Delineation of LNAPL Plumes in a Clay-Rich Site in Gyeongsangnam-do Province, South Korea, using Geophysical Surveys and Combined Interpretation with Borehole and Soil Sampling Data

Bitnarae Kim
Sejong University

Inseok Jeong
Sejong University

Huieun Yu
Sejong University

Juyeon Jeong
Sejong University

Seo Young Song
Sejong University

Jeong-sul Son
Sejong University

Youngchul Yu
KOTAM

Jehyun Shin
Korea Institute of Geoscience and Mineral Resources

Ho Young Jo
Korea University

Man Jae Kwon
Korea University

Myung Jin Nam (nmj1203@gmail.com)
Sejong University

Research Article

Keywords: Hydrocarbon, Contaminant plume, Geophysical methods, ERT, GPR, IP

Posted Date: June 1st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2992514/v1
Abstract

To effectively delineate the spatial distribution of oil contaminant plumes, geophysical methods indirectly measure the physical properties of the subsurface and can provide spatial information and images on a large scale, as opposed to traditional direct methods such as borehole drilling, sampling, and chemical analysis, which are time-consuming and costly. However, delineating geophysical responses from non-aqueous phase liquids (NAPLs) contaminated sites is not straightforward due to inconsistent responses from biodegraded oil contaminants. Additionally, the presence of clay materials can complicate the interpretation of geophysical data in NAPL-contaminated sites. In this study, we present a case study of a multi-geophysical investigation, including seismic refraction, ground-penetrating radar (GPR), electrical resistivity tomography (ERT), and complex resistivity (CR), to delineate NAPL contamination in a clay-rich site. To reduce ambiguity in discriminating between oil contaminants and clay layers, we suggest constructing a 3D geological model that incorporates borehole data and geophysical data. Based on the 3D geological model, conductive zones generally correspond to high concentrations of hydrocarbons in the unsaturated zone, but it is difficult to distinguish contaminated areas from saturated soil. The IP response rapidly decreased to close to zero in several expected highly contaminated zones, which differs from the clay soil with high IP values. Finally, we compare the expected contaminated area from geophysical data and soil sampling data and discuss how geophysical interpretation can be improved in NAPL-contaminated sites.

1. Introduction

Non-aqueous phase liquids (NAPLs), such as gasoline or industrial chlorinated solvents, are common environmental contaminants in near-surface groundwater. These chemicals are immiscible with water and exist as a separate phase in the soil, often referred to as the ‘free phase’. NAPLs can be categorized into two types: light NAPLs (LNAPLs), which are less dense than water and marginally soluble, and dense NAPLs (DNAPLs), which have a higher density than water. Therefore, LNAPL spills infiltrate the soil and initially rest on the water table. However, LNAPLs can reach deeper zones following fluctuations in groundwater level over time. Due to the lack of ionic or polar groups, ‘non-polar’ compounds, such as kerosene, form discrete droplets within the water-filled pores without direct contact with the grain surface, making NAPLs generally non-wetting, unless strong fractions of organic matter are present as a solid phase in the soil.

Conventional techniques for characterizing contaminated sites are based on direct methods such as drilling, sampling, and chemical analyses. Although necessary and informative, these techniques are time-consuming, cost-intensive, and often cannot adequately cover the full spatial extent of the contaminated site. Geophysical methods that analyze geological characteristics indirectly can play a prominent role in delineating contaminated sites at the near-surface level and over a broad scale. Ground-penetrating radar (GPR), electrical resistivity tomography (ERT), and induced polarization (IP) have been effectively applied to environmental sites where contaminants cause electrical changes (e.g., Annan,
Understanding the electrical properties of NAPL-contaminated plumes is challenging, as these properties can change both microscopically and macroscopically over time (Atekwana and Atekwana, 2010). From a microscopic perspective, NAPLs exhibit highly resistive and non-polarizable electrical properties in their initial state (Atekwana and Atekwana, 2010). However, hydrocarbon compounds feed microbes, leading to complex changes in the subsurface (Atekwana and Atekwana, 2010; Kessouri et al., 2019; Kimak et al., 2019). Activated microbes can produce electrical effects, not only from expanding bacterial cells but also from biofilm, which results in IP responses due to processes such as bio-clogging (Davis et al., 2006; Ntarlagiannis and Ferguson, 2009; Wu et al., 2014). Additionally, the byproducts of the biodegradation process can chemically and physically erode the grain boundaries of porous media and increase fluid conductivity, leading to changes in electrical resistivity (e.g., Cassidy et al., 2001; Werkema Jr et al., 2003; Atekwana et al., 2004). From a macroscopic perspective, leaked oil contamination follows the flow of groundwater and often fluctuates in response to changes in groundwater levels (Cavelan et al., 2022). Microbes are more active in disseminated areas than in the source zone, and the strongest geophysical signals caused by contamination and biodegradation occur in the upper portion of saturation zones and the capillary fringe (Cassiani et al., 2014). These factors render the electrical response from NAPL contaminants inconsistent in the field.

