Optimized survey design for the joint use of direct current resistivity and induced polarization: Monitoring of DNAPL source zone evolution at a virtual field site

Siyuan Qiang, Xiaoqing Shi, André Revil, Xueyuan Kang, Christopher Power

▶ **To cite this version:**

Siyuan Qiang, Xiaoqing Shi, André Revil, Xueyuan Kang, Christopher Power. Optimized survey design for the joint use of direct current resistivity and induced polarization: Monitoring of DNAPL source zone evolution at a virtual field site. Journal of Contaminant Hydrology, 2024, 267, 10.1016/j.jconhyd.2024.104452 . hal-04860141

HAL Id: hal-04860141
https://hal.science/hal-04860141v1

Submitted on 31 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

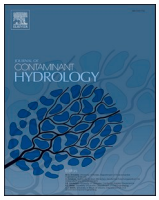
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal of Contaminant Hydrology 267 (2024) 104452

Contents lists available at ScienceDirect

Journal of Contaminant Hydrology

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



Optimized survey design for the joint use of direct current resistivity and induced polarization: Monitoring of DNAPL source zone evolution at a virtual field site

Siyuan Qiang [a], Xiaoqing Shi [a,*], André Revil [b], Xueyuan Kang [a], Christopher Power [c]

[a] Key Laboratory of Surficial Geochemistry of Ministry of Education, School of Earth Sciences and Engineering, Nanjing University, Nanjing 210023, China
[b] Université Savoie Mont Blanc, CNRS, UMR CNRS 5204, EDYTEM, 73370 Le Bourget du Lac, France
[c] Department of Civil and Environmental Engineering, Western University, London, Ontario N6A 5B9, Canada

ARTICLE INFO

Keywords:
Optimized survey design
Induced polarization
DNAPL source zone evolution
Aquifer analog
Virtual site
Hydrogeophysics

ABSTRACT

The combined application of direct current (DC) resistivity and induced polarization (IP) methods, referred to as combined DCIP method, has gained popularity for characterizing the critical zone dynamic processes such as dense non-aqueous phase liquids (DNAPLs) spreading at contaminated sites. Large-scale DCIP surveys typically require considerable durations, necessitating optimized survey designs to enhance survey resolution while controlling time and labor costs. However, to date, approaches to optimize geoelectrical survey design have focused solely on DC applications, and the efficiency of optimized survey designs for combined DCIP is yet to be investigated. Moreover, as subsurface heterogeneity would impact the geophysical observations, most field-scale numerical DCIP studies have still been conducted at artificial sites that lacked realistic aquifer heterogeneity, which could affect the validity of the DCIP survey evaluations. In this work, a virtual geoenvironmental field site based on high-resolution real aquifer analog was created to simulate a DNAPL evolution scenario with simultaneous monitoring by DCIP survey, employing both the optimized survey design and popular non-optimized survey designs (Wenner, Wenner-Schlumberger, Dipole-Dipole arrays). Results show that the optimized survey with prior information improves the monitoring accuracy of DNAPL source zone (SZ) by 8 to 19 % with respect to different DCIP characteristics (conductivity, chargeability, normalized chargeability, and relaxation time). Another ideal numerical test indicates that the optimized survey shows up to an 83 % reduction in measurement time compared to the conventional survey, while maintaining the same subsurface image resolution. Additionally, the optimized surveys designed without or with limited prior information were also shown to be more efficient than conventional survey for imaging the entire subsurface space. The findings in this study highlight the immense potential of optimized survey design methods for enhancing the efficiency of DCIP surveys on subsurface contaminants and hydrological processes.

1. Introduction

DNAPL contamination in groundwater poses a significant environmental threat (Kang et al., 2020; National Research Council, 2013). Once released to the subsurface, free-phase DNAPL tends to persist as disconnected residual ganglia or pools under the control of gravity and capillary force (Dekker and Abriola, 2000; Al-Raoush, 2009). Due to their low solubility and high resistance to degradation, DNAPL SZs have the potential to release pollutants persistently over many decades (Moretti, 2005). The dissolution pattern and environmental impact of DNAPLs are heavily controlled by the mass and architecture of DNAPL SZs (Kang et al., 2022, 2024; Wang et al., 2023). Hence, high-resolution characterization of DNAPL SZs at contaminated sites is essential for designing efficient remedial programs and assessing their performance following implementation.

Traditionally, direct investigation methods such as intrusive drilling or sampling are used at DNAPL sites, but they can be highly laborious, time-consuming, and suffer from limited spatial and temporal image resolution. Geophysical techniques are an attractive alternative since they can provide non-intrusive, spatially connected data in a shorter time at a much lower cost (Slater and Binley, 2021). Geoelectrical methods such as direct current (DC) resistivity and induced polarization

* Corresponding author.
E-mail address: shixq@nju.edu.cn (X. Shi).

https://doi.org/10.1016/j.jconhyd.2024.104452
Received 14 June 2024; Received in revised form 23 October 2024; Accepted 28 October 2024
Available online 30 October 2024
0169-7722/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

(IP) have become increasingly popular for mapping DNAPLs due to their sensitivity to the presence of resistive DNAPL (e.g., Johansson et al., 2015; Koohbor et al., 2022; Revil et al., 2012a; Han et al., 2024a). More recently, the combined DC resistivity and IP (DCIP) method have been used for DNAPL investigations, since IP characteristics such as chargeability, normalized chargeability, and relaxation time can better differentiate DNAPL SZs from surrounding soils with similar electrical resistivity (e.g., Almpanis et al., 2021; Kang et al., 2023; Sparrenbom et al., 2017). Meanwhile, IP chargeability could also help to identify the subsurface water content, and the characteristic relaxation time can be used to predict the permeability to flow of porous materials (De Lima and Niwas, 2000; Revil, 2013; Revil et al., 2015). In addition, IP characteristics can be used to image the subsurface contaminants and hydrological processes especially for clay-rich areas, because clay-rich area usually exhibits significant IP response and surface conduction that would disturb the interpretation of DC conductivity data alone. Meanwhile, the signal-to-noise ratio (SNR) in IP data acquisition often being 10–100 times smaller relative to the DC (Revil et al., 2018). Therefore, a comprehensive evaluation of combined DCIP surveys at field sites is needed to assess its applicability for a range of field survey scenarios.

The success of DCIP survey at field sites is based on the spatial image resolution that can be obtained, which is influenced by the positioning of the electrodes along a survey line and the measurement arrays used (Griffiths et al., 1990). Conventional electrode arrays such as Wenner, Wenner-Schlumberger, and dipole-dipole arrays are most commonly used due to their ease of use, but they are not always the best option for DCIP surveys due to their relatively low subsurface resolution and prolonged measurement times (Wilkinson et al., 2006). For instance, if DCIP is used to monitor transient subsurface processes that are providing a constantly changing target (e.g., fluid migration), the extended durations due to massive standalone measurements can lead to image blurring (Wilkinson et al., 2015). Therefore, DCIP surveys should be optimized to collect more subsurface information within time and labor constraints (Binley and Slater, 2020).

The 'Compare-R' method is the most widely used approach for optimizing survey design (Wilkinson et al., 2012). It selects measurements with maximum estimated resolution to provide highly efficient surveys that can improve imaging accuracy with a much lower number of measurements. The computational burden of the 'Compare-R' method can be enormous due to the required calculation of the resolution matrix, motivating Qiang et al. (2022) to propose a more computationally efficient survey design method based on Bayesian experimental design to maximize the information obtained by a standalone DC survey from the subsurface target area. While the method performed well within a 'target zone', it lacks resolution on areas outside the target which could lead to unrealistic inversions of the collected data, especially data from heterogeneous subsurface environments. To date, the 'Compare-R' and Bayesian methods have focused solely on DC surveys (e.g., electrical resistivity tomography [ERT]), with the performance of optimized survey design on IP characteristic tomographies such as chargeability, normalized chargeability and relaxation time, yet to be assessed.

A comprehensive understanding of the subsurface environment is needed to evaluate the efficiency of combined DCIP surveys using an optimized or non-optimized survey design. Most DCIP field studies have been performed at sites with little-to-no subsurface information, with the large amount of uncertainties hindering DCIP efficiency investigations. This has led to the employment of 'virtual sites', which refer to field-scale numerical model domains that contain high-resolution hydrogeological and hydrogeochemical parameters (Mumford et al., 2022; Xie et al., 2023). The transparency of virtual sites allows all subsurface parameters to be known, making them an ideal and effective approach to investigate a range of geoelectrical survey capabilities (e.g., Almpanis et al., 2021; Kang et al., 2023; Power et al., 2015). So far, most virtual sites have been created from idealized hydrogeological parameters, such as spatially correlated random permeability (Power et al., 2013), thereby lacking true subsurface heterogeneity. Since heterogeneity disturbs the geoelectrical response from subsurface targets, such as DNAPL SZs (e.g., Kang et al., 2018), it provides a significant obstacle to geoelectrical data interpretation; therefore, it is critical to implement realistic subsurface heterogeneity in numerical model domains used to evaluate geoelectrical surveys. Aquifer analogs can provide spatially explicit representations of genuine geological structures (e.g., Maji et al., 2006), meaning virtual sites created from these analogs contain high-resolution subsurface parameters while preserving subsurface heterogeneity information. These virtual sites with realistic subsurface heterogeneity have yet to be used in numerical investigations on combined DCIP performance.

The objective of this study is to evaluate the efficiency of an optimized DCIP survey design method (based on Bayesian experimental design) for characterizing DNAPL SZs. A field-scale 3D high-resolution aquifer analog was used to create a realistic virtual site with the evolution of a DNAPL SZ simultaneously monitored by DCIP survey, which used both optimized and non-optimized survey design. The imaging ability and measurement time of optimized DCIP surveys were compared to the conventional non-optimized arrays. Further, the sensitivity of different DCIP characteristics (conductivity, chargeability, normalized chargeability, and relaxation time) to subsurface heterogeneity and DNAPL SZ was explored.

2. Materials and methods

2.1. Construction of virtual site

2.1.1. High-resolution aquifer analog

The aquifer analog utilized in this study originates from the fluvial-glacial sediments located in Herten, Germany (Heinz et al., 2003). The original aquifer dataset includes hydraulic conductivity and effective porosity measured using undisturbed samples collected in the field at a resolution of 5 cm by 5 cm, featuring six vertical cross-sections sized 16 m × 7 m at 2 m intervals. Bayer et al. (2015) employed stochastic techniques to extrapolate these sections to a 3D domain. Additional adjustments were made to the 3D aquifer analog prior to implementation for DCIP monitoring of DNAPL evolution:

(1) Resolution adjustment: The resolution of the aquifer analog was changed from 5 cm to 10 cm to reduce the number of cells that need to be computed and streamline computational resources. The upscaling is done using unweighted average method.
(2) Extension: The original 16 m × 10 m × 7 m aquifer analog was expanded to enlarge the site and accommodate a typical and uniform surface DCIP survey grid. It was expanded to 16 m × 20 m × 7 m by duplicating the analog along the original *y*-direction. Then this longer '*y*-direction' was converted to the *x*-direction, with the updated analog being 20 m × 16 m × 7 m in the *x, y* and *z* directions.
(3) Scaling of hydraulic conductivity: The combined DCIP method is known to be sensitive to clay minerals (Revil et al., 2023) which usually exhibit low hydraulic conductivity. However, due to the relatively high hydraulic conductivity of the original analog (for most cells, K is larger than 10^{-5} m/s), it was difficult to assign sufficient clay contents to the domain that could provide a measurable IP response. Therefore, the hydraulic conductivity of the analog was scaled by a factor of 1×10^{-5} to align it with the normal hydraulic conductivity range in clay-rich aquifers (Kang et al., 2023; Revil et al., 2015) and allow a distribution of clay contents to be estimated. These modifications do not change the structural integrity of the aquifer. The final hydraulic conductivity K and porosity ϕ of the aquifer are present in Fig. 1.

