluvia exhibit a variant geoelectric signature between 40 and 126 $\Omega\cdot\text{m}$
288 (log10 resistivity range: 1.6-2.1). The intermediate and base clays are distinguished by an average value between ~ 8
289 and 60 $\Omega\cdot\text{m}$ (log10 resistivity range: 0.9-1.8). The combined extraction of geological and physical properties from the
290 mesh elements of the data-driven model provides the resistivity distribution of shallow lithologies, correlating the
291 electrical behavior with the geological parameter (Fig. 3).

292

293 **Fig. 3.** Frequency histograms related to the resistivity (in log scale) of the mesh elements encoding the geological
294 information for the different stratigraphic horizons

295 Although some lithologies reveal overlaps in resistivity signals, incorporation of ERT prospecting into the geological
296 model discretizes geological-physical properties in space. This leads to the differentiation of geological heterogeneities
297 with high resolution, especially in the horizontal direction, managing uncertainties arising from both sources of
298 information. In the Western sector of the site, where the intermediate clays exhibit a larger thickness, the layers
299 characterized by a low electrical resistivity correspond to the clayey horizons, whereas the levels with a higher electrical
300 resistivity coincide with the sandy layers of the shallow aquifer. In the Eastern sector, the ERT profiles reveal higher
301 electrical resistivity values (Fig. 2). In this portion of the airbase, the clayey intermediate lens is almost totally absent,
302 with the preponderance of the recent alluvia. This affects the electrical properties of the subsoil, resulting in higher
303 resistivity values. Therefore, even though ERT is not able to quantify the hydraulic conductivity of porous media, it aids
304 in the discrimination between formations marked by different electrical and hydraulic resistivities (Cassiani et al. 2014).

305 Integrating the findings of geophysical investigations into the geological model leads to the geoelectrical
306 parametrization, the qualitative and quantitative geophysical interpretation of near-surface sediments. Accordingly, the
307 combination of the data potentially reduces the misinterpretation of stratigraphic variations resulting from interpolation
308 through increased density and resolution of geophysical data (Binley et al. 2015; Crook et al. 2008; Hermans and Irving,
309 2017). Statistical data analysis experimentally explores the results of the multi-source models, correlating the
310 stratigraphy of mesh elements with their resistivity distribution and providing an explanation of the electrical behavior
311 based on the knowledge of the lithological parameter.

312 Locally, the superposition of the geoelectric response does not allow for the discretization of deeper sediments. This
313 effect is observed in the hydrogeologically downstream portions of the spills. However, in such areas, the geophysical
314 model shows locally low resistivity values in the upper levels. The latter variations can be observed in figure 4 at a
315 depth between 1.5 and 4 m, in correspondence with some bands with a resistivity between 16 and 40 $\Omega \cdot m$ (\log_{10}
316 resistivity range: 1.2-1.6). From the cross-validation of the geological-geophysical section shown in Figure 5, it appears
317 plausible to hypothesize that the pronounced increase in electrical conductivity may be associated with biodegradation
318 activity at shallow sediments impacted by petroleum hydrocarbons rather than geological heterogeneity (see Cassiani et
319 al. 2014, and references therein).

321 **Fig. 4** Comparison of a stratigraphic and a geophysical section extrapolated from the big-data package at a trace located
322 hydrogeologically downstream of the jet fuel spills

323 The conceptual geological-geophysical model demonstrates qualitative and quantitative discrimination of ERT results
324 based on lithologic characteristics and contamination effects.

325 *3.3 Evolution of Groundwater Quality*

326 Utilizing total petroleum hydrocarbons (TPH) as the revealing pollution parameter, thematic maps have been
327 constructed using the big-data package. From 2011, the detected contamination appears to be fairly widespread and has
328 been the target of years of a pump and treat action (still operational). Such remediation has allowed both a reduction of
329 the contaminant mass and a restriction of the contaminant plume, which gradually reached an asymptotic pattern
330 (Ciampi et al. 2021b). The thematic maps presented in Figure 5 reveal a decrease in pollutant concentrations and the
331 plume shrinkage over time.

332

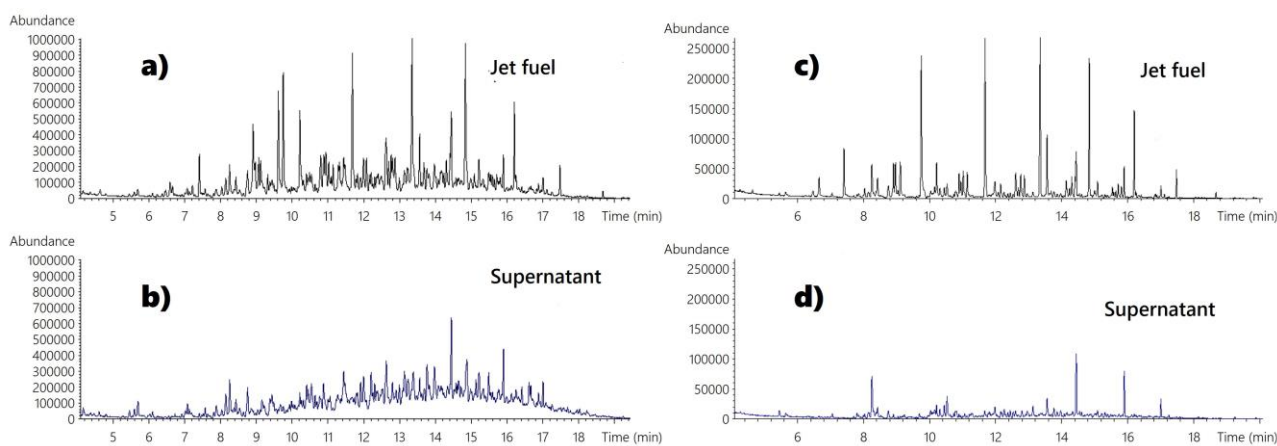
333 **Fig. 5** Contour maps of TPH concentrations in groundwater from 2011 to 2018

334 The evolutionary scenario reveals a decrease in TPH concentrations from values locally approximating 1 g/L (2011) to
335 a few mg/L (2018), hinting at the aging of the primary contamination source (Ciampi et al. 2021b). In the last
336 represented monitoring campaign, limited areas show the impact of significant dissolved concentrations in groundwater
337 (exceeding the limits established by Italian regulations). Such areas are mainly found at piezometers located within the

338 tank storage area and around the hydraulic barrier zone. Here monitoring wells locally exhibit a measured concentration
339 of TPH between 350 and 3500 $\mu\text{g/L}$. This phenomenon is likely to be linked to the production of bio-surfactants by
340 micro-organisms rendering oily substances more bio-available (see e.g., Cassiani et al. 2014). These “critical” areas
341 have been both historically affected by the presence of TPH in groundwater and by the infrequent appearance of
342 LNAPL as a separate phase. Variable apparent thicknesses of supernatant (up to 1 m) were rarely detected in
343 monitoring piezometers until 2013 (near the spill points and the hydraulic barrier). A sporadic sampling of limited
344 apparent thicknesses (less than 1 cm) of separated phase at extraction wells has been recorded since 2014.