The presence of clay in a medium makes it difficult to analyze geophysical responses in field sites (Maurya et al., 2018; Shao et al., 2019; Flores Orozco et al., 2021). Clay-rich sediments have low electrical resistivity values due to conduction mechanisms taking place at the negatively charged surface of clay minerals (Revil and Glover, 1998; Gallistl et al., 2018). IP methods are an extension of ERT technique and have been applied to various fields, including environmental sites, to address the challenges of interpreting ERT results and discriminating clay content or other conductive targets. However, high clay content is typically associated with conductive and chargeable anomalies, which can also be observed in biodegraded oil contaminants in certain conditions (e.g., Kemna et al., 2004; Flores Orozco et al., 2021).

Although the IP method can provide additional information for NAPL contamination sites, delineating oil plume distribution is very complicated due to complex IP mechanisms and inconsistent IP responses in NAPL-contaminated field sites (Sogade et al., 2006; Attwa and Günther, 2012). Published cases have reported IP data from hydrocarbon contaminants with high values in some sites (Kemna et al., 2004; Deceuster and Kaufmann, 2012) but no IP response in other sites (Blondel et al., 2014; Maurya et al., 2018). Johansson et al. (2015) suggested that the source zone has no IP anomalies, but IP responses increased next to the source zone, which is interpreted as a degradation zone. Flores Orozco et al. (2012) showed that IP responses slightly increased with increasing benzene concentration but rapidly decreased to zero values above a concentration of 1.7 g/L, indicating the presence of a free-phase plume. Therefore, the geophysical responses of hydrocarbon contaminants associated with biological processes are highly site-dependent.
In previous studies, geophysical methods have demonstrated promising potential for delineating contaminants at environmental sites, though each method exhibits distinct subsurface characteristics. While there are limited instances of various geophysical investigation methods being applied at the same NAPL-contaminated site, cases utilizing ERT, GPR, and borehole investigations in combined analyses have been reported (e.g., Cassiani et al., 2014; Shao et al., 2019). However, there is a lack of case studies comparing IP data responses with other geophysical responses and performing comprehensive interpretations in NAPL contamination sites. Additionally, integrating geophysical data with direct methods, such as borehole data and sampling data, is particularly challenging in clay-rich geological environments where electrical responses are more complex. Therefore, there is a need for a comprehensive discussion and comparison of various survey methods in NAPL-contaminated sites.

The primary objective in this study is to evaluate the geophysical response of an LNAPL-contaminated site and discuss the challenges of delineating contaminated areas in a clay-rich environment in South Korea using geophysical data. To analyze the geophysical response of oil contamination, we conducted multi-geophysical investigations, including seismic refraction survey, GPR survey, ERT, and complex resistivity (CR) survey at the site. The geophysical responses were analyzed without prior information to assess the applicability of geophysical data in the presence of clay-rich soil layers.

The second objective is to implement an integrated analysis combining geophysical data with other direct methods, such as boreholes and sampling data. To differentiate the geophysical responses between contaminants and the background medium, a 3D geological model of the site was constructed using borehole and geophysical data. Employing this geological model, the anticipated contaminated area was analyzed using geophysical data in comparison with geochemical data. Owing to the availability of dense and abundant soil sampling data, which is relatively uncommon for other environmental sites, it was possible to compare the delineated results from geophysical data with selected sets of geochemical data. Finally, the LNAPL contaminant plume within the site was delineated using geophysical data in conjunction with comprehensive borehole and soil sampling data.

2. Site Overview

The target site, which locates in Gyeongsangnam-do Province, South Korea and named as Site Y, had oil storage tanks that held approximately 50,000 barrels since the 1970s (Fig. 1). The significant oil leakage was first detected through broad-scale geochemical surveys in 2012 and 2017, and it has had a considerable impact on the porous media at the site through the effects of geochemistry and biogeophysics (KEITI, 2020). To delineate the oil contaminated region, we conducted surface geophysical and borehole surveys in the most severely contaminated region among the sites.

Site Y is characterized by a thick sequence of silty and clay sands. A backfill layer of sandy gravel material, with a thickness of about 2 m, covers most of the site. Beneath this layer lies clayey and silty sand, extending from 3.5 to 7 m. Weathered soil and rock overlie the bedrock, which is composed of sandstone and claystone, at a depth of approximately 14 m. The site is surrounded by hills, and a small
stream located in the northern region of the survey site flows from northwest to the east side. The primary groundwater flow direction is from the southern hills to the northern stream.

Soil sampling tests for the oil storage site, conducted in 2012 and 2017, revealed that most soil layers and groundwater were contaminated. Among the entire site area of 150,860 m$^2$, approximately 29,660 m$^2$ are contaminated, mainly around depths of 2 m to 4 m (KEITI, 2020). The LNAPL contaminations flow from the surrounding hills to the northern stream in a direction similar to that of the groundwater.