2.1.2. Determination of the volumetric clay content

The following equation by Marion et al. (1992) and Revil and Cathles (1999) was employed to calculate the clay content *cl* from the

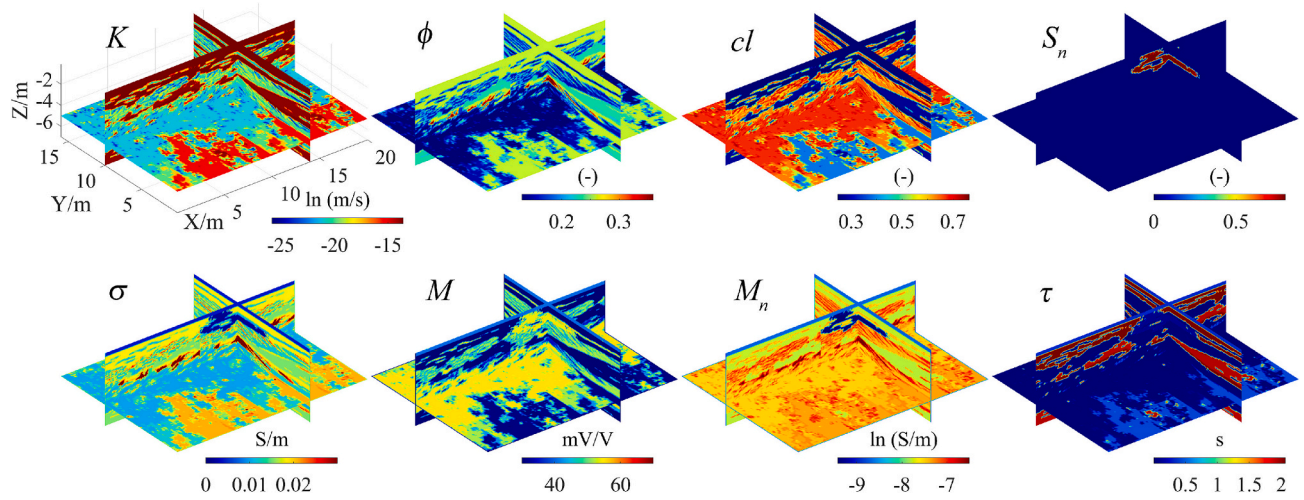


Fig. 1. Cross-sectional slices of the true parameter fields at $x = 14$ m, $y = 10$ m, and $z = -5$ m of the virtual site. *K, ϕ, cl* represent the true hydraulic conductivity, porosity, and clay content of the modified aquifer analog, respectively. S_n denotes the simulated DNAPL saturation prior to evolution, while σ, M, M_n, τ denote the conductivity, chargeability, normalized chargeability, and relaxation time distribution (estimated from *K, ϕ, cl,* S_n), respectively.

permeability (directly correlated to hydraulic conductivity) of each cell in the modified aquifer analog:

$$cl = \frac{\phi_{sd}\left(k_{sd}^{1/3m} - k^{1/3m}\right)}{k_{sd}^{1/3m}(1-\phi_{cl})} \tag{1}$$

where ϕ_{cl} (−) and ϕ_{Sd} (−) represent the porosity of pure clay and sand, respectively, m (−) denotes Archie's cementation or porosity exponent (comprised between 1 and 3), and k (m^2) is the permeability. The lower threshold permeability of clay-free sand k_{sd} (m^2) is determined by (Revil and Cathles, 1999):

$$k_{sd} = \frac{d_{sd}^2(\phi_{sd})^{3m}}{24} \tag{2}$$

where d_{sd} (m) is the grain diameter of sand. All permeabilities larger than k_{sd} were assumed to correspond to clay-free sands ($cl = 0$). The resultant clay content distribution is illustrated in Fig. 1. The value of parameters used in the determination of *cl* are shown in Table 1.

2.1.3. DNAPL percolation and evolution

A stochastic invasion-percolation (SIP) algorithm (Ewing and Berkowitz, 1998) was employed to simulate DNAPL release, infiltration, and redistribution (no dissolution for this study) into the aquifer. It allows for the generation of SZ with architectures consistent with multiphase processes at the representative element volume scale. DNAPL infiltration is modeled by simulating invasion into each model cell based on an entry pressure threshold, which is determined by the Young-Laplace equation. These thresholds are updated with the capillary pressure-saturation relationship that quantifies the capillary resistance against water and DNAPL saturation. The maximum DNAPL saturation can be achieved for each cell is imposed by the residual saturation of water, which is 0.2 in this study (Koch and Nowak, 2015). The final architecture of the source zone is influenced by the permeabilities and a stochastic component that accounts for unresolved pore-scale variability. This SIP framework efficiently generates complex, realistic SZs composed of ganglia and pools, which aligns with the intrinsic permeability field and is widely utilized in DNAPL simulations (e.g., Ayaz et al., 2020; Bandara et al., 2016; Koch and Nowak, 2015, 2016). The depth of the water table for the synthetic domain is set as 0.4 m. Since the SIP algorithm cannot deal with the DNAPL infiltration in unsaturated zone (Ewing and Berkowitz, 2001), the water saturation above the water table was set as a homogeneous value of 0.4, where the capillary fringe and the presence of DNAPL in unsaturated zone were neglected.

Table 1
Model parameters for the numerical simulation.

Parameter	Value
Aquifer analogy	
Domain Size (x, y, z) (m)	20 × 16 × 7
Grid spacing (m) (Δx, Δy, Δz,)	0.1 × 0.1 × 0.1
Water level (m)	6.6
DNAPL source zone	
DNAPL Spill point (m)	(12.6, 10, 6.6)
DNAPL density (kg m^{-3})	1496
DNAPL viscosity (Pa·s)	0.965×10^{-3}
Water residual saturation (−)[a]	0.2
Brooks-Corey shape (−)	2
Total DNAPL release mass (kg)	6000
Release rate (m^3 s^{-1})	1×10^{-6}
Petrophysical model	
d_{sd} (m)	1×10^{-4}
Cementation exponent m (−)	2
Saturation exponent n (−)	2
σ_w (S m^{-1})	0.2
Porosity (−) [b]	0.37 (sand), 0.56 (clay)
B (m^2 V^{-1} s^{-1}) [c]	1.65×10^{-9}
λ (m^2 V^{-1} s^{-1}) [c]	1.35×10^{-10}
ρ_g (kg m^{-3})	2650
CEC (C kg^{-1})	1.45 (sand), 18,652 (clay) [b]
$D_{(+)}^{s}$ (m^2 s^{-1}) [d]	3.8×10^{-12}

[a] (Koch and Nowak, 2015).
[b] (Revil et al., 2021).
[c] (Revil, 2012).
[d] (Revil, 2013).

The domain below the water table was initially saturated with water and waiting for DNAPL invasion.

This aquifer analog-based DNAPL SZ simulation aims to assess the optimized survey design under more complex SZs conditions with practical significance, rather than simple blocky targets as used in previous optimization studies. Although there are some simplifications in the setting such as the neglect of the effects of DNAPL dissolution and external groundwater flow. In fact, the effect of dissolved DNAPL phase on the electrical signal is negligible due to their lack of charge carriers (Sauck, 1998; Song et al., 2024). Even if the dissolved phase were considered, their conductive characteristics would enhance the contrast between the resistive DNAPL SZ and the backgrounds (Power et al., 2014), which do not affect the applicability of DCIP surveys on these sites.

A progressive release of 6000 kg DNAPL mass at a randomly

determined spill center located at $x = 12$ m, $y = 10$ m, and $z = -0.4$ m (level with the water table) was simulated by the SIP algorithm. During the release, datasets of DNAPL distribution and saturation were recorded at 100 kg, 500 kg, 1000 kg, 2000 kg, 4000 kg, and 6000 kg individually. These DNAPL distributions were then combined in reverse order to reconstitute the evolution of a DNAPL SZ from 6000 kg over six time steps (T0 to T5). The DNAPL saturation (S_n) of each time step within the aquifer analog is depicted in Fig. 2. Due to the strong longitudinal heterogeneity within this site, DNAPL is trapped immediately at the near-surface low permeability (clay-rich) layer and formed rounded pool structures. It is noted that these subtracting following-up DNAPL release and distribution processes may not fully reflect actual DNAPL SZ remediation, but serve as synthetic numerical simulations of DNAPL SZ evolution. For example, the organic compound dissolution and the corresponding effect of DNAPL morphology change on the petrophysical relationships (Qiang et al., 2024) are not considered. These simplifications are reasonable here because the ultimate goal of this work is to evaluate the performance of DCIP survey design.

2.1.4. Determination of DCIP characteristics

Petrophysical models were employed to determine the electrical conductivity and chargeability of the aquifer analog with DNAPL SZ from the hydrogeological parameters. In addition to the electrical properties of the pore fluid, contributions from the electrical double layer (EDL) (Revil, 2012) were incorporated into the calculation of the parameters. The bulk electrical conductivity σ (Ω-m) was derived from the low-frequency conductivity model (Revil, 2013; Revil et al., 2017) as follows:

$$\sigma = \frac{1}{F} S_w^n \cdot \sigma_w + \frac{1}{F \cdot \phi} S_w^{n-1} \cdot \rho_g \cdot (B - \lambda) \cdot CEC \tag{3}$$

while the chargeability M (mV/V) was calculated using the equation (Revil et al., 2018):

$$M = \frac{\rho_g \cdot \lambda \cdot CEC}{\sigma_w \cdot \phi \cdot S_w + \rho_g \cdot B \cdot CEC} \tag{4}$$

where $F = \phi^{-m}$ (−) is the formation factor. S_w (−) denotes the water saturation, σ_w (S m^{-1}) denotes the conductivity of the pore water. ρ_g (kg m^{-3}) is the density. B and λ (m^2 V^{-1} s^{-1}) denote the apparent mobility of the counterions for surface conduction and polarization, respectively. The cation exchange capacity (CEC, C kg^{-1}) for each cell in the domain could be determined by the clay content in weight φ_w as (Patchett, 1975; Revil et al., 1998):

$$CEC = CEC_{sd} \cdot (1 - \varphi_w) + CEC_{cl} \cdot \varphi_w \tag{5}$$

where CEC_{sd} and CEC_{cl} denote the CEC of pure sand and clay, respectively. The clay content in weight φ_w and volume cl can be related through the grain densities of pure clay and sand as well as the clay porosity (Rabaute et al., 2003; Revil et al., 2002). For simplicity, we assumed the φ_w here to be equal to the cl calculated by Eq. (1), which is not true in an absolute way but is a common practice in some numerical DCIP studies (Kang et al., 2023; Almpanis et al., 2021). Finally, the normalized chargeability of each cell M_n (S m^{-1}) was determined by $M_n = M \bullet \sigma$.

According to Revil (2013), the characteristic relaxation time of a partially saturated clayey material can be determined by:

$$\tau = \frac{\Lambda^2}{2D^s_{(+)}} S_w^n \tag{6}$$

where $D^s_{(+)}$ (m^2 s^{-1}) denotes the diffusion coefficient of the counterions, Λ (m) is characteristic pore size. Following the relationship proposed by Johnson et al. (1987), Λ can be estimated from the permeability k using the formation factor F:

$$k = \frac{\Lambda^2}{8F} \tag{7}$$

which yields a direct determination of the relaxation time from the permeability (Revil, 2013; Revil et al., 2012b) by:

$$\tau = \frac{4kF}{2D^s_{(+)}} S_w^n \tag{8}$$

The parameters used in the calculations of DCIP characteristics are presented in Table 1. All these equations have a good correlation with the experimental and field observations and have been used in various hydrogeological studies involving DCIP. Finally, the numerical domain that contains aquifer analog, DNAPL saturation, and electrical parameter information is defined as a 'virtual site'.