345 3.4 Analytical evidence disclosing the source-aging scenario

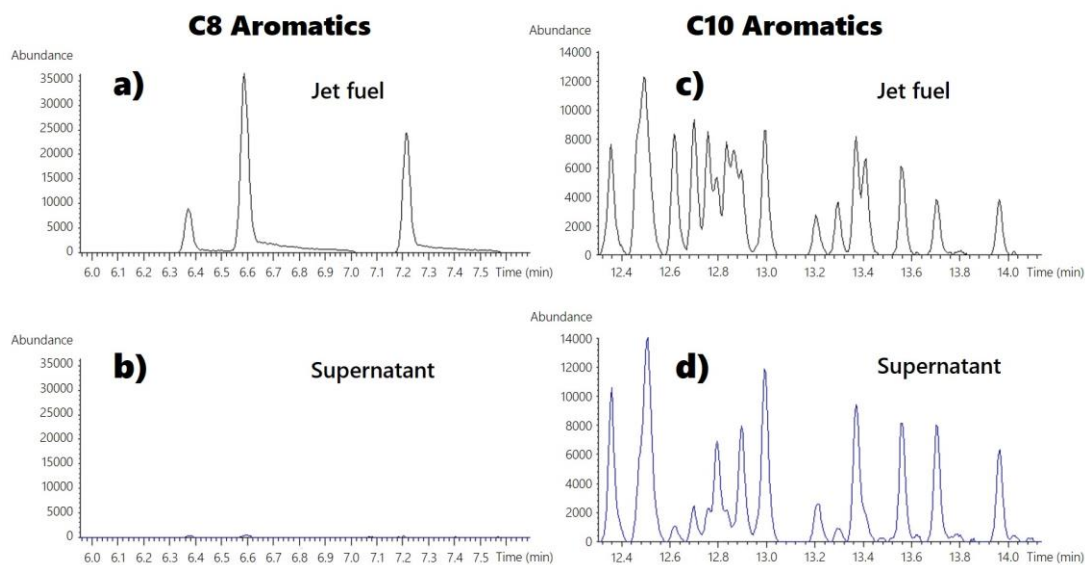
346 The multi-source big-data package was enriched by detailed speciation, through GC-MS (Vozka et al. 2019), of
347 supernatant that has been occasionally detected in the piezometric monitoring network. The chromatograms of jet fuel
348 and supernatant samples show a significant difference in the fingerprint regarding the peaks of more volatile fractions,
349 less present in the supernatant, coherently with the expected aging of the contaminant separate phase (Fig. 6a, b).
350 Besides, comparing the GC/MS chromatographic fingerprint of jet fuel and supernatant samples, a clear difference is
351 noted for all linear components that are drastically reduced in the supernatant. This is not surprising, as these fractions
352 are known to be more bioavailable to biodegradation (Tran et al. 2018) (Fig. 6c, d).



353

354 **Fig. 6** GC/MS chromatographic fingerprints (total ion) related to jet fuel (a) and supernatant (b) that was occasionally
355 collected as a free phase in the piezometric monitoring network. GC/MS chromatographic fingerprint relative to linear
356 aliphatics (C6 – C16) measured in the “fresh” jet fuel (c) and the supernatant (d)

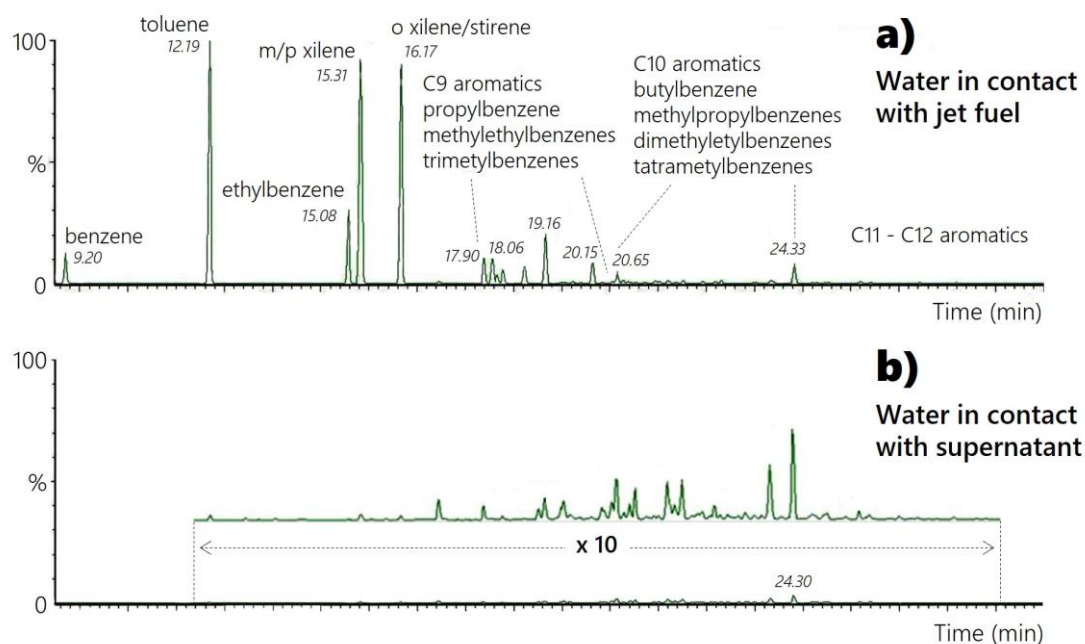
357 GC-MS analysis demonstrates that the medium-light aromatic fractions (C8- such as toluene, xylene, etc.) are nearly
358 completely absent in the supernatant sample (Tran et al. 2018). For medium-heavy compounds (C10-es. butylbenzene,
359 tetramethylbenzene) the phenomenon is much less pronounced (Fig. 7).



360

361 **Fig. 7** GC-MS analysis of medium-light (a, c) and medium-heavy (b, d) aromatic fractions measured in the jet fuel (a,
 362 b) and the ex-situ extracted free phase (c, d)

363 The analysis of dissolved components in water confirms the aging of the LNAPL phase. These laboratory tests were
 364 performed on a sample of water previously partitioned with supernatant and jet fuel in equilibrium conditions (Fig. 8).