3. Borehole and Soil Sampling Data

3.1 Borehole Data

The borehole survey and soil sampling were conducted in previous works (KEITI, 2020) (Fig. 1c). A total of 7 boreholes (BH1-7) were drilled with 100 mm radius PVC casings to depths of 17–19 m before the bedrock (Table 1). Three boreholes of BH1-3 locates in the north, while BH4-6 in the south with on horizontal offset of 2.5 m from BH7. The average groundwater level is between 2 and 3 m below the ground surface, and the overall flow of groundwater in the study area is from southwest to northeast. Free-phase oil was detected at the groundwater level. On August 26 and September 8, 2020, free-phase LNAPL floating on the groundwater was detected with a thickness ranging from 10 to a maximum of 50 cm in BH1, 4, and 5. Free-phase oil was not detected in BH2, 3, and 6, which are located on the right side of BH7. The floating LNAPL in BH7 was detected only on August 26, 2020. From this data, we can infer that the free-phase LNAPL distribution is mostly on the western side of the survey region and changes with groundwater fluctuations. Soil core sampling was conducted beside each borehole point. Generally, clay or silty sand is distributed from 3.5 m to a maximum of 7.5 m. The clay sand layer in BH1 was the shortest from 3.5 m to 5.5 m, and in BH5, the clayey distribution was less due to the sand layer between 6–7 m.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
</tr>
<tr>
<td>Drilling depth [m]</td>
</tr>
<tr>
<td>Soft rock [m]</td>
</tr>
</tbody>
</table>

3.2 Geochemical data on total petroleum hydrocarbon (TPH)

Soil sampling for geochemical analysis was performed in the area covered by geophysical profiles (KEITI, 2020). Boreholes A1 to A20 were dilled with a maximum depth of 9 m and a distance of 0.5 m between them. The soil sampling data showed that the TPH concentration is highest between A13 to A16, with the maximum value being 39,188 mg/L in A14. LNAPL contamination was found to be thickest in the range
of groundwater fluctuations. More than 2000 mg/L of contaminated plume was found in the range of groundwater levels between A9 to A17 (Fig. 2). However, TPH concentration in A1 to A8 is not significantly higher compared to the other side, and the contaminations in this region showed sparse distributions. For example, 1668 mg/L was found in A2 at a depth of 3.5 m and more than 2400 mg/L in A8 at depths of 3.5 m and 5 m.

When comparing TPH concentration and electrical conductivity (EC) of groundwater, a linear relationship was observed (Fig. 2). TPH data was collected from A11, 9, and 10, while EC of groundwater was collected from BH1, 6, and 7, which are positioned similarly. TPH in A11 was only concentrated above the groundwater level and less than 100 mg/L, and EC in BH1 was nearly constant at around 360 µS/cm. TPH in A9 had a peak value at depths of 4 m and 4.5 m, while EC in BH6 increased around those depths. TPH in A10 was found to be the most contaminated among the sampling data, and EC in BH7 was the most conductive with a maximum value of about 700 µS/cm. From the EC of groundwater data, it can be speculated that the LNAPL-contaminated region is conductive due to the presence of many ions in the pore liquid resulting from bio-degradation.

4. Geophysical Surveys and Interpretation

4.1 Geophysical Surveys

Considering the borehole locations and geological characteristics, we conducted ERT, CR, GPR, and seismic refraction surveys (Table 2). ERTL1-3 as well as CRL1-3 are inline profiles for ERT and CR surveys, respectively, passing through BH1, BH7, and BH6 in the South-East direction. ERTL4 is a crossline profile passing through all three boreholes. Seismic refraction data were collected along five profiles: two inlines (SRL1 and SRL2) and three crosslines (SRL3-5). Two parallel GPR survey profiles were conducted, with GPRL1 overlapping ERTL2, while GPRL2 was set at the northeast edge of the main survey area.

<table>
<thead>
<tr>
<th>Survey methods</th>
<th>Date</th>
<th>Device</th>
<th>No. Line</th>
<th>spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrical Resistivity Tomography</td>
<td>2020.08.26</td>
<td>AGI Superstring R8</td>
<td>4</td>
<td>1 m</td>
</tr>
<tr>
<td>Complex Resistivity</td>
<td>2020.09-10</td>
<td>Zonge ZT-30 (Tx), GDP 26ch (Rx)</td>
<td>3</td>
<td>2 m</td>
</tr>
<tr>
<td>Ground Penetrating Radar</td>
<td>2020.08.11</td>
<td>Sensor &amp; soft PulseEkko</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Seismic refraction</td>
<td>2020.08.26</td>
<td>Geometrics SmartSeis 24ch</td>
<td>5</td>
<td>3 m geophone</td>
</tr>
</tbody>
</table>

4.1.1 Seismic Refraction
Seismic refraction surveys were conducted in October 2000 to delineate the geological structure of the main survey area using SmartSeis 24 ch (Geometrics) and collected a 32,000 square-foot, multi-offset, 3D survey data. A hammer of 5 kg was used as a transmitting seismic source, and geophones were inserted into the earth every 3 m for each of the five survey lines with a length of 69 m. For each source, data were measured for 0.125 sec with a sampling rate of 0.125 msec.