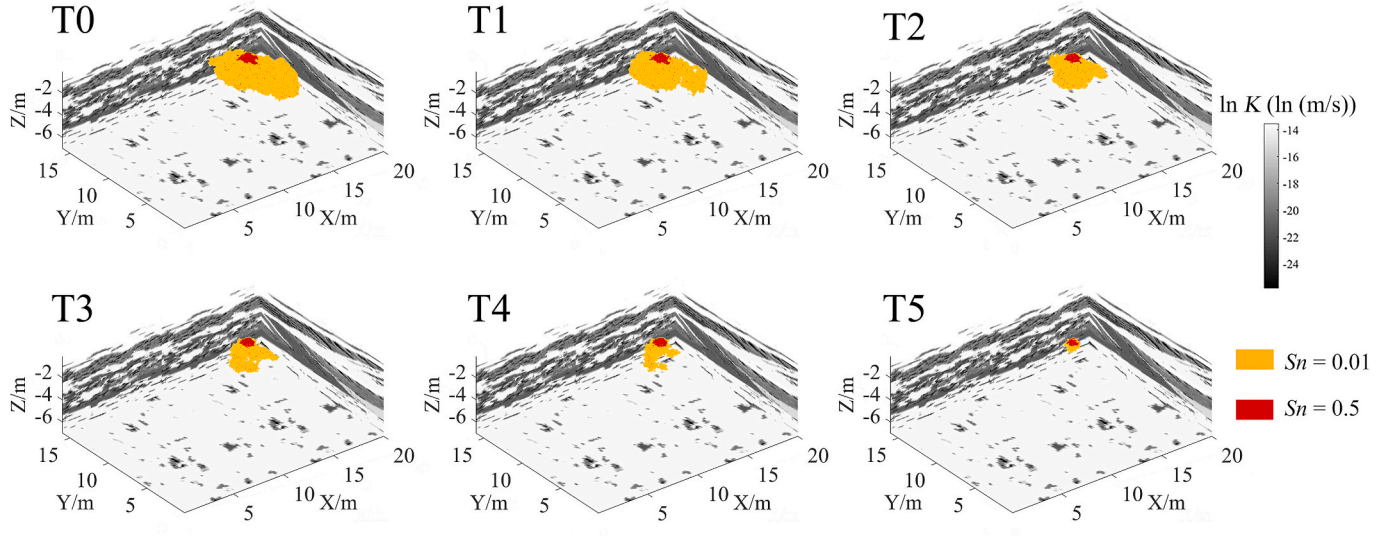


Fig. 2. DNAPL distribution at different time steps shown as the isosurface of $S_n = 0.01$ and $S_n = 0.5$. The spill center is located at $x = 12$ m, $y = 10$ m, and $z = -0.4$ m. The darker area in the hydraulic conductivity distribution denotes higher clay content. T0 signifies the DNAPL SZ before evolution (a case where DNAPL release mass = 6000 kg). Notably, the DNAPL formed a pool structure on the top low permeability (clay-rich) layer.

2.2. DCIP survey implementation

2.2.1. DCIP survey settings

The DCIP surveys performed in this study were completed with a survey grid consisting of nine parallel lines with interline spacing of 2 m (*y*-direction) and 21 electrodes along each line with inline spacing of 1 m (*x*-direction) (see Fig. 3). The 20 m × 16 m × 7 m virtual site was implemented within a larger 300 m × 300 m × 30 m domain to mitigate the boundary effect, this perimeter zone was assigned with background parameters indicated as $\sigma = 0.015$ S m^{-1}, M= 8 mV/V, $\tau = 0.4$ s, which were consistent with their common values of sand dominated cells within the virtual site.

2.2.2. Optimized survey design method and survey types evaluated

The electrodes used to inject/accept electrical current and measure potential for each DCIP measurement are controlled by the survey design, which influences the subsurface resolution distribution. The Bayesian experimental design method employed to generate an optimized survey selects the measurements that have high relative entropy (the difference between prior and posterior parameter distributions caused by a measurement) and minimizes the predicted post-experimental uncertainties on model parameters, while satisfying some other constraints (Lindley, 1956). The expected relative entropy $U_{(s)}$ for each measurement in the DCIP survey was estimated according to Bayes' rule (Huan and Marzouk, 2013):

$$U_{(s)} \approx \frac{1}{n_{out}} \sum_{i=1}^{n_{out}} \left\{ ln\left[p\left(\mathbf{d}^i|\mathbf{m}^i, s\right)\right] - ln\left[p\left(\mathbf{d}^i|\mathbf{s}\right)\right] \right\}, \tag{9}$$

where $\mathbf{m}^i$ is the Monte Carlo (MC) samples randomly generated with prior information of the model parameters $\mathbf{m}$, $\mathbf{d}^i$ are the simulated observations contaminated by measurement error, and n_{out} is the number of samples in the outer loop of the MC simulation, which is used to calculate the likelihood function $p\left(\mathbf{d}^i|\mathrm{m}^i, \mathrm{s}\right)$. The evidence factor $p\left(\mathbf{d}^i|\mathbf{s}\right)$ can be estimated by the following formula:

$$p\left(\mathrm{d}^i|\mathrm{s}\right) \approx \frac{1}{n_{in}} \sum_{j=1}^{n_{in}} p\left(\mathbf{d}^i|\mathbf{m}^{ij}, \mathbf{s}\right), \tag{10}$$

where n_{in} is the number of samples used in the inner loop of MC simulation. To alleviate the computational cost of the MC simulation, the same samples for both the outer and inner loop of MC simulation were used, and consequently $n_{in} = n_{out}$ (Huan and Marzouk, 2013). For the DCIP measurement with multiple types of model parameters (conductivity, chargeability, and relaxation time) and observations, conductivity and the corresponding voltage difference between electrodes were chosen as the model parameters $\mathbf{m}$ and observations $\mathbf{d}$, respectively in Eq. (9) to calculate the relative entropy of each measurement due to their computation rapidity. The principle of superposition was also used here to reduce the number of forward modelings needed for each optimization (see details in Qiang et al., 2022).

The implementation of Bayesian optimized survey design usually necessitates a specific area of interest as the survey focus target. The determination of the target area range for each site typically relies on the prior site information. For example, we need to investigate potential contaminant zones through the site's history of production and facility operations. Borehole sampling and preliminary geophysical surveys can also provide prior subsurface information to aid in target area identification. For the virtual site used in this study, we assume a scenario where the only preliminary information available is the location of the spill point. This is a common situation we might encounter in realistic field works, such as the detection of contaminant plume range from a known leakage. Given the uncertainty of the subsurface properties and processes, a vague and large potential contaminated area is often determined for these sites with limited historical information. Thus, to restore the real-world scenario, we chose an 8 m × 8 m × 5 m target area (as illustrated in Fig. 3) starting at $x = 8$ m, $y = 6$ m, $z = -5$ m that surround the DNAPL spill center for the focused optimized survey. This target area occupies nearly a quarter of the entire model, which is relatively large compared with the DNAPL SZ generated in the same aquifer analog with much higher (original) permeability (Maji and Sudicky, 2008; Maji et al., 2006) and the DNAPL SZ encountered in most field or synthetic surveys (Almpanis et al., 2021; Wang et al., 2014; Kang et al., 2023). Note that the target area here only encompasses a potential area where the target DNAPL can distributed, rather than an exact inclusion of the potential DNAPL SZ, as the resolution could still be improved at cells outside the target area as they are located closely.

We may also encounter cases where little or vacant historical site information is available. If we still assign the survey design with a specific target area, locations far away from these selected zones will have lower resolution and poor imaging quality due to the absence of data points (lower than the conventional survey), hence limiting the investigation efficiency. To address this, we propose two strategies here. The first routine is the unfocused optimization strategy, which was

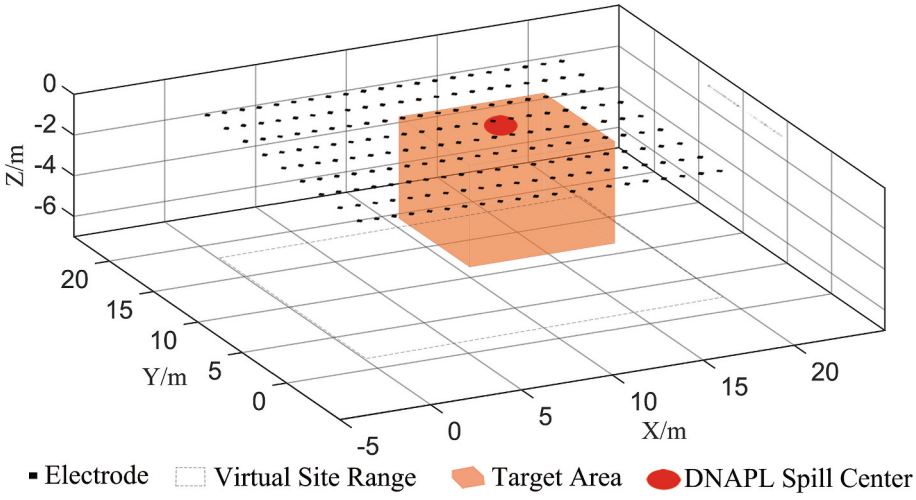


Fig. 3. DCIP survey settings used in this study. An additional surrounding 300 × 300 × 30 m domain outside of this model was employed to mitigate the boundary effect. The black dots indicate the positions of electrodes, the black dashed line shows the range of the virtual site incorporated. The orange cube is the target area for the focused optimized survey design and the red dot illustrates the DNAPL spill center. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

applied without defining a specific target area. Thus, the configurations with the highest entropy across the entire model are selected. The resulting optimized survey is then supposed to enhance the resolution throughout the subsurface. The second routine is to replace some of the measurements in the focused optimized survey with conventional measurements that cover the entire domain, which serves as compensation for the resolution outside the target area.

Meanwhile, as the magnitude of the IP signal in field measurements is relatively lower than DC survey (see Kemna et al., 2012; Zarif et al., 2017), the optimization should be also aimed at improving the signal-to-noise ratio of the DCIP survey. To implement this, we introduced voltage-dependent noise to the simulated observations in Eq. (9). Compared with traditional Gaussian noises, it comprises an extra absolute minimum noise that indicates the constant background noise. This type of noise can reduce the relative entropy of the measurements with small measure voltage magnitude, hence improving the signal-to-noise ratio of the selected optimized measurements.

The measurements in the optimized surveys are usually selected from a comprehensive measurement set that contains all the possible independent measurements (Wilkinson et al., 2006). However, there are $N(N-1)(N-2)(N-3)/8$ possible independent measurement for a survey with N electrodes (Xu and Noel, 1993), which is 1.5×10^8 for the survey grids we used. The computation and storage burden when considering the complete comprehensive measurement set is unacceptable in this study. Therefore, we chose a reduced comprehensive measurement set with 842,091 measurements as the pending library of the optimized survey. This comprehensive set comprising inline 4-pole Alpha type (Wenner, Wenner-Schlumberger) and Beta type (Dipole-Dipole) arrays with their offset version (electrodes A, B, M, N located at different survey lines), as well as the equatorial Dipole-Dipole arrays (see Fig. 4). From the comprehensive set, four types of measurement sets were constructed:

1. The *conventional survey* consists of 3486 inline Wenner, Wenner-Schlumberger, and Dipole-Dipole measurements along both the x- and y-direction, with a maximum geometric factor of 660 m. This low geometric factor constraint ensures a mean investigation depth around −1.5 m (Edwards, 1977) that proportional to DNAPL SZs depth in the conceptual model. The conventional survey was generated to simulate the conventional DCIP monitoring approach at the virtual site, providing a reference for comparison.
2. The *focused optimized survey* comprises 3486 measurements which have the highest relative entropy in the comprehensive measurement set when the target area (8 m × 8 m × 5 m) is set around the spill center. Note that the term "target area" throughout this paper indicates this small area for the focused optimized survey design
3. The *unfocused optimized survey* comprises 3486 measurements which have the highest relative entropy in the comprehensive measurement set for the entire model space
4. The *optimized + base survey* replaced 20 % measurements in the focused optimized survey with the randomly selected conventional (base) measurements to make up for the resolution outside the target area. As a result, the *optimized + base survey* comprises 2948 optimized measurements with the highest relative entropy and 736 conventional measurements

2.2.3. Forward modeling and inversion of DCIP characteristics

The forward modeling of apparent conductivity for the virtual site utilized an elliptic equation computed using the finite element method by COMSOL Multiphysics® (Version 5.4):

$$\nabla\cdot(\sigma\nabla V) = -I\delta(\mathbf{r}) \tag{11}$$

where δ denotes the Dirac delta function and $\mathbf{r}$ is a single current electrode which is a point source at the origin with current strength I (A).