365

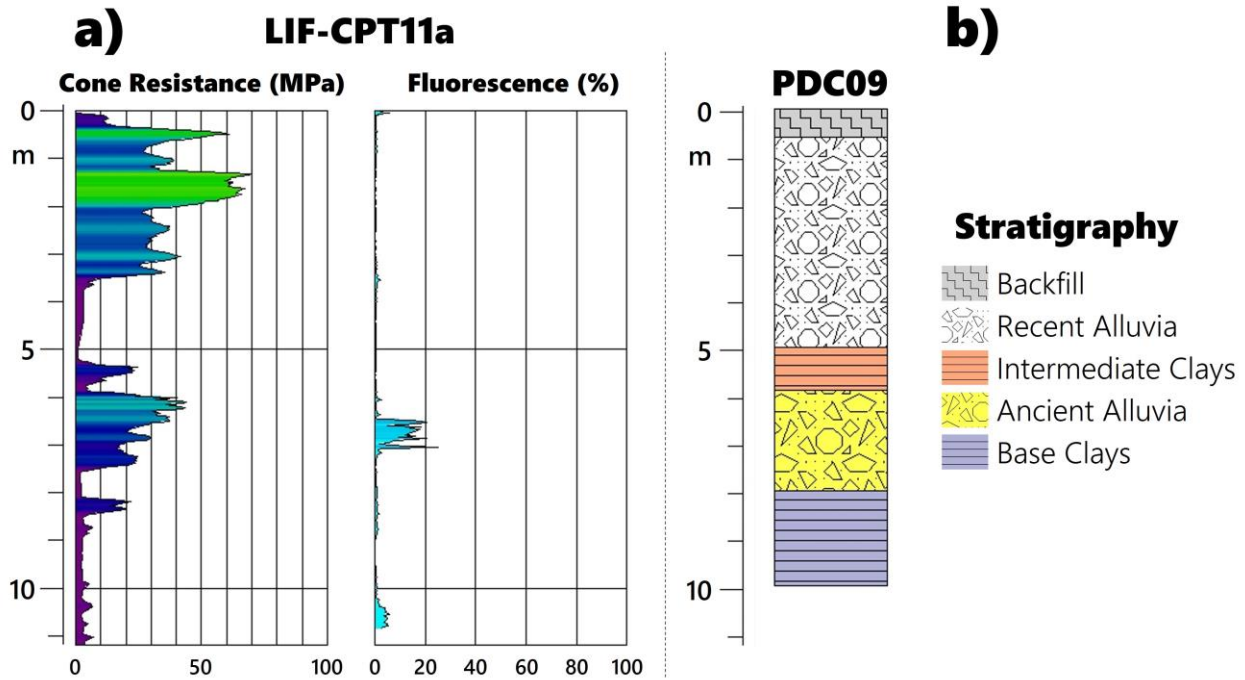
366 **Fig. 8** Solubilization of components in water for jet fuel (a) and the supernatant (b)

367 The analytical investigations for water samples in contact with the supernatant and the jet fuel exhibit a very low
 368 presence of light aromatic fractions in the first case. Such an aspect reveals the exhaustion of the soluble fraction in the
 369 recovered separated phase sample. The exhaustion of the more mobile and degradable components with the

370 accumulation of the heavier fraction is associated with the natural and progressive "aging" of the contamination source
371 (Lekmine et al. 2017). This source-aging scenario is highly representative of petroleum hydrocarbon pollution, since the
372 composition of these mixtures, which contain compounds distinguished by widely differing chemical/physical and
373 biodegradable characteristics, is highly complex (Vozka et al. 2019). The lighter, more soluble hydrocarbon fractions
374 (e.g., BTEX) are mobilized into groundwater in the initial step of the primary pollution event, and aerobic
375 biodegradation processes act on the more degradable components. The comparative analyses conducted on "fresh"
376 product (original jet fuel) and supernatant recovered during the monitoring campaigns unmasked the presence of a
377 weathered and aged product in the residual phase (Lekmine et al. 2017). This residual fraction, albeit still present in the
378 environmental matrices, is not able to release significant quantities of soluble substances into groundwater. This latter
379 aspect is confirmed by the total absence of the aromatic fraction in the supernatant. However, the residual, insoluble
380 fraction of higher molecular weight hydrocarbons persists in the primary source area. This is sporadically "mobilized"
381 and hence caught during dynamic sampling activities (Ciampi et al. 2021b).

382 *3.5 The Findings from LIF-UVOST and CPT Surveys*

383 The investigation through the LIF technique delineated the presence of the residual fraction of spilled fuel in the
384 subsurface. Qualitative calibration of direct push profiles with spatially adjacent stratigraphic logs validates the
385 lithotechnical interpretation (Gruiz et al. 2017). Although geologic model interpolation is not constrained by CPT data,
386 the overlap of such data differing in nature and resolution accounts for vertical heterogeneity. The extraction of such
387 data from the multi-source model captures lithotechnical parameterization and spatial variability of the stratigraphic
388 profile (Einarson et al. 2018; Pepper et al. 2002) (Fig. 9).



389

390 **Fig. 9** Resistance to penetration of the cone resulting from CPT and fluorescence signals detected by LIF-UVOST
 391 technology along with a vertical profile (a). Adjacent stratigraphic log representing the calibration borehole (b)

392 The LIF-CPT11a presented in Figure 9 reveals several peaks in a depth range between 6.48 m and 7.05 m, with an
 393 intensity reaching a maximum value of 25%. Some appreciable fluorescence signals, observed at depths between 10.42
 394 m and 10.82 m with an average intensity of 4%, suggest that the spilled product may have potentially and locally
 395 reached the basement clays. Also, figure 9 displays an example of the CPT response as a function of depth in presence
 396 of different lithologies, depicting a major difference between coarse-grained (recent and ancient alluvia) and fine-
 397 grained (intermediate and base clays) deposits. Coarse-grained deposits exhibit cone penetration resistance values
 398 generally between 20 and 60 Mpa. Intermediate and base clays always show cone penetration resistance values below
 399 10 MPa. The direct geotechnical investigations validate the geological and geophysical characterization, thus,
 400 improving, in general, the integrated multidisciplinary model. The geotechnical voxel-based model exhibits an excellent
 401 correlation between the low cone resistance bands and the levels ascribable to the intermediate clays in the Western
 402 area. Differently, in the Eastern portion, cone resistance increase is due to the presence of the gravel levels belonging to
 403 the ancient and recent alluvia, as presented in Fig. 10.

404

405 **Fig. 10** Two viewpoints of the 3D geotechnical model illustrating the resistance to penetration (MPa) of the CPT cone
406 arising from the LIF-CPT surveys and representation of the realized geological boreholes

407 The execution of 30 LIF-UVOST soundings permitted to identify with high vertical resolution the presence of
408 contaminants as free-phase droplets or adsorbed on the solid matrix. Accordingly, the LIF investigations delineate the
409 areas impacted by secondary and residual contamination and thus recognize the subsoil thickness affected by the
410 presence of aged product in the geological domain (Algreen et al. 2015), as illustrated in Fig. 11.

411

412 **Fig. 11** 3D model of fluorescence measured by LIF-UVOST probes in the geological framework of the site. The
413 executed stratigraphic boreholes are portrayed in the three-dimensional solid and multi-source picture

414

3.6 *The Joint Integration of Multi-Source Data Revealing the Contamination Dynamics*

415 Extracting the overlapped hydrogeophysical knowledge from the big-data package yields important information about
416 the contamination mechanisms which is not accessible without the proposed approach. In the primary source area, the
417 free phase contaminant is present both as oil droplets trapped in the pore space and adsorbed onto the solid matrix
418 (Trulli et al. 2016). Such contaminant is distributed across the so-called smear zone, often with a thickness of 4 meters
419 (Fig. 12).

420

421 **Fig. 12** Stratigraphic section, piezometric level, and fluorescence peaks detected along the track reported in the map
422 Relatively high fluorescence signals are measured in the LIF-UVOST15 survey at depths ranging from 5.58 to 6.77 m,
423 with a maximum peak (54% of fluorescence) at 6.02 m depth. The LIF-UVOST16 survey records moderate signals (a

424 maximum fluorescence peak of 10%), over a depth range of 3.76 to 5.57m. The fluorescence peaks unveil the presence
425 of residual free-phase/adsorbed hydrocarbons in the region surrounding the water table fluctuation range, which varies
426 between about 19 m and 14 m above sea level (Fig. 2 of Supplementary Material). Fig. 12 also reveals a potential
427 contribution of the extraction wells on the pollution dynamics. The depression of the piezometric surface due to the
428 pumping operated by the extraction wells and the seasonal oscillation of the water table favored the redistribution of the
429 product in the residual phase across the smear zone (as also observed in Trulli et al. 2016) as well as laterally. The
430 redistribution of the LNAPL along the smear zone was favored by the absence of intermediate clays in the area of the
431 storage tanks and the pumping wells. The intercalations of intermediate clays could have limited the vertical dispersion
432 of the contaminants caused by the changes in the water table. In the vicinity of the hydraulic barrier, the multi-source
433 model evidences hydraulic perturbation on contamination mechanisms. The pronounced cone of depression induced by
434 the intensive pumping favored redistribution of the aged product to the base of the aquifer and in the base clays (as
435 observed in Fig. 13).