4.1.2 GPR

Surface GPR survey were carried out using a Sensors & Software PulseEkko Pro system with a 100 MHz unshielded antenna and a 250 MHz shielded antenna along two profiles, GPRL1-2. GPRL1 has a length of 43 m with the same position as ERTL2, which passes through BH7. GPRL2 has a length of 90 m and was set apart from the main geophysical site to check the possibility of LNAPL leakage. We processed GPR data using RADPRO (Kim, 2005) and adapted DC filter, Dewowing, deconvolution, and filtering in the frequency domain to remove direct wave signals. We set the gain differently to amplify weak signals from deep depths.

4.1.3 ERT

AGI Superstring R8 and a dipole-dipole array were used to acquire ERT data. Three inlines (ERTL1-3) and one crossline (ERTL4) were deployed with a 1 m electrode spacing for each profile, reaching a total length of 39 m with 40 stainless-steel electrodes (Fig. 3b). The surface was quite wet due to heavy rain prior to the ERT survey. Inversion of resistivity data was made using an in-house algorithm based on the non-linear Gauss-Newton method for 3D inversion (Sasaki, 1994) and a commercial software of Dipro (KIGAM, 2001) for 2D inversion.

4.1.4 CR

Three parallel CR lines (CRL1-3) were set identically on the ERT inline profiles, but the CR line lengths were 27 m, which were shorter than those of the ERT (Fig. 3c). A dipole-dipole array was used with a dipole spacing of 2 m, but an electrode spacing of 1 m was used to improve data resolution. To reduce capacitive coupling, cable layout separation between the transmitter and receiver was used (Dahlin and Leroux, 2012). Stainless steel electrodes were used for current transmission, and CuSO4 non-polarizable electrodes from AGI were used for potential measurement to secure data from the galvanic coupling effect. The source frequencies were 0.125, 1, and 8 Hz. CR measurements were made using a ZT-30 in Zonge for a multi-purpose EM transmitter and a GDP-26 ch in Zonge for a receiver. To recover 3D complex resistivity models, an in-house algorithm that can consider the effect of adjacent frequency data was used (Kim et al., 2022).

4.2 Interpretation of Geophysical Data

The geophysical data were first interpreted without considering any prior information. Seismic refraction revealed layered structures up to a depth of 20 m, while GPR, ERT, and CR data showed some anomalous features near the surface.
4.2.1 Seismic Refraction

Inverted seismic refraction sections along SRL1 and SRL4 (Fig. 4) reveal gradually increasing velocity structure without any remarkable faults. The section along SRL1 passing through the ERT and GPR profiles (Fig. 3a), clearly shows velocity structures of three layers. The first layer extended to a depth of 5 m with a velocity of approximately 1.2 km/s, and the basement layer was located near a depth of 15 m with a velocity of more than 4.8 km/s. In the cross line SRL4, the basement structure was shown near a depth of 15 m, but the boundary between the first and second layers was less clear than in SRL1. The seismic refraction sections provided approximate geological structures and depths of the basement, but it was difficult to obtain information about contamination.

4.2.2 GPR

Although GPR signal attenuation can be a good indicator for LNAPL contaminants with high resolution, GPR responses in Site Y provide limited information from deep depths (Fig. 5). Although responses along GPRL1 at 100 MHz showed good resolution in identifying the contaminated region, proper interpretation of GPR data was challenging in other lines and frequencies (Fig. 5a). The section along GPRL2 at 100 MHz shows sparse signal attenuation, but these are from puddles of water on the surface and not from contamination (Fig. 5b). Also, large hyperbolas from 0–20 m and 65–90 m in line 2 occurred from surrounding metallic containers and train rails that interrupted signal interpretation from the subsurface. The 250 MHz frequency data has many small noise hyperbolas, which occur due to gravels in the backfills. The wavelength decreases as the frequency increases, and the resolution appears for a medium of about a quarter of the wavelength. In the case of 100 MHz signals, which have larger wavelengths than 250 MHz and less resolution, small gravels cannot be detected. This effect makes it difficult to interpret the contamination field in the 250 MHz sections. Below the direct wave, the signal is entirely weakened and then amplified after 80 nsec. The result can be interpreted that the difference in impedance occurs not only in LNAPL but also in other areas. Water, with the highest electrical permittivity, affects the signal to change rapidly at the boundary of the water table, leading to a completely increased signal interpreted as groundwater.