For the IP effect, both intrinsic and integrated (an integration of the whole voltage decay curve) chargeability should be acquired in this study to calculate the inverted relaxation time tomography. Therefore, the time domain IP effects were solved by assuming it a time-dependent self-potential-type problem (Soueid and Revil, 2018), which allows time-saving in the numerical acquisition of the whole secondary voltage decay curve after the primary current is shut down. The total current density $\mathbf{J}$ in the simulation domain is written as (Seigel, 1959):

$$\mathbf{J} = [1 - M(t)]\mathbf{J}_p \tag{12}$$

where $\mathbf{J_p}$ denotes the secondary current density. $M(t)$ denotes the time-dependent chargeability. From a Debye model (Revil et al., 2014) we have:

$$M(t) = M\left[1 - exp\left(-\frac{t}{\tau}\right)\right] \tag{13}$$

After forward modeling, Gaussian random noise with a standard deviation of 5 % of the absolute value was added to the apparent conductivity and chargeability prior to inversion. The conductivity was inverted using the standard smoothness-constrained least-squares method (L2-norm, Loke et al., 2003):

$$\left(\mathbf{G}^{\mathrm{T}}\mathbf{G} + \lambda\mathbf{C}\right)\Delta\mathbf{r_i} = \mathbf{G}^{\mathrm{T}}\mathbf{d} - \lambda\mathbf{C}\mathbf{r_{i-1}} \tag{14}$$

where $\mathbf{G}$ is the Jacobian matrix, λ denotes the damping factor. $\Delta\mathbf{r_i} = \mathbf{r_i} - \mathbf{r_{i-1}}$ is the model parameters change vector. $\mathbf{d}$ denotes the difference between the observed data and the model response (the discrepancy vector). $\mathbf{C}$ contains the roughness filter constraint. The intrinsic and

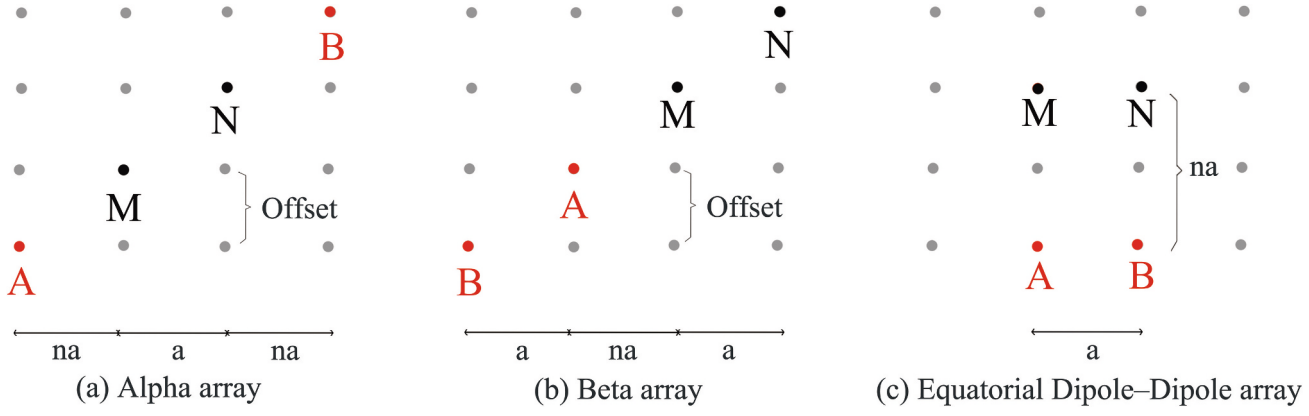


Fig. 4. The sketch of array configurations in the comprehensive measurement set. (a) Alpha type array, (b) Beta type array, and (c) Equatorial Dipole-Dipole array. The electrode spacings are controlled by the dipole length *a*, dipole separation *na*, and the number of *offset* lines, which vary among different electrodes and measurements.

integrated chargeability were then inverted using a linear approach based on the conductivity model (Oldenburg and Li, 1994). The inverted normalized chargeability was determined by the inverted $M \bullet \sigma$.

The relaxation time was usually inverted by fitting complex conductivity models (Loke et al., 2006) for each cell in the domain, which may lead to computation burdens. In this work, relaxation time distribution was approximated in a more computationally effective way by the ratio of the acquired integrated chargeability to the intrinsic chargeability tomography cell by cell through the Debye model:

$$\tau \approx \frac{M_i}{M} = \frac{\int_0^\infty M[1 - exp(-t/\tau)]\mathrm{d}t}{M} \tag{15}$$

3. Results

3.1. Resolution of the conventional and optimized surveys

The resolution of a geoelectrical survey represents its capability to delineate the subsurface characteristics (Day-Lewis et al., 2005). Fig. 5 illustrates the resolution distributions of the conventional and optimized surveys of each model cell, which is estimated using the diagonal elements of the resolution matrix **R** (Loke et al., 2010) calculated by:

$$\mathbf{R} = \left(\mathbf{G}^{\mathrm{T}}\mathbf{G} + \lambda\mathbf{C}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{G} \tag{16}$$

where $\lambda = 1 \times 10^{-6}$ (−). The Jacobian matrix **G** of a 4-pole measurement is derived by aggregating the 3-D Frechet derivatives (Mcgillivray and Oldenburg, 1990) for 2-pole measurements where one electrode is equivalently located in (a,0,0):

$$F_{3D}(x,y,z) = \frac{1}{4\pi^2}\cdot\frac{x(x-a)+y^2+z^2}{\left[x^2+y^2+z^2\right]^{1.5}\left[(x-a)^2+y^2+z^2\right]^{1.5}} \tag{17}$$

where x, y, and z are the relative coordinates of the other electrode in the 2-pole array. The Jacobian matrix of these 2-pole measurements is then used to calculate the 4-pole measurement Jacobian matrix with the superposition principle.

Figs. 5a and b show that the subsurface resolution from the conventional and unfocused optimized survey is distributed uniformly and decreases with depth, despite the resolution for the unfocused optimized survey being higher throughout the domain. In comparison, the resolution of the focused optimized + base and focused optimized survey is clearly confined around the target area, reaching almost 1 near the surface (< 0.5 m) and exhibiting significant resolution enhancement throughout the target area. The mean resolution for cells within the target area at −2 m depth acquired by the focused optimized survey (0.23) is almost 3.4 times larger than that by the conventional survey (0.068), following the unfocused optimized survey with 0.18 and the optimized + base survey with 0.21 at the same range.

Fig. 6 presents the mean resolution values across various areas. The focused optimized survey shows nearly 2 times higher mean resolution within the target area than the conventional survey. This does involve a small (but worthwhile) cost because the resolution of the focused optimized survey outside the target area is slightly lower than that of the conventional survey (see Fig. 6b). Compared with the focused optimized survey, the unfocused optimized survey could provide a moderated resolution enhancement in the target area around the spill center, and the highest mean resolution throughout the entire model. While the focused optimized + base survey provides a remarkable resolution enhancement within the target area and also improves a bit the resolution throughout the entire area. Interestingly, the three optimized surveys turn out to have lower geometrical factor than the conventional survey (see Fig. 6d). This is caused by the specific array types in these surveys. For example, 75.95 % of the measurements in the focused optimized survey are Wenner-Schlumberger (alpha) or equatorial Dipole-Dipole array with dipole length $a \geq 5$, and dipole separation $n \leq 0.4$. These arrays share the feature that the distance between two current electrodes is much larger than the distance between the current and potential electrodes, which corresponds to very strong signal strength and small geometrical factors. These measurements are easier to be selected for the optimized survey because of the voltage-dependent Gaussian noise (Qiang et al., 2022) added to the simulated observations in Eq. (9), which assigns higher relative entropy to measurements with better signal-to-noise ratios.

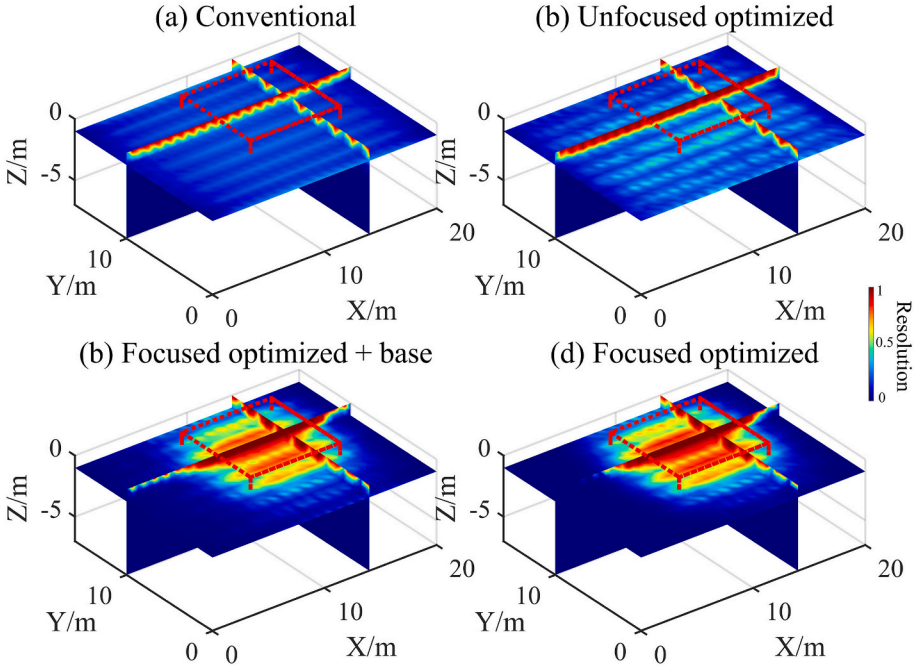


Fig. 5. The subsurface resolution distribution for (a) conventional survey, (b) unfocused optimized survey, (c) focused optimized + base survey, and (d) focused optimized survey. The red dash line delineates the target area for the focused optimized surveys. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

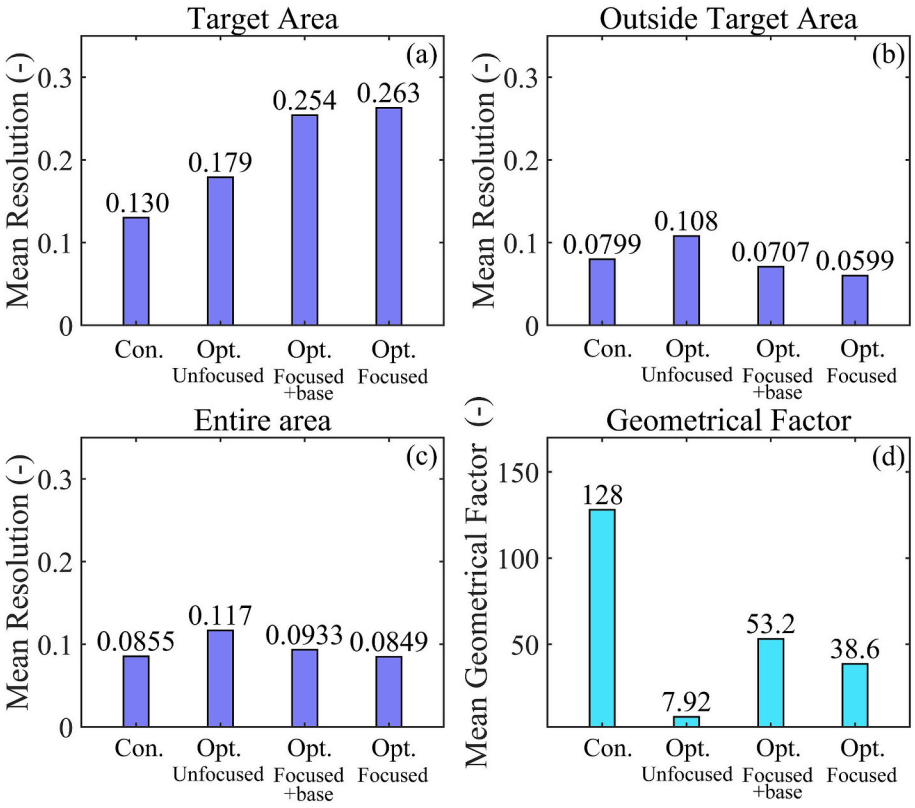


Fig. 6. Mean resolution and geometrical factor for different surveys. (a) Mean resolution within the target area, (b) mean resolution outside the target area, (c) mean resolution of the entire area. Correspondingly, (d) illustrates the mean geometrical factor for different surveys. 'Con.' refers to conventional survey and 'Opt.' refers to optimized surveys with different features as follows below.

3.2. Static DCIP survey

Fig. 7 presents the true and inverted conductivity, chargeability, normalized chargeability, and relaxation time distributions acquired by the conventional survey and optimized surveys for the virtual site at T0 (i.e., 6000 kg DNAPL mass). The conventional survey can roughly outline the DNAPL SZ via different tomography, while the optimized surveys could better locate the DNAPL SZ with more detailed and obvious structural information. In particular, the focused optimized surveys (row 5) could reveal a bit of the DNAPL SZ heterogeneity with more accurate values of different DCIP characteristics within SZ, whereas the conventional survey shows the DNAPL SZ as a blurred anomaly.