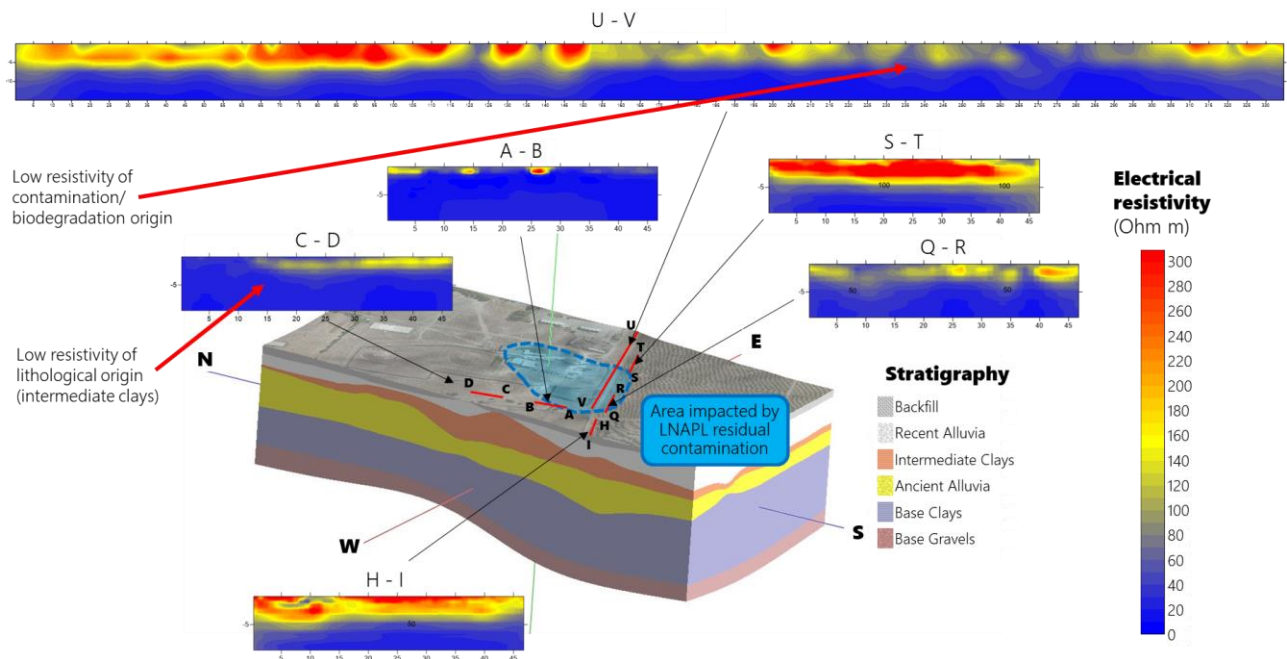
436

437 **Fig. 13** Comparison between the stratigraphic and the resistivity sections which are extracted along the track reported in
438 the map. The stratigraphy intersected by the wells, the measured piezometric level, and the recorded fluorescence peaks
439 are overlapped on the multi-modality profiles

440 Although the presence of a peak in the low-permeability layer bounding the aquifer is unexpected, calibrating the
441 stratigraphic profile via the ERT permits to improve the reconstruction of the aquifer conformation, by defining the
442 geometry of the basal clay shallow interface. In such an area with a high density of input data, the geological model
443 produced with an exact interpolator assumes the role of a training tool to develop spatial links between geological
444 properties and geophysical signals. Although the control boreholes lie 5 m away from the ERT line and the geological
445 model is affected by the correct interpretation of lithological data, geophysical findings reveal a substantial consistency
446 with the geological observations at known points as well as local and abrupt deepening of the low-permeability basal
447 layer where interpolation failed to delineate the undulating surface of this level between stratigraphic boreholes. At
448 borehole PB02 the resistivity section of Fig. 13 suggests a deeper contact of the base clays compared to the stratigraphic
449 profile. Such local deviation of the geologic data from the ERT image may be related to marked lateral geologic
450 variability over short distances or may delineate potential misinterpretations of the borehole data. The LIF detector
451 tracks two remarkable percent fluorescence peaks at depths of 5.67 m (24%) and 7.77 m (50%). Such signals disclose
452 the occurrence of aged product within the ancient alluvia and base clays. The overlap and interference of multiple radii
453 of influence for intensive pumping reduced hydraulic head, dewatering the aquifer horizon. A part of the LNAPL that
454 was originally mobile was smeared to the base of the aquifer and within the base clays due to piezometric surface
455 depression over time (Fig. 2 of Supplementary Material). Such residual LNAPL is adsorbed to the soil particles and
456 trapped into the pore of the saturated domain when the water table rises for aquifer recharge or recovery system
457 pumping is reduced, providing a persistent source of groundwater contamination.

458 **4. Discussion**

459 The joint use of point data coming from piezometric surveys, hydrogeochemical samplings, vertical profiling of
460 geotechnical and hydrocarbon presence (via LIF), and of spatially distributed data from ERT (i.e., geophysical) surveys
461 led to the construction of a comprehensive 3D conceptual model concerning both (a) the hydro-geophysical structure of
462 the site subsurface, and (b) the distribution of jet fuel contamination. This model is a tool through which the user can
463 analyze geospatial data, giving a rapid and intuitive way to access a vast amount of data. Such an approach has also
464 been discussed in other studies for different areas (Ciampi et al. 2019a; Harvey et al. 2017; Jones et al. 2009). One of
465 the main results obtained using the integrated geodatabase has been to provide evidence for an improved interpretation
466 of the ERT results based on physical information. In this regard, for our site, low electrical resistivity may be caused
467 either by lithological features (such as clayey formations) or by contamination effects (as a result of bio attenuation)
468 (Fig. 14).