4.2.3. ERT

To recover resistivity models of the Y site, we performed 2D inversion with Dipro software for the entire profiles (ERTL1-4) (Fig. 6a) and 3D inversion with our in-house algorithm for inline profiles (ERTL1-3) (Fig. 6b, c, and d). In the resistivity sections, the top layer shows high resistive zones with a thickness of about 2 m at 16–23 m on the x-axis. Some regions exhibit much lower resistivity, especially near –10 m to 3 m in the x-axis of ERTL1 and ERTL2 and from –15 m to –5 m of L3. In the crossline (ERTL4) section, we can confirm that the resistive zones near the surface are distributed at 16 m ~ 18 m and 22 m ~ 23 m in the y-axis where ERTL3 is crossing. When comparing the ERT sections with the section of GPRL1 for 100 MHz in Fig. 5a, we observe that GPR signal attenuation around –13 m to 0 m is consistent with the
conductive area in the ERTL2 section. However, low resistivity regions with little changes from a depth of 2 m do not provide meaningful information to identify contaminants from the background medium.

4.2.4. CR

We performed 3D inversion of 0.125, 1, and 8 Hz CR data and plotted sections of real ($\sigma'$) and imaginary ($\sigma''$) conductivities and phase ($\phi$) (Fig. 7). The $\sigma'$ sections in the conductive area near the surface showed similar trends to the resistivity sections from ERT data, but the recovered $\sigma'$ models were smoother due to the larger $a$-spacing, 2 m in CR compared to 1 m in ERT. Considering the real component of CR implies the bulk electrical resistivity of the porous media, we can expect contaminated areas with low resistivity above the groundwater level. However, even with double the $a$-spacing of ERT, delineating subsurface structures of $\sigma'$ below about 2 m depths is still challenging due to the less change in the $\sigma'$ values.

On the other hand, the $\sigma''$ and $\phi$ sections showed some variations below the depth of 2 m, unlike the $\sigma'$ sections. IP effects increased from approximately 2.5 m, especially in CRL3 sections, and very low IP anomalies were also noticeable both near the surface and around 5 m depth. However, analyzing the imaginary conductivity and phase sections without any geological background can lead to misinterpretation. This is because IP responses are affected not only by geological factors such as clay and microbial byproducts but also by the condition of oil contamination, such as wettability and the phase of oil (free-phase, disseminated). To interpret CR data reliably, geological information is necessary to distinguish the responses of contaminants from the background medium.

5. Estimation of LNAPL Plumes Using Geophysical Data and Discussion

In this section, we present our approach to delineate LNAPL plumes in the study site. Firstly, we interpret the geophysical data independently to evaluate the information that can be obtained from geophysical methods. Subsequently, we incorporate borehole and geophysical data into a 3D geological model to consider the site's geological setting. Finally, we perform a comprehensive interpretation of the geophysical data combined with all direct method data, including borehole data and soil sampling data.

5.1 Interpretation Based Solely on Geophysical Data

From the inverted seismic refraction sections (Fig. 4), we can assume that Site Y has a layered structure with a bedrock at a depth of 14 m. Since LNAPL contamination typically occurs near groundwater level, more distinct responses from oil contaminants near the surface are deplaned in the resistivity sections (Fig. 6): high-resistive zones from the surface to a depth of 2 m might be attributed to the unsaturated zone or fresh free-phase oil with non-conductive properties. Since oil leakage had occurred over 10 years ago, interpreting the resistive zone as an unsaturated zone is more reasonable due to the biodegradation effect. Additionally, GPR data showed strong reflections at a depth of 3 m (Fig. 5a), which likely indicates the groundwater level near the surface.
In contrast to ERT and GPR data, IP data \((\sigma'', \phi)\) display remarkable responses below a depth of 3 m, with some strong IP regions or areas with values close to 0 (Fig. 7). Nonetheless, interpreting IP signals in NAPL-contaminated sites solely based on these responses is nearly impossible, as IP responses show very contradictory results in real field cases (Flores Orozco et al., 2021). However, more accurate interpretation of IP data needs additional information such as borehole or soil/groundwater sampling data.

5.2 Geophysical Interpretation Considering Geological Setting of the Site with a Clay-Rich Soil Layer

Before estimating LNAPL plumes from geophysical data, we first constructed a 3D geological model based on borehole data and geophysical data. Based on the 3D geological model, the ERT and CR data are interpreted.

5.2.1 Construction of a 3D geological model integrating borehole data and geophysical data

We used an open source code called ‘GemPy’ (de la Varga et al., 2019) to construct a 3D geological model from the geological labeled points, which were obtained from the input data, i.e., borehole data and geophysical data. Firstly, borehole core sampling data were used to establish the basic stratigraphy model providing soil and basement rock distribution with depths, while groundwater level data to divide the unsaturated and saturated zones. Based on the borehole data, we identified the following geological labels: unsaturated backfill, saturated backfill, clay, weathered soil, and bedrock (Fig. 8).