Fig. 8 presents the cross-sectional images at $y = 10$ m of the true and inverted conductivity, chargeability, normalized chargeability, and relaxation time fields at T0. The inverted conductivity fields acquired by both the conventional and optimized surveys effectively identify the DNAPL SZ, but reveal limited background stratigraphic information. In contrast, the IP characteristics such as chargeability and normalized chargeability, better delineate the spatial boundaries between the top sand and below clay-rich layers (identified by dashed gray lines). Due to the poor resolution outside the target area, the focused optimized survey failed to capture the complete stratigraphic boundary of the clay-rich zones, which is rather evident in the chargeability, normalized chargeability and relaxation time tomography. The issue is resolved by the unfocused optimized and focused optimized + base survey, which provide clearer stratigraphic boundaries with their sensitivity across the entire domain while keeping resolution enhancements for the DNAPL SZ targets. Meanwhile, the conventional survey seems to overestimate the chargeability of the clay-rich layer below the DNAPL SZ, making the high chargeability center improperly shift away from the original DNAPL SZ to the underlying clay layers.

3.3. DCIP survey for DNAPL SZ evolution

Fig. 9 shows the true DNAPL SZ evolution (row 1), inverted conductivity (rows 2–4), and inverted chargeability (rows 5–7) for the unfocused and focused optimized survey at all time steps. All surveys could successfully characterize the DNAPL SZ in the first stage. While some changes indicated by the conventional survey are not as evident as the optimized surveys. For instance, the conventional survey almost loses sight of the DNAPL SZ in conductivity tomography T1 and chargeability tomography T2 (indicated as the red dashed box), which is still obvious in the two optimized surveys.

An indicator called misfit (Loke et al., 2015) is used here to quantitatively evaluate the differences between the true model parameters $\boldsymbol{c_t}$ and the inverted model $\boldsymbol{c_i}$:

$$\text{Misfit} = \left(\frac{1}{N} \sum_{i=1}^{N} [log(\boldsymbol{c_t}(i)) - log(\boldsymbol{c_i}(i))]^2 \right)^{0.5}. \quad (18)$$

The value of misfit is dependent on the magnitude of the evaluated parameter so it can only be compared when evaluating the same parameters. The misfits of the inverted characteristics within the target area acquired by the different surveys across the six time steps are shown in Fig. 10 and normalized by the maximum value of each parameter themselves, respectively. For all time steps evaluated, the misfit for the conductivity, chargeability, and normalized chargeability acquired by the focused optimized surveys within the target area decreased by an average of 8 %, 19 %, and 12 %, respectively, compared to the conventional survey. For the relaxation time tomography, although the optimized surveys provide an improved qualitative delineation of the DNAPL SZ range above the clay-rich area (see Fig. 8, row 2–5, column 4, above the dashed lines), its misfits in Fig. 10 are abnormally larger than those acquired by the conventional survey. This anomaly is largely due

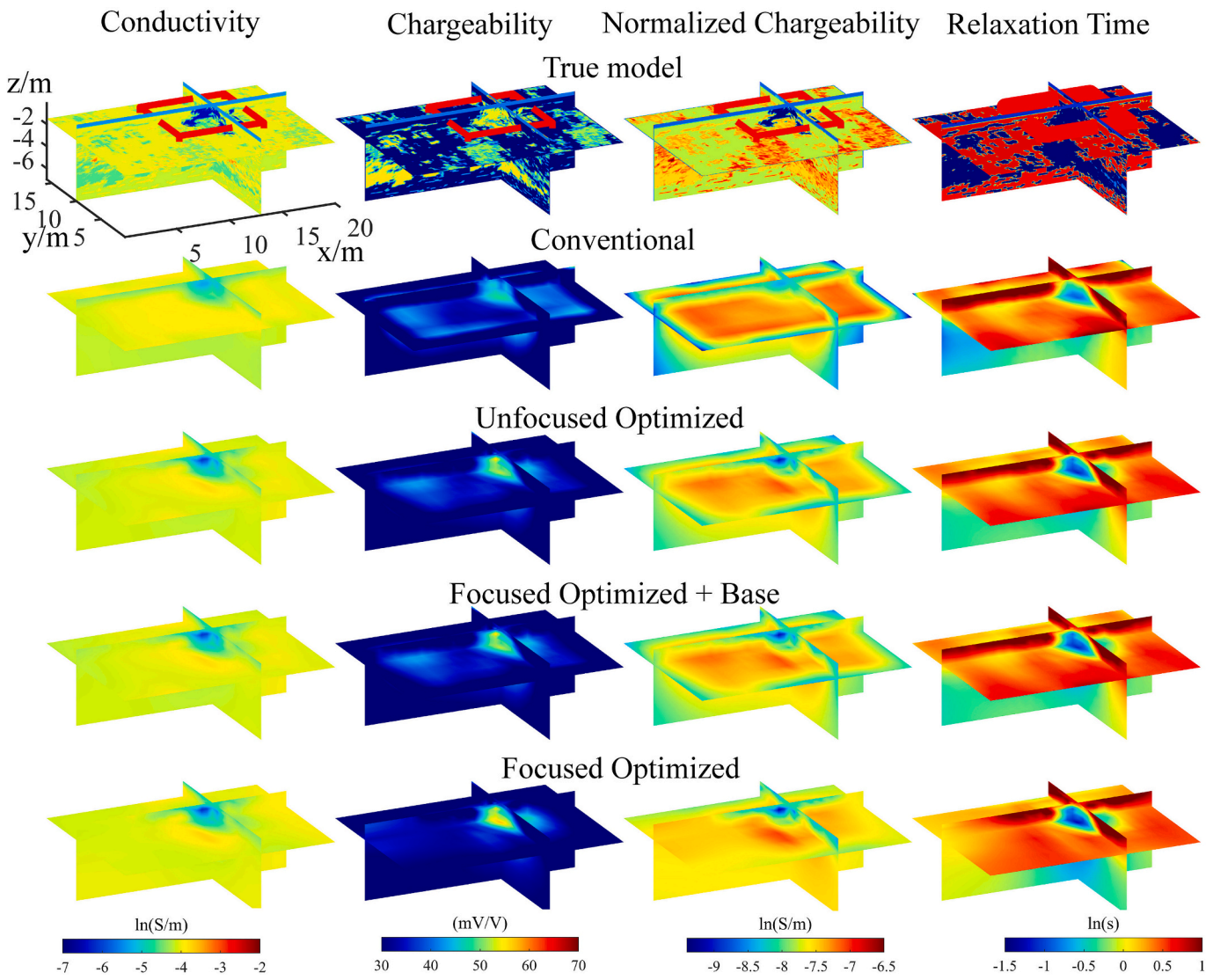


Fig. 7. True and inverted conductivity, chargeability, normalized chargeability, and relaxation time fields at the initial time step T0 for conventional, unfocused optimized, focused optimized + base, and focused optimized surveys. The target area for the focused optimized surveys is indicated as red box in the true models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the overestimation of the relaxation time value during inversions. This shows that the indicator misfit may not always be a good representative of the image due to its extremely high sensitivity to the magnitude of the parameters, while neglecting the overall conformity in distribution. Nevertheless, the significant difference between the values of the true and inverted relaxation time would indeed invalidate most similarity indicators, even the structural similarity index metrics (SSIM, see Han et al., 2024b).

Fig. 11 presents the change ratio of conductivity, chargeability, normalized chargeability, and relaxation time in logarithm based on the inversions at T5 and T0. These proportional images remove the 'constant' background geology (i.e., aquifer structure), allowing all DCIP characteristics to highlight the changes associated with the DNAPL SZ evolution. It is then evident from Fig. 11 that the DNAPL was trapped above the clay rich layer (identified by dashed gray lines). The optimized surveys reveal the change ratios of various DCIP characteristics that more closely approximate their true values, while the conventional survey shows varying degrees of underestimation. This deviation is particularly pronounced in normalized chargeability (M_n) tomography, whereas the conventional survey almost loses sight of the M_n changes. This is because M_n, as the product of conductivity (decrease with time steps) and chargeability (increase with time steps), does not exhibit as significant change ratios as other characteristics in nature, making it more difficult to detect. Meanwhile, M_n results have to involve the uncertainties from both the conductivity and chargeability inversions, further contributing to poor characterization quality. Fig. 11 also provides insight into why the optimized surveys have larger misfits for relaxation times compared to the conventional survey, as the optimized surveys seem to overestimate the relaxation time changes of the area below the DNAPL SZs.

4. Discussion

4.1. Sensitivity of different DCIP characteristics to survey resolution

The enhanced extent of target area characterization accuracy attributed to the optimized surveys appears to vary with different DCIP characteristics. As in Fig. 10, the mean misfit reduction for chargeability acquired by the focused optimized survey (19 %) is larger than for conductivity (8 %). This is because the chargeability measurement is more sensitive to the survey resolution due to its susceptibility to subsurface structures. For example, the chargeability tomography in

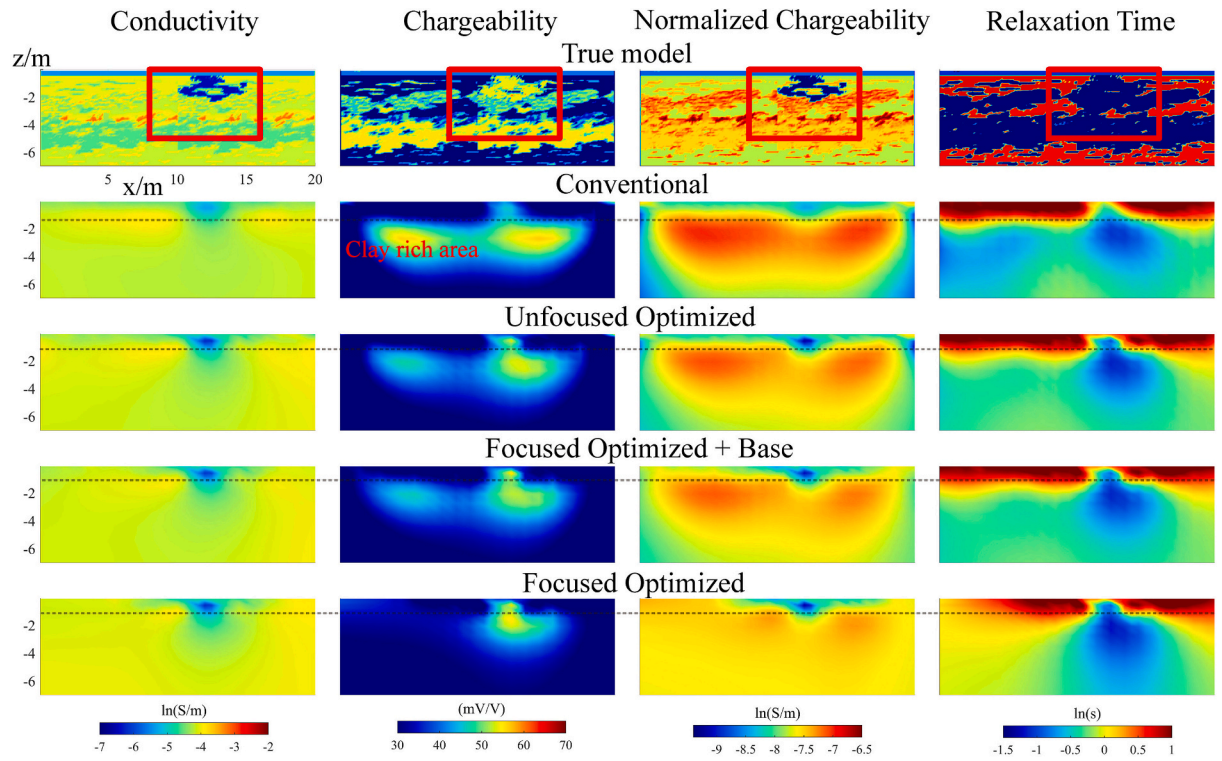


Fig. 8. Cross-sectional images of the true and inverted conductivity, chargeability, normalized chargeability, and relaxation time field at $y = 10$ m at the initial time step T0. The target area for the focused optimized surveys is indicated as a red box in the true models. The black dashed lines in the inverted images indicate the boundary between the sand and clay-rich area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