469

470 **Fig. 14** Low resistivity anomalies at the Decimomannu site, caused by either (a) lithology (clays) and (b) contamination
 471 (via biodegradation)

472 Figure 14 shows the cross-analysis of ERT surveys and geo-stratigraphic reconstruction from borehole cores, which
 473 allows the identification of contaminated areas. Note that without such cross-analysis the interpretation of ERT results
 474 would be impossible, in particular, to distinguish between contaminants and clays as the cause of the low electrical
 475 resistivity values. While low resistivity caused by clay is a barrier to contamination spreading, low resistivity caused by
 476 biodegradation of petroleum hydrocarbons is a viable signal of contaminant presence (e.g., Cassiani et al. 2014). The
 477 multi-source CSM provides qualitative-quantitative indicators to reduce uncertainties associated with subsurface
 478 interpretation by separating the signatures of geologic material in the absence of LNAPL (Hermans and Irving, 2017)
 479 and the substantial increase in electrical conductivity caused by petroleum hydrocarbon biodegradation (Cassiani et al.
 480 2014). The conceptual model offers a window into in situ bioattenuation at the LNAPL-affected site and represents a
 481 tool for sharing robust evidence of microbiological activity to policymakers, who very often do not recognize natural
 482 attenuation (NA) as a remediation technology and oppose its application for limited information on natural attenuation
 483 processes (Declercq et al. 2012; Lari et al. 2019). The methods advocated in this paper could help promote a high
 484 degree of confidence and return of experience from the CSM, so NA technology could be seen as eligible by
 485 environmental authorities. Besides, the multi-source model led to the understanding of both the real pollutant
 486 characteristics and the contamination mechanisms depending on the hydraulic dynamics (Ciampi et al. 2021a).
 487 Laboratory tests proved necessary to verify the occurrence of aged product persisting in the residual phase (Lekmine et
 488 al. 2017; Vozka et al. 2019). The smearing of LNAPL caused by the water table fluctuation (Gatsios et al. 2018), the

489 entrapment of an "immobile" phase at the base of the aquifer linked to intensive pumping, and the detection of
490 lithotypes affected by the presence of residual product may be unveiled only by data fusion into a hydrogeophysical
491 clone (Ciampi et al. 2021a). Again, ERTs are crucial to delineate the geometry of the base clay shallow interface, which
492 is poorly solved by relying on borehole data exclusively. In the absence of sediment geoelectric signature spatialization,
493 the association of a fluorescence signal to the base clays would remain approximate and uncertain. The joint-modeling
494 findings identify the secondary contamination source that sporadically and slowly releases constituents into
495 groundwater as a result of both piezometric surface fluctuation and horizontal groundwater flow that can cross LNAPL
496 accumulations in the saturated aquifer (Gatsios et al. 2018; Lari et al. 2018). Such a conceptual reconstruction also
497 explains the sporadic presence of a TPH plume downstream of the barrier, coherently with the evidence derived by
498 Flores Orozco et al. (2021) via multi-frequency complex conductivity imaging. Future integration of complex
499 conductivity imaging data into the multi-source model may exploit the full potential of the method adopted in this study
500 by providing a quasi-continuous link between textural information, aquifer hydraulic properties, preferential plume
501 transport pathways, hydrocarbon concentrations, and biogeochemical transformations via quantitative interpretation of
502 electrical signatures of subsurface phenomena in addition to geologic contrast. Following the principles of Binley et al.
503 (2015), Jones et al. (2009), and Crook et al. (2008), the confluence of disparate types of hydrogeophysical geomodeling
504 develops a picture linking hydrologically relevant properties and measurable geophysical parameters of the
505 contaminants. The fusion of multiple data sources into the data-driven model is critical to understand the underlying
506 mechanisms that influence contamination dynamics (Ciampi et al. 2021a; Kueper et al. 2014). The fusion, exchange,
507 and extraction of knowledge from multi-source data pursue the concepts of Breunig et al. (2019), enhancing the
508 interoperability of multi-modal information and further advancing the utility of merged data to explain the contaminant-
509 physicochemical behavior and guide the design of a remediation strategy tailored to site-specific characteristics. In this
510 sense, Ciampi et al. (2021b) exploit the capabilities of the big-data package and conceptual model confined to the scale
511 of a pilot test for 4D time-lapse monitoring of decontamination dynamics induced by reagent injection in the source
512 area via two additional ERT profiles. Although the above study does not account for the LIF and ERT investigations
513 reported in this work, it unveils the potential performance of the data-driven model in handling end-of-process
514 remediation strategies, by interpreting the physicochemical modifications in space-time induced by the remediation
515 process at the field scale and revealing the mobilization of the immobile material constituting the residual phase of
516 hydrocarbons.

517 **5. Conclusions**

518 The 3D hydrogeophysical model exploits information from different sources to discretize the different causes
519 influencing the measured physicochemical properties, differentiating the signature of geologic features from the
520 contamination effects and explaining pollution dynamics in space-time. GC-MS analyses unveil a source-aging scenario
521 of petroleum hydrocarbon contamination while LIF investigations delineate the subsoil volume impacted by the
522 presence of residual spilled fuel fraction. Incorporating electrical models from geophysical surveys into the
523 hydrogeochemical model surmounts the limitations of spatial aliasing associated with conventional geological
524 investigations and permits to improve the geophysical interpretation. In particular, our approach allows us to
525 discriminate low conductivity values related to clay layers and due to aging hydrocarbon contaminants. At a
526 contaminated site subject to remediation action through groundwater extraction wells, bridging such a gap and
527 capturing the spatial variations of the data permits to understand the pollution mechanisms within the geological and
528 hydraulic framework. The integrated analysis and joint data modeling approach unmask the LNAPL weathering and
529 reveal both the trapping of residual phase hydrocarbons across the smear zone and locally to the aquifer base, due to
530 water table fluctuation and hydraulic perturbations triggered by extraction wells. On the one hand, the redistribution and
531 sequestration of aged contaminants in the separated phase by hydraulic processes is in agreement with the geological
532 units and the presence of low permeability layers. On the other hand, LNAPL aging reduces the mobility of pollutants
533 both trapped in pore spaces and adsorbed onto the solid matrix. The geodatabase-driven and multi-modality portrayal
534 emphasize the need for a large amount of multi-source data to build a reliable and high-resolution conceptual model, an
535 indispensable prerequisite for planning an effective remediation strategy.

536 **Funding**

537 This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit
538 sectors.

539 **Acknowledgments**

540 We gratefully acknowledge the Italian Air Force for their support and collaboration during the data acquisition
541 investigation and stages. We also thank the reviewers for their comments that helped to improve the quality of this
542 manuscript.

543 **References**

544 Abbaspour, K., Matta, V., Huggenberger, P., Johnson, C.A., 2000. A contaminated site investigation: comparison of
545 information gained from geophysical measurements and hydrogeological modeling. *J. Contam. Hydrol.* 40, 365-380.
546 [https://doi.org/10.1016/S0169-7722\(99\)00055-8](https://doi.org/10.1016/S0169-7722(99)00055-8).

547 Algreen, M., Kalisz, M., Stalder, M., Martac, E., Krupanek, J., Trapp, S., Bartke, S., 2015. Using pre-screening methods
548 for an effective and reliable site characterization at megasites. *Environ. Sci. Pollut. Res.* 22, 14673–14686.
549 <https://doi.org/10.1007/s11356-015-4649-6>.

550 Arato, A., Wehrer, M., Biró, B., Godio, A., 2014. Integration of geophysical, geochemical and microbiological data for
551 a comprehensive small-scale characterization of an aged LNAPL-contaminated site. *Environ. Sci. Pollut. Res.* 21, 8948-
552 8963. <https://doi.org/10.1007/s11356-013-2171-2>.

553 Åslund, M.W., Stephenson, G.L., Simpson, A.J., Simpson, M.J., 2013. Comparison of earthworm responses to
554 petroleum hydrocarbon exposure in aged field contaminated soil using traditional ecotoxicity endpoints and 1H NMR-
555 based metabolomics. *Environ. Pollut.* 182, 263-268. <https://doi.org/10.1016/j.envpol.2013.07.026>.