Geophysical data were then used to extend the model from the borehole data by providing geological tendencies on a broad scale. Seismic refraction data showed that the field had smooth layered structures without any faults. After comparing the seismic refraction data with the borehole data, we set the boundaries of the bedrock layer with velocity structures greater than 2.7 km/s. GPR data also showed that the groundwater level remained at quite steady depths along the profiles. The final geological model has a dimension of 40 m × 10 m × 20 m (Fig. 8) to cover the ERT and CR profiles, which are the main geophysical data used to analyze the contamination zone. In the \(xz\) sections of the geological model, the stratigraphy is generally smooth, but the clay zone is about 2 m thinner between \(x\)-distances \(-30\) to \(-3\) m than the right side of the model, especially in L2 and L3 sections, which are along the ERTL2 and 3.

5.2.2. Interpretation of ERT and CR data considering electrical effects of clay on Geophysical data

For resistivity sections, the unsaturated zone generally has higher resistivity values than the saturated zone if the soil material is the same. Thus, the low resistive zone in Fig. 6 can be interpreted as the contaminated area, and the area between from \(-10\) to \(3\) m in the \(x\)-axis is more contaminated than the
right side of the profiles, especially in ERTL1 and 2. On the other hand, distinguishing the response at approximately a depth of 2 m is not clear even if the geological model is known due to the less variation of the resistivity values in sections.

In CR data, the general IP effect of $\sigma''$ and $\phi$ sections in Fig. 7 matches with the clay material of the site except for the unsaturated zone with little IP response. Most of the medium has more than 0.2 mS/m, and we can suppose the response is from clayey material which occurs in IP responses. Some regions had very low values ($\sigma'' < 0.1$ mS/m) at the unsaturated zone where the soil sampling revealed the high TPH concentration. The anomalous lower $\sigma''$ values are also obtained from 2 m to $-13$ m in the $x$-axis, which can be due to the high TPH concentration. The phase sections generally show a similar response with $\sigma''$ but some regions have high $\phi$ values where $\sigma''$ values are close to 0 due to the low $\sigma'$ values. Comprehensively, the resistivity and $\sigma'$ reveal the contaminated zone only for the unsaturated zone because the contaminated zone has high conductivity like the clay layer in the saturated zone. On the other hand, the IP response ($\sigma''$, $\phi$) indicates a high TPH zone by very small IP values contrary to the clay zone even under the groundwater depth.

### 5.3 Interpretation of Multi-Geophysical Data in Conjunction with Soil Sampling Data

#### 5.3.1 Comparison of LNAPL Plume Delineation Using Partial Sets of Soil Sampling Data

To assess the effectiveness of geophysical data in the delineation of LNAPL plumes at Site Y, we compared the interpolation results of several sets of soil sampling data with the interpretation of geophysical data considering the geological setting of the site shown in Fig. 9. While soil sampling data may have limitations in representing the spatial distribution of oil contaminants on a broad scale (Cassiani et al., 2014; Shao et al., 2019; Flores Orozco et al., 2021), we consider the interpolated TPH distribution using the whole soil sampling data of the site (Fig. 9a) as the actual spatial TPH distribution due to the comprehensive sampling conducted at the site. We plotted interpolated TPH sections from three different sets of six sampling points (A1-3 and A13-15, A1-3 and A9-11, A5-7 and A13-15) in Fig. 9b, c, and d, respectively. The interpolated TPH xz sections (L1, L2, and L3) parallel to ERT profiles ERTL1, L2, and L3, respectively, were compared with the geophysical interpretation Fig. 9e and f).

- Interpretation using A1-3 and A13-15 soil sampling data

In the interpolated TPH cross-section (Fig. 9b), the contamination source near $x = 0$ m at L1 and L2 was not clearly visible, and the interpolated region at $x = 10$ m which showed less contaminated did not correspond with highly contaminated areas depicted in full sampling data section (Fig. 9a). This is attributed to the fact that excessive smoothing TPH values from few sampling points to recover a large area. In the ERT sections (Fig. 9e), low-resistivity areas corresponded well with highly contaminated regions in Fig. 9a, with the exception of the area near $x = 0$ m at the L3 section.
In the \( \sigma'' \) cross-section Fig. 9f), \( \sigma'' \) values neared 0 in the region with the highest TPH value in the A14 data. From this interpretation, the severely contaminated areas aligned well at \( x = 0 \) m, located in the middle of the survey line. Nonetheless, the alignment was less accurate at \( x = 10 \) m, particularly at A6, where the TPH concentration was measured as the lowest. The inadequate correspondence between the response at \( x = 10 \) m and the \( \sigma'' \) of other areas might be due to the limited resolution of the investigation depth at the end of the CR survey line. Another explanation for the discrepancy could be the influence of the contamination plume measured near \( x = 20 \) m, which is beyond the survey line.

- Interpretation using A1-3 and A9-11 soil sampling data

The TPH section showed continuous contamination plumes near 10 m in the \( x \)-axis at sections L2 and L3 due to interpolation from A6 and A7 data (Fig. 9c), despite these areas being relatively less contaminated in full sampling data section (Fig. 9a). In ERT sections (Fig. 9e), low resistivity areas generally correspond to regions with high TPH concentrations, while high resistivity areas correspond to regions with less contamination. Consequently, resistive area near 10 m in the \( x \)-axis at ERTL2 and ERTL3 can be interpreted as areas with relatively less contamination.