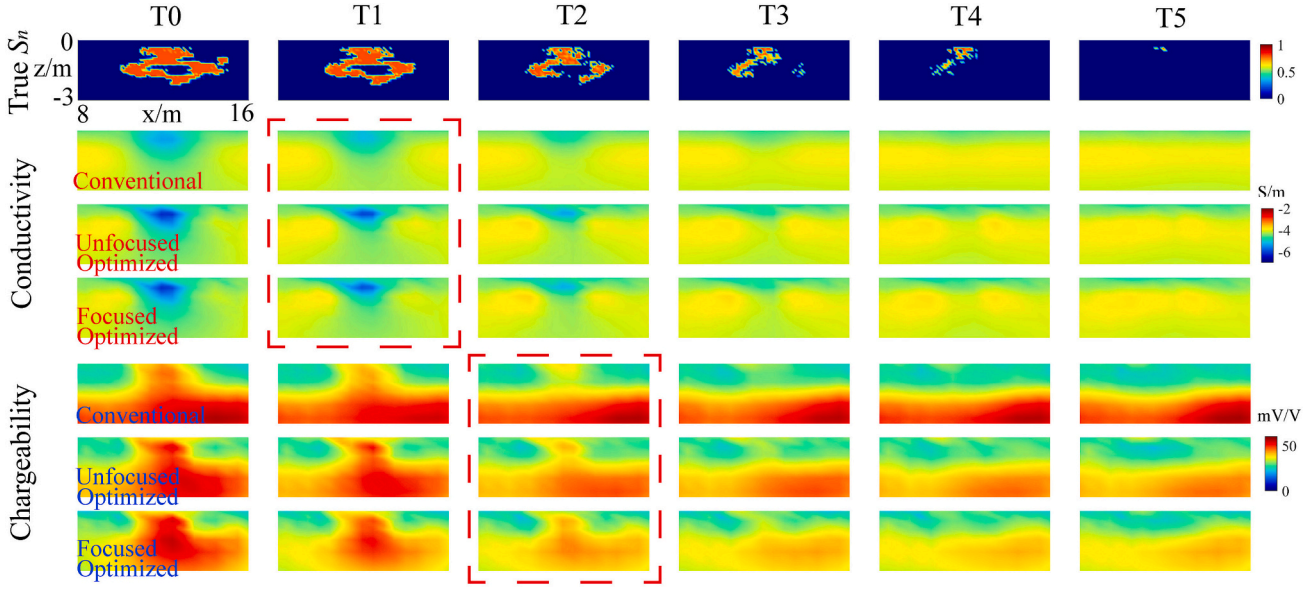


Fig. 9. Cross-sectional images at $y = 10$ m of the true and inverted conductivity and chargeability fields acquired at different time steps. To better highlight the details of the DNAPL SZ, the fields are magnified to show an 8 m × 3 m region within the target area. The dashed red boxes indicate the stages where the conventional surveys start to lose sight of the DNAPL source zone. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

conventional survey confused the high chargeability area caused by DNAPL SZ and the clay-rich layer, and overestimates the chargeability values at around −2 m, while the optimized surveys with higher resolutions could better characterize the chargeability values of both DNAPL SZ and the clay-rich layer (see Fig. 8). Meanwhile, all the misfit for conductivity data in Fig. 10 decreases with time steps, indicating that the conductivity signal is mainly controlled by the DNAPL SZ mass decay. While chargeability is more sensitive to the background aquifer information, their misfit is almost fixed over time since the background aquifer contributes more to the electrical signals than the DNAPL SZs. The normalized chargeability, acquired by the product of conductivity and chargeability, exhibits a neutral sensitivity to both the DNAPL SZ

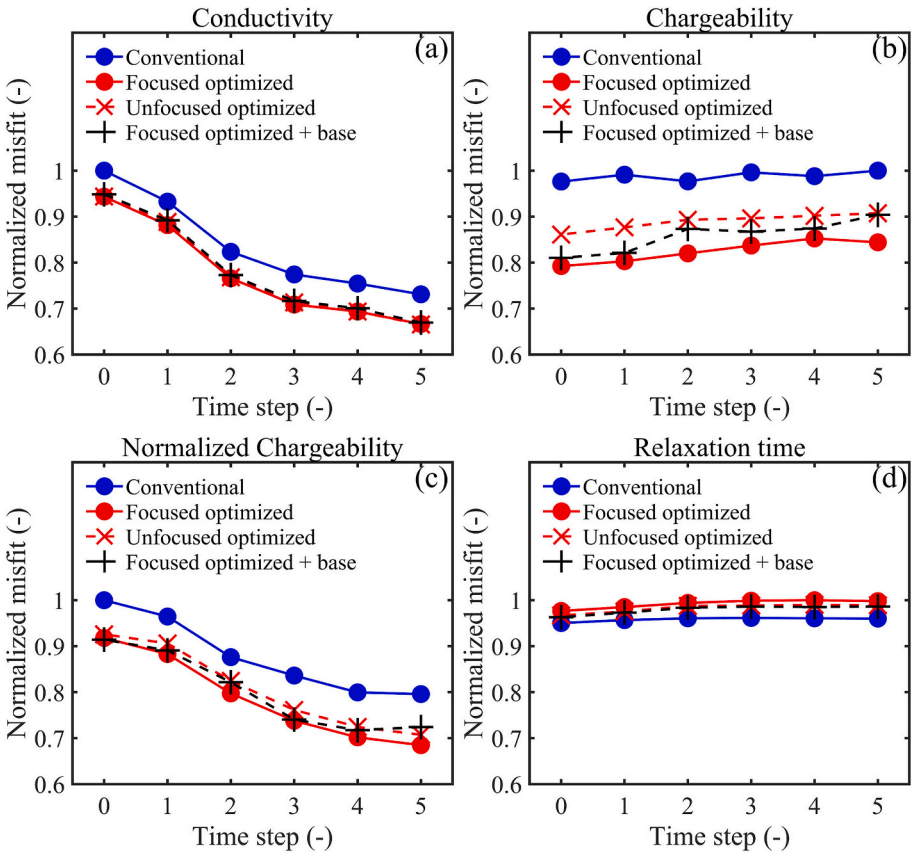


Fig. 10. The normalized data misfit within the target area for different characteristics and surveys across the six time steps. (a) Inverted conductivity, (b) chargeability, (c) normalized chargeability, and (d) relaxation time. The misfits are normalized by the maximum value of each parameter series themselves, respectively.

and the background information, resulting in a moderated mean misfit reduction for the optimized surveys (12 %).

The distinct sensitivity of different characteristics to resolutions also leads to the different performance of the three types of optimized surveys. The focused optimized survey shows a similar misfit reduction in conductivity estimations to the unfocused optimized and the focused optimized + base surveys, but a more significant misfit reduction in the chargeability and normalized chargeability tomography. These outcomes underscore the importance of optimized survey design in DCIP measurements, especially for the sites with complex subsurface conditions that demand higher data quality and resolution.

4.2. *Influence of optimized survey design on DCIP survey duration*

Survey duration is an important factor to consider for DCIP surveys at field sites since it is usually proportional to the human and economic costs. Prolonged measuring time can also lead to image blurring if monitoring quickly changing or moving targets (Wilkinson et al., 2015). The optimized survey can provide equal or higher subsurface resolution within the target area compared to a conventional survey, even though it uses much fewer measurements and therefore much shorter survey duration. Fig. 12 presents the mean resolution associated with different survey types and the number of measurements used in a DCIP survey at the virtual site. Since the conventional survey only contains in-line Wenner, Wenner-Schlumberger, and Dipole-Dipole measurements, the number of measurements in the conventional survey is limited for a certain electrode grid (e.g., 21 × 9 electrodes in this study). To show the benefits of an optimized survey over a larger range of the measurement numbers used in a DCIP survey, an *arbitrary survey* is introduced that made by measurements selected randomly from the comprehensive set, while the optimized survey selects the measurements with the largest relative entropy under the focused (with target area 8 m × 8 m × 5 m around the spill center) and unfocused conditions.

As shown in Fig. 12, the optimized surveys exhibit higher mean resolution compared to the conventional or arbitrary survey across varying numbers of measurements. For example, the focused optimized survey with 3486 measurements achieves a mean resolution of 0.263 within the target area, while the arbitrary survey can only obtain a similar resolution with more than 21,000 measurements, which is an ~83 % difference. This reduction of measurements is less significant for the unfocused optimized survey (Fig. 12b), but it could still save up to 61 % of measurements for the same mean resolution throughout the entire model space compared to the arbitrary survey. Assuming a single DCIP measurement in the field usually takes 0.5 s, the focused optimized survey on the concerned domain can lower the measuring time equivalently by 2.4 h, while maintaining a similar resolution within the target area.

Fig. 12 shows that with slight or even no prior information, the proposed survey design method can provide a significant balance with the survey resolution within the target area and survey duration, hence improving the efficiency of DCIP surveys in the field. This time saving could be more and more significant in future DCIP surveys using instruments with increasing amounts of supported electrodes and possible measurements.

4.3. *Optimize survey design implementation suggestions*

Our results highlight the advantages of different types of optimized surveys. Nevertheless, it is important to note that the advantages of the specific focused optimized survey are based on a premise that the location of the concerned subsurface target is roughly known, so that it is included in the selected target area that has high resolution. For sites

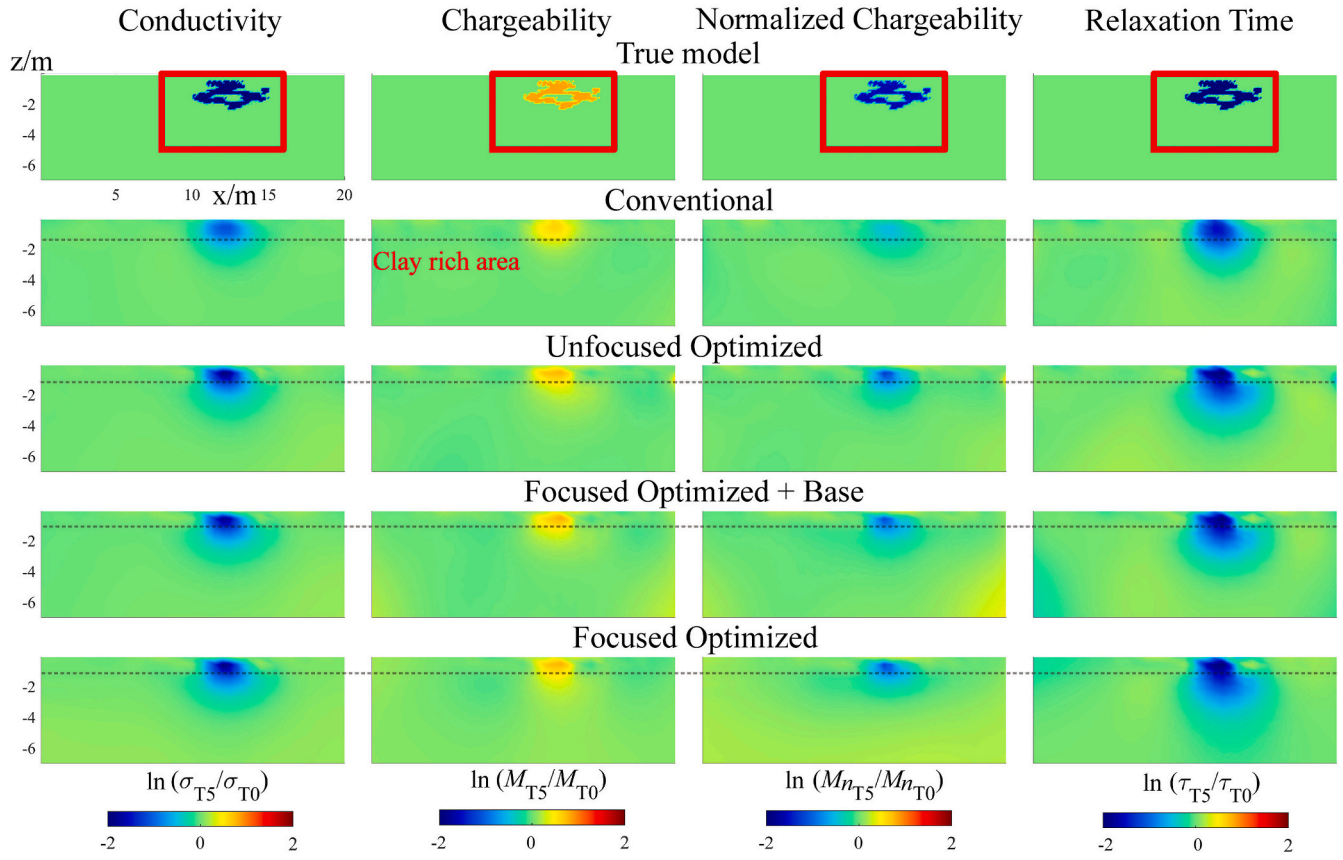


Fig. 11. Cross-sectional difference (ratio) images at $y = 10$ m of the true and inverted conductivity, chargeability, normalized chargeability, and relaxation time between T5 and T0. The target area for the focused optimized surveys is indicated as a red box in the true models. The black dashed lines in the images indicate the boundary between the sand and clay-rich area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