556 Bini, C., 2013. Geology and Geomorphology, in: Costantini, E., Dazzi, C. (Eds), *The Soils of Italy*. World Soils Book
557 Series. Springer, Dordrecht, pp. 39-56. https://doi.org/10.1007/978-94-007-5642-7_3.

558 Binley, A., Cassiani, G., Deiana, R., 2010. Hydrogeophysics: opportunities and challenges. *Boll. Geofis. Teor. Appl.*
559 51, 267-284.

560 Binley, A., Hubbard, S.S., Huisman, J.A., Revil, A., Robinson, D.A., Singha, K., Slater, L.D., 2015. The emergence of
561 hydrogeophysics for improved understanding of subsurface processes over multiple scales. *Water Resour. Res.* 51,
562 3837-3866. doi:10.1002/2015WR017016.

563 Binley, A., Kemna, A., 2005. DC Resistivity and Induced Polarization Methods, in: Rubin, Y., Hubbard, S.S. (Eds.),
564 *Hydrogeophysics*. Water Science and Technology Library, vol 50. Springer, Dordrecht, pp. 129-156.
565 https://doi.org/10.1007/1-4020-3102-5_5.

566 Binley, A., Ramirez, A., Daily, W., 1995. Regularised image reconstruction of noisy electrical resistance tomography
567 data, in: Beck, M.S., Hoyle, B.S., Morris, M.A., Waterfall, R.C., Williams, R.A. (Eds.), *Process Tomography*.
568 *Proceedings of the 4th Workshop of the European Concerted Action on Process Tomography*, Bergen, 6-8 April 1995,
569 pp. 401–410.

570 Blanchy, G., Saneiyani, S., Boyd, J., McLachlan, P., Binley, A., 2020. ResIPy, an intuitive open source software for
571 complex geoelectrical inversion/modeling. *Comput. Geosci.* 137, 104423. <https://doi.org/10.1016/j.cageo.2020.104423>.

572 Breunig, M., Bradley, P.E., Jahn, M., Kuper, P., Mazroob, N., Rösch, N., Al-Doori, M., Stefanakis, E., Jadidi, M., 2020.
573 *Geospatial Data Management Research: Progress and Future Directions*. *ISPRS Int. J. Geo-Inf.* 9(2), 95.
574 <https://doi.org/10.3390/ijgi9020095>.

575 Brusseau, M.L., 2019. Soil and Groundwater Remediation, in: Brusseau, M.L., Pepper, I.L., Gerba, C.P. (Eds.),
576 Environmental and Pollution Science. Academic Press, New York, pp. 329–35. doi:10.1016/B978-0-12-814719-
577 1.00019-7.

578 Bücken, M., Flores Orozco, A., Hördt, A. Kemna, A., 2017. An analytical membrane-polarization model to predict the
579 complex conductivity signature of immiscible liquid hydrocarbon contaminants. *Near Surf. Geophys.* 15(6), 547-562.
580 <https://doi.org/10.3997/1873-0604.2017051>.

581 Cassiani, G., Bruno, V., Villa, A., Fusi, N., Binley, A.M., 2006. A saline tracer test monitored via time-lapse surface
582 electrical resistivity tomography, *J. Appl. Geophys.* 59, 244-259. <https://doi.org/10.1016/j.jappgeo.2005.10.007>.

583 Cassiani, G., Binley, A., Kemna, A., Wehrer, M., Flores Orozco, A., Deiana, R., Boaga, J., Rossi, M., Dietrich, P.,
584 Werban, U., Zschornack, L., Godio, A., Jafargandomi, A., Deidda, G.P., 2014. Noninvasive characterization of the
585 Trecate (Italy) crude-oil contaminated site: links between contamination and geophysical signals. *Environ. Sci. Pollut.*
586 *Res.* 21, 8914–8931. DOI 10.1007/s11356-014-2494-7.

587 Caterina, D., Flores Orozco, A. Nguyen, F., 2017. Long-term ERT monitoring of biogeochemical changes of an aged
588 hydrocarbon contamination. *J. Contam. Hydrol.* 201, 19-29. <https://doi.org/10.1016/j.jconhyd.2017.04.003>.

589 Chambers, J.E., Wilkinson, P.B., Wealthall, G.P., Loke, M.H., Dearden, R., Wilson, R., Allen, D., Ogilvy, R.D., 2010.
590 Hydrogeophysical imaging of deposit heterogeneity and groundwater chemistry changes during DNAPL source zone
591 bioremediation. *J. Contam. Hydrol.* 118. <https://doi.org/10.1016/j.jconhyd.2010.07.001>.

592 Ciampi, P., Esposito, C., Petrangeli Papini, M., 2019a. Hydrogeochemical Model Supporting the Remediation Strategy
593 of a Highly Contaminated Industrial Site. *Water* 11, 1371. <https://doi.org/10.3390/w11071371>.

594 Ciampi, P., Esposito, C., Viotti, P., Boaga, J., Cassiani, G., Petrangeli Papini, M., 2019b. An Integrated Approach
595 Supporting Remediation of an Aquifer Contaminated with Chlorinated Solvents by a Combination of Adsorption and
596 Biodegradation. *Appl. Sci.* 9(20), 4318. <https://doi.org/10.3390/app9204318>.

597 Ciampi, P., Esposito, C., Bartsch, E., Alesi, E.J., Petrangeli Papini, M., 2021a. 3D dynamic model empowering the
598 knowledge of the decontamination mechanisms and controlling the complex remediation strategy of a contaminated
599 industrial site. *Sci. Total Environ.* 793, 148649. <https://doi.org/10.1016/j.scitotenv.2021.148649>.

600 Ciampi, P., Esposito, C., Cassiani, G., Deidda, G.P., Rizzetto, P., Petrangeli Papini, M., 2021b. A field-scale
601 remediation of residual light non-aqueous phase liquid (LNAPL): chemical enhancers for pump and treat. *Environ. Sci.*
602 *Pollut. Res.* 28, 35286–35296. <https://doi.org/10.1007/s11356-021-14558-2>.

603 Crook, N., Binley, A., Knight, R., Robinson, D.A., Zarnetske, J., Haggerty, R., 2008. Electrical resistivity imaging of
604 the architecture of substream sediments. *Water Resour. Res.* 44, W00D13. <https://doi.org/10.1029/2008WR006968>.

605 Declercq, I., Cappuyns, V., Duclos, Y., 2012. Monitored natural attenuation (MNA) of contaminated soils: State of the
606 art in Europe-A critical evaluation. *Sci. Total Environ.* 426, 393-405. <https://doi.org/10.1016/j.scitotenv.2012.03.040>.

607 Deiana, R., Cassiani, G., Kemna, A., Villa, A., Bruno, V., Bagliani, A., 2007. An experiment of non-invasive
608 characterization of the vadose zone via water injection and cross-hole time-lapse geophysical monitoring. *Near. Surf.*
609 *Geophys.* 5(3), 183-194.

610 Einarson, M., Fure, A., Germain, R., Chapman, S., Parker, B., 2018. DyeLIFT™: A New Direct-Push Laser-Induced
611 Fluorescence Sensor System for Chlorinated Solvent DNAPL and Other Non-Naturally Fluorescing NAPLs. *Ground*
612 *Water Monit. Remediat.* 38, 28-42. doi:10.1111/gwmmr.12296.