Conversely, the \( \sigma'' \) value approaches nearly 0 at a depth of about 2 m, where the contamination is most severe in A9-11 data (Fig. 9f). Based on this interpretation, the \( \sigma'' \) value decreases significantly in areas where the TPH value exceeds 100 mg/L at the A13-15 points located at \( -10 \) m in the \( x \)-axis and the A7 point located at 10 m. Nevertheless, very small \( \sigma'' \) area appears near a depth of 5 m at the A6 point, which has a comparatively low TPH concentration, leading to potential misinterpretation as an area with a high level of contamination.

- Interpretation using A5-7 and A13-15 soil sampling data

In interpolated TPH sections using A5-7 and A13-15 data (Fig. 9d), the general distribution of TPH values is similar with the interpolated TPH sections from all sampling points (Fig. 9a) but several points of highly contaminated regions at the near surface especially at 0 m in \( x \)-axis are missing. In ERT sections (Fig. 9e), the conductive zones can be expected to contaminated zone in the near surface by comparing A13-15 data with conductive regions near \( -10 \) m in the \( x \)-axis and A5-7 data with resistive regions near 10 m. As verifying in 0 m in the \( x \)-axis, highly contaminated plumes in L1 and L2 could be expected by conductive responses from ERT data but not in L3 due to the resistive zone.

5.3.2 Final estimation of LNAPL plumes considering all of the soil sampling data

TPH concentrations from soil sampling data are correlated with the inversion results of ERT and CR data at the same depth, even though they have different measurement scales. To account for the variation in electrical characteristics of background materials with depth, we divided the data into unsaturated soil at depths of 0 ~ 2 m and clay layer at depths of 4 ~ 6 m and plotted the relations between TPH
concentrations from soil sampling data and electrical properties from inversed ERT and CR models (Fig. 10).

In the unsaturated soil depth, we observed a significant decrease in resistivity values from recovered ERT models with increasing TPH concentration up to about 100 mg/L, after which there was no significant change in resistivity responses. The inversed CR models showed similar trends with electrical resistivity. However, some points were observed with high electrical conductivity (σ') at a relatively low TPH concentration of about 30 ~ 70 mg/L. On the other hand, σ showed a gradual decrease to close to 0 with increasing TPH values. In the clay layer, while the values of electrical resistivity and σ' did not show much variation with TPH concentration, σ'' showed a sharp attenuation trend from approximately 70 mg/L. This result can be attributed to the flow path of ions related to the polarization phenomenon being blocked by highly polarized NAPL contamination in clay layers.

The depth slices of electrical resistivity (σ'') and interpolated soil sampling data are plotted together for each of unsaturated soil, saturated soil and clay layers (Fig. 11). The resistivity depth slices (Fig. 11b) show that conductive zones at 0 m in the x-axis are generally well matched with highly concentrated TPH zone (< 100 mg/L) except for the resistive region near the location of -2.5 m in y axis near A11. This result can be interpreted by inferring the comparisons between TPH concentration from soil sampling data and groundwater EC measured from borehole data in Fig. 2. The highly contaminated region near A11 was resistive in ERT data due to the effect of the highly concentrated TPH, which is located apart from the groundwater level, having little effect on the groundwater EC measured.

In the depth slice of the σ'' anomaly (Fig. 11c), very low IP responses from -10 to 0 m in the x-axis are matched with high concentrated TPH area not only at unsaturated zone but also saturated zone and clay layer. However, relatively less contaminated region around A6 and 7 (x = 10 m) was not detected in the area where IP response was very low. This is possibly attributed to the fact that the edge of the survey area have lower resolution with depths, even though it could also indicate the presence of contamination plume between x = 10 ~ 20 m because high TPH concentration was detected at x = 20 m in the soil sampling data.

6. Conclusions

In this study, we demonstrated the application of various geophysical methods to LNAPL-contaminated sites and discussed the challenges of interpreting geophysical data in clay-rich environments. Our results revealed that interpreting geophysical data alone can be ambiguous due to the inconsistent electrical response of biodegraded oil and the presence of clayey materials. To distinguish NAPL contaminants from the background medium, a 3D geological model was constructed by integrating borehole and geophysical data. Seismic refraction data, less affected by oil contamination, proved particularly useful for extending the geological structure from borehole data.

Based on the geological model, we discovered that resistive properties from ERT data and CR real conductivity data generally aligned well with oil-contaminated regions in the unsaturated backfill zone.
However, resistive responses appeared sensitive to dissolved ion concentrations in pore liquids rather than the contamination plume and exhibited less deviation in the saturated backfill and clay layers. CR data provided additional information by highlighting areas with very low IP values in the clay layer, which could be interpreted as highly contaminated regions. Nevertheless, these areas did not consistently correspond perfectly with high TPH zones from soil sampling data.