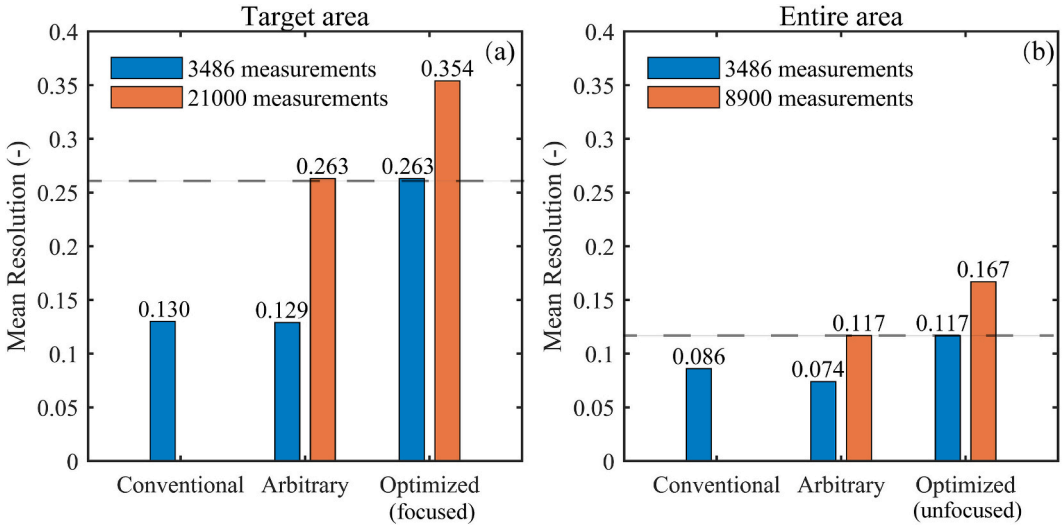


Fig. 12. The mean resolution within the target and entire area versus the survey types and number of measurements used in a DCIP survey. The *arbitrary survey* is made by measurements selected randomly from the comprehensive set. The optimized measurements are selected from the comprehensive set in order of their relative entropy with different focus ranges. The dashed gray lines indicate that different surveys provide similar resolutions but with different amounts of measurements.

with scarce preliminary information, it is recommended to define a relatively larger target area, or use the unfocused optimized / focused optimized + base survey first to prescreen the site, which can further constrain the potential targets for the subsequent focused optimized survey design.

In addition, this study neglects the possible inductive coupling between cables in DCIP survey (Wang et al., 2021), as well as the parallel configuration of survey commands for multi-channel systems. These

considerations may introduce additional principles for the optimized survey design that prioritize measurements with certain current electrode numbers or injection sequences, but would not affect the underlying advantages of the optimized survey. Nevertheless, they may be considered in future studies for efficient DCIP surveys regarding different site conditions.

5. Conclusions

In this study, numerical DCIP monitoring of DNAPL evolution at the field scale was performed to evaluate the efficiency of a Bayesian optimized survey design method. The virtual site was developed from a high-resolution heterogeneous clay-rich aquifer analog and then used to simulate a 6-stages DNAPL evolution process and simultaneous monitoring by DCIP. The results demonstrated that with the same time cost, the optimized survey with known prior target area information showed an average enhancement of 8 % to 19 % in target DNAPL SZs monitoring for various evaluated DCIP characteristics, including conductivity, chargeability, and normalized chargeability. In the meantime, to provide the same subsurface resolution, the optimized survey with a prior target can save up to 83 % in measuring time compared to conventional surveys. To further, we evaluated the optimized surveys designed with limited or no prior site information, which can provide efficient resolution enhancement throughout the entire subsurface area. The inverted results also highlight the high sensitivity of the IP characteristics to subsurface heterogeneity caused by clay distributions.

The findings in this study emphasize a broad application prospect of optimized survey design in field scale DCIP surveys, which promote the monitoring of critical zone hydrological processes such as organic contamination remediation and solute transports. Moving forward, the development and utilization of more accurate relaxation time modeling and inversion methods, together with the optimized survey design method for inductive coupling and multi-channel geoelectrical systems, warrant attention in future research endeavors.

CRediT authorship contribution statement

Siyuan Qiang: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xiaoqing Shi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **André Revil:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Xueyuan Kang:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Christopher Power:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

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

Acknowledgments

The work was supported by National Key Research and Development Program of China (2022YFC3703101) and the National Natural Science Foundation of China (42272276, 42202267, 41977157). Siyuan Qiang thanks the China Scholarship Council (202306190219) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX24_0209) for their support. We are also grateful to the High-Performance Computing Center (HPCC) of Nanjing University for doing the numerical calculations in this paper on its blade cluster system. We thank Meiyu Song and Fangzhou Du for their help in virtual site preparation. We sincerely appreciate Editor Diogo Bolster and two anonymous reviewers for their insightful comments and suggestions, which significantly improve the quality of this work.

Data availability

Data will be made available on request.

References

Almpanis, A., Gerhard, J., Power, C., 2021. Mapping and monitoring of DNAPL source zones with combined direct current resistivity and induced polarization: a field-scale numerical investigation. Water Resour. Res. 57 (11), e2021WR031366. https://doi.org/10.1029/2021WR031366.

Al-Raoush, R.I., 2009. Impact of wettability on pore-scale characteristics of residual nonaqueous phase liquids. Environ. Sci. Technol. 43 (13), 4796–4801. https://doi.org/10.1021/es802566s.

Ayaz, M., Toussaint, R., Schäfer, G., Måløy, K.J., 2020. Gravitational and finite-size effects on pressure saturation curves during drainage. Water Resour. Res. 56 (10), e2019WR026279. https://doi.org/10.1029/2019WR026279.

Bandara, U.C., Palmer, B.J., Tartakovsky, A.M., 2016. Effect of wettability alteration on long-term behavior of fluids in subsurface. Comput. Part. Mech. 3 (2), 277–289. https://doi.org/10.1007/s40571-015-0098-8.

Bayer, P., Comunian, A., Höyng, D., Mariethoz, G., 2015. High resolution multi-facies realizations of sedimentary reservoir and aquifer analogs. Sci. Data 2 (1), 1–10. https://doi.org/10.1038/sdata.2015.33.

Binley, A., Slater, L., 2020. Resistivity and Induced Polarization: Theory and Applications to the Near-Surface Earth. Cambridge University Press. https://doi.org/10.1017/9781108685955.

Day-Lewis, F.D., Singha, K., Binley, A.M., 2005. Applying petrophysical models to radar travel time and electrical resistivity tomograms: resolution-dependent limitations. J. Geophys. Res. Solid Earth 110 (B8). https://doi.org/10.1029/2004JB003569.

De Lima, O., Niwas, S., 2000. Estimation of hydraulic parameters of shaly sandstone aquifers from geoelectrical measurements. J. Hydrol. 235 (1–2), 12–26. https://doi.org/10.1016/S0022-1694(00)00256-0.

Dekker, T.J., Abriola, L.M., 2000. The influence of field-scale heterogeneity on the surfactant-enhanced remediation of entrapped nonaqueous phase liquids. J. Contam. Hydrol. 42 (2–4), 219–251. https://doi.org/10.1016/S0169-7722(99)00091-1.

Edwards, L., 1977. A modified pseudosection for resistivity and IP. Geophysics 42 (5), 1020–1036. https://doi.org/10.1190/1.1440762.

Ewing, R.P., Berkowitz, B., 1998. A generalized growth model for simulating initial migration of dense non-aqueous phase liquids. Water Resour. Res. 34 (4), 611–622. https://doi.org/10.1029/97WR03754.

Ewing, R.P., Berkowitz, B., 2001. Stochastic pore-scale growth models of DNAPL migration in porous media. Adv. Water Resour. 24 (3–4), 309–323. https://doi.org/10.1016/S0309-1708(00)00059-2.

Griffiths, D., Turnbull, J., Olayinka, A., 1990. Two-dimensional resistivity mapping with a computer-controlled array. First Break 8 (4). https://doi.org/10.3997/1365-2397.1990008.

Han, Z., Kang, X., Singha, K., Wu, J., Shi, X., 2024a. Real-time monitoring of in situ chemical oxidation (ISCO) of dissolved TCE by integrating electrical resistivity tomography and reactive transport modeling. Water Res. 252, 121195.

Han, Z., Kang, X., Wu, J., Shi, X., Jiang, J., 2024b. Improved solute transport modeling through joint estimation of hydraulic conductivity and dispersivities from tracer and ERT data. Adv. Water Resour. 104655.

Heinz, J., Kleineidam, S., Teutsch, G., Aigner, T., 2003. Heterogeneity patterns of quaternary glaciofluvial gravel bodies (SW-Germany): application to hydrogeology. Sediment. Geol. 158 (1–2), 1–23. https://doi.org/10.1016/S0037-0738(02)00239-7.

Huan, X., Marzouk, Y.M., 2013. Simulation-based optimal Bayesian experimental design for nonlinear systems. J. Comput. Phys. 232 (1), 288–317.

Johansson, S., Fiandaca, G., Dahlin, T., 2015. Influence of non-aqueous phase liquid configuration on induced polarization parameters: conceptual models applied to a time-domain field case study. J. Appl. Geophys. 123, 295–309. https://doi.org/10.1016/j.jappgeo.2015.08.010.

Johnson, D.L., Plona, T., Kojima, H., 1987. Probing porous media with 1st sound, 2nd sound, 4th sound, and 3rd sound. In: Paper Presented at AIP conference proceedings. American Institute of Physics. https://doi.org/10.1063/1.36398.

Kang, X., Shi, X., Deng, Y., Revil, A., Xu, H., Wu, J., 2018. Coupled hydrogeophysical inversion of DNAPL source zone architecture and permeability field in a 3D heterogeneous sandbox by assimilation time-lapse cross-borehole electrical resistivity data via ensemble Kalman filtering. J. Hydrol. 567, 149–164. https://doi.org/10.1016/j.jhydrol.2018.10.019.

Kang, X., Kokkinaki, A., Kitanidis, P.K., Shi, X., Revil, A., Lee, J., Soueid Ahmed, A., Wu, J., 2020. Improved characterization of DNAPL source zones via sequential hydrogeophysical inversion of hydraulic-head, self-potential and partitioning tracer data. Water Resour. Res. 56 (8), e2020WR027627. https://doi.org/10.1029/2020WR027627.

Kang, X., Kokkinaki, A., Shi, X., Yoon, H., Lee, J., Kitanidis, P.K., Wu, J., 2022. Integration of deep learning-based inversion and Upscaled mass-transfer model for DNAPL mass-discharge estimation and uncertainty assessment. Water Resour. Res. 58 (10), e2022WR033277. https://doi.org/10.1029/2022WR033277.

Kang, X., Power, C., Kokkinaki, A., Revil, A., Wu, J., Shi, X., Deng, Y., 2023. Characterization of DNAPL source zones in clay-sand media via joint inversion of DC resistivity, induced polarization and borehole data. J. Contam. Hydrol. 258, 104240. https://doi.org/10.1016/j.jconhyd.2023.104240.

Kang, X., Kokkinaki, A., Shi, X., Lee, J., Guo, Z., Ni, L., Wu, J., 2024. Modeling upscaled mass discharge from complex DNAPL source zones using a Bayesian neural network: prediction accuracy, uncertainty quantification and source zone feature importance. Water Resour. Res. 60 (7), e2023WR036864. https://doi.org/10.1029/2023WR036864.

Kemna, A., Binley, A., Cassiani, G., Niederleithinger, E., Revil, A., Slater, L., Williams, K. H., Orozco, A.F., Haegel, F.H., Hördt, A., 2012. An overview of the spectral induced polarization method for near-surface applications. Near Surface Geophys. 10 (6), 453–468. https://doi.org/10.3997/1873-0604.2012027.

Koch, J., Nowak, W., 2015. Predicting DNAPL mass discharge and contaminated site longevity probabilities: conceptual model and high-resolution stochastic simulation. Water Resour. Res. 51 (2), 806–831. https://doi.org/10.1002/2014WR015478.

Koch, J., Nowak, W., 2016. Identification of contaminant source architectures—a statistical inversion that emulates multiphase physics in a computationally practicable manner. Water Resour. Res. 52 (2), 1009–1025. https://doi.org/10.1002/2015WR017894.