613 Falivene, O., Cabrera, L., Tolosana-Delgado, R., Alberto, S., 2010. Interpolation algorithm ranking using cross-
614 validation and the role of smoothing effect. A coal zone example. *Comput. Geosci.* 36 (4), 512-519.
615 <https://doi.org/10.1016/j.cageo.2009.09.015>.

616 Fedotov, Y., Belov, M., Kravtsov, D., Gorodnichev, V., 2019. Laser fluorescence method for detecting oil pipeline
617 leaks at a wavelength of 355 nm. *J. Opt. Technol.* 86, 81-85. <https://doi.org/10.1364/JOT.86.000081>.

618 Flores Orozco, A., Kemna, A., Oberdörster, C., Zschornack, L., Leven, C., Dietrich, P., Weiss, H., 2012, Delineation of
619 subsurface hydrocarbon contamination at a former hydrogenation plant using spectral induced polarization imaging. *J.*
620 *Contam. Hydrol.* 136-137, 131-144. <https://doi.org/10.1016/j.jconhyd.2012.06.001>.

621 Flores Orozco, A., Velimirovic, M., Tosco, T., Kemna, A., Sapion, H., Klaas, N., Sethi, R., Leen, B. 2015. Monitoring
622 the injection of microscale zero-valent iron particles for groundwater remediation by means of complex electrical
623 conductivity imaging. *Environ. Sci. Technol.* 49, 5593–5600. <https://doi.org/10.1021/acs.est.5b00208>.

624 Flores Orozco, A., Kemna, A., Binley, A., Cassiani, G., 2019a. Analysis of time-lapse data error in complex
625 conductivity imaging to alleviate anthropogenic noise for site characterization. *Geophysics* 84(2), B181-B193.
626 <https://doi.org/10.1190/geo2017-0755.1>.

627 Flores Orozco, A., Micić, V., Bücker, M., Gallistl, J., Hofmann, T., Nguyen, F., 2019. Complex-conductivity
628 monitoring to delineate aquifer pore clogging during nanoparticles injection. *Geophys. J. Int.* 218(3), 1838-1852.
629 <https://doi.org/10.1093/gji/ggz255>.

630 Flores Orozco, A., Ciampi, P., Katona, T., Censini, M., Petrangeli Papini, M., Deidda, G.P., Cassiani, G., 2021.
631 Delineation of hydrocarbon contaminants with multi-frequency complex conductivity imaging. *Sci. Total Environ.*
632 768,144997. <https://doi.org/10.1016/j.scitotenv.2021.144997>.

633 Gatsios, E., García-Rincón, J., Rayner, J.L., McLaughlan, R.G., Davis, G.B., 2018. LNAPL transmissivity as a
634 remediation metric in complex sites under water table fluctuations. *J. Environ. Manage.* 215, 40-48.
635 <https://doi.org/10.1016/j.jenvman.2018.03.026>.

636 Ghosh, J., Tick, G.R., Akyol, N.H., Zhang, Y., 2019. A pore-scale investigation of heavy crude oil trapping and
637 removal during surfactant-enhanced remediation. *J. Contam. Hydrol.* 223, 103471.
638 <https://doi.org/10.1016/j.jconhyd.2019.03.003>.

639 Gruiz, K., Fenyvesi, E., Molnár, M., Feigl, V., Vaszita, E., Tolner, M., 2017. In-situ and real-time measurements for
640 effective soil and contaminated site management, in: Gruiz, K., Meggyes, T., Fenyvesi, E. (Eds.), *Engineering Tools for*
641 *Environmental Risk Management. Site Assessment and Monitoring Tools*. Taylor, Francis Group, CRC Press, London,
642 pp. 245-341. <https://doi.org/10.1201/b19954>.

643 Harris, M.K., Looney, B.B., Jackson, D.G., 2004. *Geology and environmental remediation: Savannah River Site, South*
644 *Carolina*. *Environ. Geosci.* 11, 191-204. <https://doi.org/10.1306/eg.06150404015>.

645 Harvey, A.S., Fotopoulos, G., Hall, B., Amolins, K., 2017. Augmenting comprehension of geological relationships by
646 integrating 3D laser scanned hand samples within a GIS environment. *Comput. Geosci.* 103, 152-163.
647 <https://doi.org/10.1016/j.cageo.2017.02.008>.

648 Hermans, T., Irving, J., 2017. Facies discrimination with electrical resistivity tomography using a probabilistic
649 methodology: Effect of sensitivity and regularization. *Near Surf. Geophys.* 15, 13-25. [https://doi.org/10.3997/1873-](https://doi.org/10.3997/1873-0604.2016047)
650 [0604.2016047](https://doi.org/10.3997/1873-0604.2016047).

651 Høyer, A.S., Jørgensen, F., Sandersen, P.B.E., Viezzoli, A., Møller, I., 2015. 3D geological modelling of a complex
652 buried-valley network delineated from borehole and AEM data. *J. Appl. Geophys.* 122, 94-102.
653 <https://doi.org/10.1016/j.jappgeo.2015.09.004>.

654 Huysegoms, L., Cappuyns, V., 2017. Critical review of decision support tools for sustainability assessment of site
655 remediation options. *J. Environ. Manage.* 196, 278–296. doi:10.1016/j.jenvman.2017.03.002.

656 Jones, R.R., Mccaffrey, K.J.W., Clegg, P., Wilson, R.W., Holliman, N.S., Holdsworth, R.E., Imber, J., Waggott, S.,
657 2009. Integration of regional to outcrop digital data: 3D visualisation of multi-scale geological models. *Comput. Geosci.*
658 35, 4-18. <https://doi.org/10.1016/j.cageo.2007.09.007>.

659 Kaliraj, S., Chandrasekar, N., Peter, T.S., Selvakumar, S., Magesh, N.S., 2015. Mapping of coastal aquifer vulnerable
660 zone in the south west coast of Kanyakumari, South India, using GIS-based DRASTIC model. *Environ. Monit. Assess.*
661 187, 4073. <https://doi.org/10.1007/s10661-014-4073-2>.

662 Lari, K.S., Rayner, J.L., Davis, G.B., 2018. Towards characterizing LNAPL remediation endpoints. *J. Environ. Manage.*
663 224, 97-105. <https://doi.org/10.1016/j.jenvman.2018.07.041>.

664 Lari, K.S., Davis, G.B., Rayner, J.L., Bastow, T.P., Puzon, G.J., 2019. Natural source zone depletion of LNAPL: A
665 critical review supporting modelling approaches. *Water Res.* 157, 630-646.
666 <https://doi.org/10.1016/j.watres.2019.04.001>.

667 Lekmine, G., Lari, K.S., Johnston, C.D., Bastow, T.P., Rayner, J.L., Davis, G.B., 2017. Evaluating the reliability of
668 equilibrium dissolution assumption from residual gasoline in contact with water saturated sands. *J. Contam. Hydrol.*
669 196, 30-42. <https://doi.org/10.1016/j.jconhyd.2016.12.003>.

670 Liu, H., Chen, S., Hou, M., Liang, H., 2020. Improved inverse distance weighting method application considering
671 spatial autocorrelation in 3D geological modeling. *Earth Sci. Inform.* 13, 619-632. [https://doi.org/10.1007/s12145-019-](https://doi.org/10.1007/s12145-019-00436-6)
672 00436-6.