When compared to interpolated TPH values from sampling data, geophysical data could identify highly contaminated areas located between the sampling points. Even though applying geophysical surveys to NAPL-contaminated sites necessitates careful consideration, especially in clayey environments, meaningful information on NAPL plumes can be obtained if one can make interpretation of geophysical data with a thorough understanding of the geological structure of the target site and the influence of NAPL on geophysical parameters.

**Declarations**

**Ethical Approval**

This article does not contain any studies with human participants or animals performed by any of the authors.

**Consent to Participate**

All the authors are agreed to be listed as authors in the current version of the manuscript.

**Consent to Publish**

All the authors agreed for the publication of the manuscript in the Environmental Science and Pollution Research.

**Authors Contributions**

Conceptualization: Bitnarae Kim, Myung Jin Nam; Methodology: Bitnarae Kim, Myung Jin Nam; Formal analysis and investigation: Bitnarae Kim, Inseok Jeong, Huieun Yu, Juyeon Jeong, Seo Young Song, Jeong-sul Son, Youngchul Yu; Visualization: Bitnarae Kim, Inseok Jeong, Huieun Yu, Juyeon Jeong, Seo Young Song; Validation: Bitnarae Kim, Jeong-sul Son, Ho Young Jo, Man Jae Kwon; Writing - original draft preparation: Bitnarae Kim; Writing - review and editing: Bitnarae Kim, Myung Jin Nam; Funding acquisition: Myung Jin Nam, Jehyun Shin, Ho Young Jo; Resources: Jeong-sul Son, Youngchul Yu, Man Jae Kwon; Supervision: Myung Jin Nam

**Funding**

This work was supported by Korea Ministry of Environment as “The SEM projects” (Grant No. 2018002440005; 2018002440001; and 2018002440002)
Competing Interests

The authors declare no competing interests.

Availability of data and materials

Not applicable.

References


Figures
(a) Location of Site Y, (b) a drone image of the site, and (c) the locations of sampling points and boreholes in the site. The site was contaminated with LNAPL throughout the region.

Figure 1
Figure 2

(a) Scatter plot of total petroleum hydrocarbon (TPH) values at different depths of each soil sampling position. Groundwater levels varied from a depth of about 2.3 m to 3 m. (b) Comparisons between TPH from soil samplings (A-11, 10, and 9) and electrical conductivity (EC) of groundwater from boreholes (BH1, 7, and 6).
Figure 3

(a) Map of the study area showing the locations of the geophysical profiles determined by seismic refraction (SR1-4, red lines), ground penetrating radar (GPR) (GPR L1-2, black lines), electrical resistivity tomography (ERT) (ERT L1-4, blue lines), and complex resistivity (CR) (CR L1-3, yellow lines). The arrow indicates the direction of measurement progress. Detailed locations of (b) ERT and (c) CR profiles with soil sampling and borehole points.
Figure 4

Seismic refraction inversion results for (a) SR1 and (b) SR4.

Figure 5
GPR 100 MHz data from profile (a) GPR L1 and (b) GPR L2, 250 MHz data from (c) profile GPR L1 and (d) GPR L2.

Figure 6

(a) Fence diagram of 2D inversion results with Dipro (KIGAM, 2001) for the entire profiles (ERT L1-4) and (b) xz sections of 3D inversion results with the in-house algorithm for in-line profiles (ERT L1-3).
Figure 7

xz sections of recovered CR models by 3D inversion algorithm. The electrical response is expressed in terms of the real, imaginary, and phase of the complex conductivity at different frequencies: (a) 0.125 Hz, (b) 1 Hz, and (c) 8 Hz.
Figure 8

(a) 3D geological model of the site Y incorporating borehole and geophysical data, along with corresponding xz sections of (b) L1, (c) L2, and (d) L3 which are parallel to ERT L1-3, respectively. The red boxes in (b), (c), and (d) indicate the main target depths for the geophysical surveys.
Figure 9

Interpolated TPH distribution in xz sections from sampling data: (a) all sampling points, (b) A5-7 and A13-15, (c) A1-3 and A9-11, or (d) A1-3, A13-15. Expected LNAPL contaminants from (e) ERT sections, and (f) imaginary components of complex conductivity sections.
Figure 10

Plot of electrical parameters against the measured TPH concentrations from selected sampling points (A5, 6, 7, 13, 14, and 15). The ERT resistivity, and real component and imaginary component of the complex conductivity (from left to right) were extracted from the inverted values at depths of 0.5 to 1.5 m and 3.5 to 5 m, which correspond to the unsaturated and clay zones identified from borehole data, respectively.

Figure 11
Depth slices for (a) interpolated TPH distribution from soil sampling data, (b) resistivity, and (c) imaginary conductivity. The depth slices from top to bottom correspond to the unsaturated zone and clay zone identified from borehole data, respectively.