Koohbor, B., Deparis, J., Leroy, P., Ataie-Ashtiani, B., Davarzani, H., Colombano, S., 2022. DNAPL flow and complex electrical resistivity evolution in saturated porous media: a coupled numerical simulation. J. Contam. Hydrol. 248, 104003. https://doi.org/10.1016/j.jconhyd.2022.104003.

Lindley, D.V., 1956. On a measure of the information provided by an experiment. Ann. Math. Stat. 27 (4), 986–1005. https://doi.org/10.1214/aoms/1177728069.

Loke, M., Acworth, I., Dahlin, T., 2003. A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. Explor. Geophys. 34 (3), 182–187. https://doi.org/10.1071/EG03182.

Loke, M., Chambers, J., Ogilvy, R., 2006. Inversion of 2D spectral induced polarization imaging data. Geophys. Prospect. 54 (3), 287–301. https://doi.org/10.1111/j.1365-2478.2006.00537.x.

Loke, M., Wilkinson, P., Chambers, J., 2010. Fast computation of optimized electrode arrays for 2D resistivity surveys. Comput. Geosci. 36 (11), 1414–1426. https://doi.org/10.1016/j.cageo.2010.03.016.

Loke, M., Wilkinson, P., Chambers, J., Uhlemann, S., Sorensen, J., 2015. Optimized arrays for 2-D resistivity survey lines with a large number of electrodes. J. Appl. Geophys. 112, 136–146. https://doi.org/10.1016/j.jappgeo.2014.11.011.

Maji, R., Sudicky, E., 2008. Influence of mass transfer characteristics for DNAPL source depletion and contaminant flux in a highly characterized glaciofluvial aquifer. J. Contam. Hydrol. 102 (1–2), 105–119. https://doi.org/10.1016/j.jconhyd.2008.08.005.

Maji, R., Sudicky, E., Panday, S., Teutsch, G., 2006. Transition probability/Markov chain analyses of DNAPL source zones and plumes. Groundwater 44 (6), 853–863. https://doi.org/10.1111/j.1745-6584.2005.00194.x.

Marion, D., Nur, A., Yin, H., Han, D.-H., 1992. Compressional velocity and porosity in sand-clay mixtures. Geophysics 57 (4), 554–563. https://doi.org/10.1190/1.1443269.

Mcgillivray, P.R., Oldenburg, D., 1990. Methods for calculating fréchet derivatives and sensitivities for the non-linear inverse problem: a comparative study. Geophys. Prospect. 38 (5), 499–524. https://doi.org/10.1111/j.1365-2478.1990.tb01859.x.

Moretti, L., 2005. In Situ Bioremediation of DNAPL Source Zones. In: National Network of Environmental Management Studies Fellows for US Environmental Protection Agency.

Mumford, K.G., Bryck, S., Kueper, B.H., Mancini, S., Kavanaugh, M., Reynolds, D., 2022. Virtual site investigation to evaluate conceptual site model development at DNAPL-impacted sites. Groundwater Monitor. & Remediat. 42 (3), 44–58. https://doi.org/10.1111/gwmr.12537.

National Research Council, 2013. Alternatives for Managing the nation's Complex Contaminated Groundwater Sites. National Academies Press.

Oldenburg, D.W., Li, Y., 1994. Inversion of induced polarization data. Geophysics 59 (9), 1327–1341. https://doi.org/10.1190/1.1443692.

Patchett, J.G., 1975. An Investigation of Shale Conductivity. SPWLA Annual Logging Symposium.

Power, C., Gerhard, J.I., Tsourlos, P., Giannopoulos, A., 2013. A new coupled model for simulating the mapping of dense nonaqueous phase liquids using electrical resistivity tomography. Geophysics 78 (4), EN1–EN15. https://doi.org/10.1190/geo2012-0395.1.

Power, C., Gerhard, J.I., Karaoulis, M., Tsourlos, P., Giannopoulos, A., 2014. Evaluating four-dimensional time-lapse electrical resistivity tomography for monitoring DNAPL source zone remediation. J. Contam. Hydrol. 162, 27–46. https://doi.org/10.1016/j.jconhyd.2014.04.004.

Power, C., Gerhard, J.I., Tsourlos, P., Soupios, P., Simyrdanis, K., Karaoulis, M., 2015. Improved time-lapse electrical resistivity tomography monitoring of dense non-aqueous phase liquids with surface-to-horizontal borehole arrays. J. Appl. Geophys. 112, 1–13 https://doi.org/j.jappgeo.2014.10.022.

Qiang, S., Shi, X., Kang, X., Revil, A., 2022. Optimized arrays for electrical resistivity tomography survey using Bayesian experimental design. Geophysics 87 (4), E189–E203 https://doi.org/geo2021-0408.1.

Qiang, S., Shi, X., Revil, A., Kang, X., Song, Y., Xing, K., 2024. Quantitative evaluation of the effect of pore fluids distribution on complex conductivity saturation exponents. J. Geophys. Res. Solid Earth 129 (8), e2024JB028689. https://doi.org/10.1029/2024JB028689.

Rabaute, A., Revil, A., Brosse, E., 2003. In situ mineralogy and permeability logs from downhole measurements: application to a case study in chlorite-coated sandstones. J. Geophys. Res. Solid Earth 108 (B9). https://doi.org/10.1029/2002JB002178.

Revil, A., 2012. Spectral induced polarization of shaly sands: influence of the electrical double layer. Water Resour. Res. 48 (2). https://doi.org/10.1029/2011WR011260.

Revil, A., 2013. Effective conductivity and permittivity of unsaturated porous materials in the frequency range 1 mHz–1GHz. Water Resour. Res. 49 (1), 306–327. https://doi.org/10.1029/2012WR012700.

Revil, A., Cathles, L.M., 1999. Permeability of shaly sands. Water Resour. Res. 35 (3), 651–662. https://doi.org/10.1029/98WR02700.

Revil, A., Cathles Iii, L., Losh, S., Nunn, J., 1998. Electrical conductivity in shaly sands with geophysical applications. J. Geophys. Res. Solid Earth 103 (B10), 23925–23936. https://doi.org/10.1029/98JB02125.

Revil, A., Grauls, D., Brévart, O., 2002. Mechanical compaction of sand/clay mixtures. J. Geophys. Res. Solid Earth 107 (B11). https://doi.org/10.1029/2001JB000318.

Revil, A., Karaoulis, M., Johnson, T., Kemna, A., 2012a. Some low-frequency electrical methods for subsurface characterization and monitoring in hydrogeology. Hydrogeol. J. 20 (4), 617. https://doi.org/10.1007/s10040-011-0819-x.

Revil, A., Koch, K., Holliger, K., 2012b. Is it the grain size or the characteristic pore size that controls the induced polarization relaxation time of clean sands and sandstones? Water Resour. Res. 48 (5). https://doi.org/10.1029/2011WR011561.

Revil, A., Kessouri, P., Torres-Verdín, C., 2014. Electrical conductivity, induced polarization, and permeability of the Fontainebleau sandstone. Geophysics 79 (5), D301–D318. https://doi.org/10.1190/GEO2014-0036.1.

Revil, A., Binley, A., Mejus, L., Kessouri, P., 2015. Predicting permeability from the characteristic relaxation time and intrinsic formation factor of complex conductivity spectra. Water Resour. Res. 51 (8), 6672–6700. https://doi.org/10.1002/2015WR017074.

Revil, A., Coperey, A., Shao, Z., Florsch, N., Fabricius, I.L., Deng, Y., Delsman, J., Pauw, P., Karaoulis, M., De Louw, P., 2017. Complex conductivity of soils. Water Resour. Res. 53 (8), 7121–7147. https://doi.org/10.1002/2017WR020655.

Revil, A., Tartrat, T., Abdulsamad, F., Ghorbani, A., Coperey, A., 2018. Chargeability of porous rocks with or without metallic particles. Petrophysics 59 (04), 544–553. https://doi.org/10.30632/PJV59V4-2018a8.

Revil, A., Schmutz, M., Abdulsamad, F., Balde, A., Beck, C., Ghorbani, A., Hubbard, S., 2021. Field-scale estimation of soil properties from spectral induced polarization tomography. Geoderma 403, 115380. https://doi.org/10.1016/j.geoderma.2021.115380.

Revil, A., Ghorbani, A., Jougnot, D., Yven, B., 2023. Induced polarization of clay-rich materials—part 1: the effect of desiccation. Geophysics 88 (4). https://doi.org/10.1190/GEO2022-0510.1. MR195-MR210.

Sauck, W.A., 1998. High conductivities associated with an LNAPL plume imaged by integrated geophysical techniques. J. Environ. Eng. Geophys. 2, 203–212.

Seigel, H.O., 1959. Mathematical formulation and type curves for induced polarization. Geophysics 24 (3), 547–565. https://doi.org/10.1190/1.1438625.

Slater, L., Binley, A., 2021. Advancing hydrological process understanding from long-term resistivity monitoring systems. Wiley Interdiscip. Rev. Water 8 (3), e1513. https://doi.org/10.1002/wat2.1513.

Song, Y., Shi, X., Revil, A., Deng, Y., 2024. Influence of dissolved and non-aqueous phase toluene on spectral induced polarization signatures of soils. J. Hazard. Mater. 135973. https://doi.org/10.1016/j.jhazmat.2024.135973.

Soueid, Ahmed A., Revil, A., 2018. 3-D time-domain induced polarization tomography: a new approach based on a source current density formulation. Geophys. J. Int. 213 (1), 244–260. https://doi.org/10.1093/gji/ggx547.

Sparrenbom, C.J., Åkesson, S., Johansson, S., Hagerberg, D., Dahlin, T., 2017. Investigation of chlorinated solvent pollution with resistivity and induced polarization. Sci. Total Environ. 575, 767–778. https://doi.org/10.1016/j.scitotenv.2016.09.117.

Wang, F., Annable, M.D., Schaefer, C.E., Ault, T.D., Cho, J., Jawitz, J.W., 2014. Enhanced aqueous dissolution of a DNAPL source to characterize the source strength function. J. Contam. Hydrol. 169, 75–89. https://doi.org/10.1016/j.jconhyd.2014.07.004.

Wang, H., Huisman, J.A., Zimmermann, E., Vereecken, H., 2021. Experimental design to reduce inductive coupling in spectral electrical impedance tomography (sEIT) measurements. Geophys. J. Int. 225 (1), 222–235. https://doi.org/10.1093/gji/ggaa594.

Wang, Z., Yang, P., Guo, H., Wang, Y., 2023. Enhanced mass transfer of residual NAPL by convection in stagnant zone. J. Hydrol. 625, 130050. https://doi.org/10.1016/j.jhydrol.2023.130050.

Wilkinson, P.B., Meldrum, P.I., Chambers, J.E., Kuras, O., Ogilvy, R.D., 2006. Improved strategies for the automatic selection of optimized sets of electrical resistivity tomography measurement configurations. Geophys. J. Int. 167 (3), 1119–1126. https://doi.org/10.1111/j.1365-246X.2006.03196.x.

Wilkinson, P.B., Loke, M.H., Meldrum, P.I., Chambers, J.E., Kuras, O., Gunn, D.A., Ogilvy, R.D., 2012. Practical aspects of applied optimized survey design for electrical resistivity tomography. Geophys. J. Int. 189 (1), 428–440. https://doi.org/10.1111/j.1365-246X.2012.05372.x.

Wilkinson, P.B., Uhlemann, S., Meldrum, P.I., Chambers, J.E., Carrière, S., Oxby, L.S., Loke, M., 2015. Adaptive time-lapse optimized survey design for electrical resistivity tomography monitoring. Geophys. J. Int. 203 (1), 755–766. https://doi.org/10.1093/gji/ggv329.

Xie, Q., Mumford, K.G., Kueper, B.H., 2023. Simulating field-scale thermal conductive heating with the potential for the migration and condensation of vapors. J. Hazard. Mater. 453, 131439. https://doi.org/10.1016/j.jhazmat.2023.131439.

Xu, B., Noel, M., 1993. On the completeness of data sets with multielectrode systems for electrical resistivity survey. Geophys. Prospect. 41 (6), 791–801.

Zarif, F., Kessouri, P., Slater, L., 2017. Recommendations for field-scale induced polarization (IP) data acquisition and interpretation. J. Environ. Eng. Geophys. 22 (4), 395–410. https://doi.org/10.2113/JEEG22.4.395.