673 Mccall, W., Christy, T.M., Pipp, D.A., Jaster, B., White, J., Goodrich, J., Fontana, J., Doxtader, S., 2018. Evaluation
674 and application of the optical image profiler (OIP) a direct push probe for photo-logging UV-induced fluorescence of
675 petroleum hydrocarbons. *Environ. Earth. Sci.* 77, 374. <https://doi.org/10.1007/s12665-018-7442-2>.

676 Mirzaei, R., Sakizadeh, M., 2016. Comparison of interpolation methods for the estimation of groundwater
677 contamination in Andimeshk-Shush Plain, Southwest of Iran. *Environ. Sci. Pollut. Res.* 23, 2758-2769.
678 <https://doi.org/10.1007/s11356-015-5507->.

679 Pepper, J.W., Wright, A.O., Kenny, J.E., 2002. In situ measurements of subsurface contaminants with a multi-channel
680 laser-induced fluorescence system. *Spectrochim. Acta A* 58, 317-331. [https://doi.org/10.1016/S1386-1425\(01\)00557-1](https://doi.org/10.1016/S1386-1425(01)00557-1).

681 Pollard, S.J.T., Brookes, A., Earl, N., Lowe, J., Kearney, T., Nathanail, C.P., 2004. Integrating decision tools for the
682 sustainable management of land contamination. *Sci. Total Environ.* 325, 15-28.
683 <https://doi.org/10.1016/j.scitotenv.2003.11.017>.

684 Prasanna, M.V., Chidambaram, S., Pethaperumal, S., Srinivasamoorthy, K., Peter, A.J., Anandhan, P., Vasanthavigar,
685 M., 2008. Integrated geophysical and chemical study in the lower subbasin of Gadilam River, Tamilnadu, India.
686 *Environ. Geosci.* 15, 145-152. <https://doi.org/10.1306/eg.11080707015>.

687 Reuter, M., Auer, G., Brandano, M., Harzhauser, M., Corda, L., Piller, W.E., 2017. Post-rift sequence architecture and
688 stratigraphy in the Oligo–Miocene Sardinia Rift (Western Mediterranean Sea). *Mar. Pet. Geol.* 79, 44-63.
689 <https://doi.org/10.1016/j.marpetgeo.2016.10.025>.

690 Revil, A., Coperey, A., Shao, Z., Florsch, N., Fabricius, I.L., Deng, Y., Delsman, J.R., Pauw, P.S., Karaoulis, M., de
691 Louw, P.G.B., van Baaren, E.S., Dabekaussen, W., Menkovic, A., Gunnink, J.L., 2017. Complex conductivity of soils.
692 *Water Resour. Res.* 53(8), 7121-7147. <https://doi.org/10.1002/2017WR020655>.

693 Ruggeri, P., Gloaguen, E., Lefebvre, R., Irving, J., Holliger, K., 2014. Integration of hydrological and geophysical data
694 beyond the local scale: Application of Bayesian sequential simulation to field data from the Saint-Lambert-de-Lauzon
695 site, Québec, Canada. *J. Hydrol.* 514, 271 - 280.

696 Safarbeiranvand, M., Amanipoor, H., Battaleb-Looie, S., Ghanemi, K., Ebrahimi, B., 2018. Quality Evaluation of
697 Groundwater Resources using Geostatistical Methods (Case Study: Central Lorestan Plain, Iran). *Water Resour. Manag.*
698 32, 3611–3628. <https://doi.org/10.1007/s11269-018-2009-2>.

699 Samouelian, A., Cousin, I., Tabbagh, A., Bruand, A., Richard, G., 2005. Electrical resistivity survey in soil science: a
700 review. *Soil Tillage Res.* 83, 173-193. <https://doi.org/10.1016/j.still.2004.10.004>.

701 Shannon, C.E., 1949. Communication in the presence of noise, *Proceedings of the Institute of Radio Engineers* 37(1),
702 10-12.

703 Shao, S., Guo, X., Gao, C. 2019. Fresh underground light non-aqueous liquid (LNAPL) pollution source zone
704 monitoring in an outdoor experiment using cross-hole electrical resistivity tomography. *Environ. Sci. Pollut. Res.* 26,
705 18316–18328. <https://doi.org/10.1007/s11356-019-05039-8>.

706 Suthersan, S.S., Horst, J., Schnobrich, M., Welty, N., Mcdonough, J., 2016. *Remediation Engineering, Design*
707 *Concepts*, 2nd edn. CRC Press, Taylor, Francis Group, Boca Raton, pp. 107-135.

708 Teramoto, E.H., Isler, E., Polese, L., Baessa, M.P.M., Chang, H.K., 2019. LNAPL saturation derived from laser induced
709 fluorescence method. *Sci. Total Environ.* 683, 762-772. <https://doi.org/10.1016/j.scitotenv.2019.05.262>.

710 Totsche, K.U., Kogel-Knabner, I., Haas, B., Scheibke, R., Geisen, S., 2003. Evidence for preferential flow and ageing
711 of NAPL and NAPL borne contaminants in the unsaturated soil zone of a creosote contaminated site: a field study. *J.*
712 *Plant Nutr. Soil Sci.* 166, 102-110.

713 Tran, H., Vu, H.C., Lin, C., Bui, X., Huang, W., Vo, T., Hoang, H., Liu, W., 2018. Remediation of highly fuel oil-
714 contaminated soil by food waste composting and its volatile organic compound (VOC) emission. *Bioresour. Technol.*
715 *Rep.* 4, 145-152. <https://doi.org/10.1016/j.biteb.2018.10.010>.

716 Trulli, E., Morosini, C., Rada, E.C., Torretta, V., 2016. Remediation in Situ of Hydrocarbons by Combined Treatment
717 in a Contaminated Alluvial Soil due to an Accidental Spill of LNAPL. *Sustainability* 8, 1086.

718 Vasudevan, M., Johnston, C.D., Bastow, T.P., Lekmine, G., Rayner, J.L., Nambi, I.M., Suresh Kumar, G., Ravi
719 Krishna, R., Davis, G.B., 2016. Effect of compositional heterogeneity on dissolution of non-ideal LNAPL mixtures. *J.*
720 *Contam. Hydrol.* 194, 10-16. <https://doi.org/10.1016/j.jconhyd.2016.09.006>.

721 Vozka, P., Modereger, B.A., Park, A.C., Zhang, W., Trice, R.W., Kenttämä, H.I., Kilaz, G., 2019. Jet fuel density via
722 GC- \times -GC-FID. *Fuel* 235, 1052-1060. <https://doi.org/10.1016/j.fuel.2018.08.110>.

723 Wang, W., Huang, L., 2012. 3D geological modeling for mineral resource assessment of the Tongshan Cu deposit,
724 Heilongjiang Province, China. *Geosci. Front.* 3 (4), 483-491. <https://doi.org/10.1016/j.gsf.2011.12.012>.

725 Xia, T., Dong, Y., Mao, D., Meng, J., 2021. Delineation of LNAPL contaminant plumes at a former perfumery plant
726 using electrical resistivity tomography. *Hydrogeol. J.* 29, 1189–1201. <https://doi.org/10.1007/s10040-021-02311-5>.