

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**CHARACTERIZING AND DETECTING URBAN
UNDERGROUND HAZARDS BY TEMPORAL
GROUND PENETRATING RADAR MEASUREMENTS**

YIMIN ZHOU

PhD

The Hong Kong Polytechnic University

2026

The Hong Kong Polytechnic University
Department of Land Surveying and Geo-informatics

**Characterizing and detecting urban underground hazards by
temporal ground penetrating radar measurements**

Yimin Zhou

**A thesis submitted in partial fulfillment of the requirement for the
degree of Doctor of Philosophy**

April 2025

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

(Signed)

Yimin Zhou

(Name of student)

Abstract

The increasing urban population has intensified demands for buried water supplies, drainage systems, energy distribution, and telecommunication networks. Repetitive road openings for new installations and replacement of aging utilities have made underground infrastructure more complex and congested. Aging utility infrastructure often leads to issues like ground subsidence caused by voids and water main leakages. Effective imaging and diagnostics are critical for decision-making on rehabilitation or replacement strategies. Ground penetrating radar (GPR) has emerged as the most suitable non-destructive technology for underground imaging. However, traditional GPR surveys using single- or dual-channel systems are labor-intensive, time-consuming, and limited to small-scale applications, making them impractical for large-scale urban hazard detections of void and leakage. Multichannel GPR (MCGPR), as an evolving technology in recent years, enhances efficiency by enabling large-scale and temporal data acquisition because vehicle-towed antennae can be arranged side-by-side together. But is the GPR processing simply an exercise of channel adding up? Can the conventional GPR survey scheme be simply adopted? The answers to the above questions are all negative, and migration of the data processing approach from GPR to MCGPR is required but underexplored in the GPR community.

This thesis bridges critical gaps in MCGPR-based hazard detection through three interconnected contributions:

(1). Leakage Signature **Database** & Frequency-Domain **Characterization**

A comprehensive database of 78 real-world leakage scenarios was created using a 2,400 sq.m well-designed experimental site ‘Q-Leak’. This database reduces reliance on simulated data by capturing leakage patterns across different pavement types and pipe attributes. Each leakage scenario was designed and measured in a before-and-after

emulation scheme, enabling the database to capture both static leakage patterns and dynamic leakage patterns through temporal GPR measurements. To facilitate dynamic pattern recognition, three sub-datasets were created. This database provides a solid foundation for advancing the recognition of temporal leakage patterns.

Additionally, GPR-Instantaneous frequency slices (IFS) in the time-frequency domain were developed as a new method for characterizing water leakage patterns. The developed method addresses the limitations of traditional amplitude slice analysis, and the restrictions of signal aliasing and windowing in short-time Fourier transform (STFT) and wavelet transform methods.

(2). MCGPR **Error Mitigation** Framework

Two error mitigation approaches were developed to reduce geo-location misalignment of MCGPR trajectory (above-ground) and harmonize signal variations in MCGPR antennae channels (underground). The first method focuses on correcting geolocation misalignments, which are a common yet often neglected issue in MCGPR trajectory data. A novel post-processing method that, for the first time, leverages cross-channel MCGPR measurements was proposed to enhance positioning accuracy. This approach achieves centimeter-level accuracy and is compatible with various positioning systems, making it a versatile solution for improving geolocation reliability. For signal harmonization, A two-stage correction method was implemented to improve signal consistency, ensuring that data from multiple surveys and channels are comparable. These error mitigation techniques enhance the overall quality of MCGPR data, enabling more accurate analysis and hazard detection.

(3). Change-detection **Algorithms**

There are two foci of algorithm development for identifying underground hazards, i.e., voids and water leakages. For void detection, an unsupervised superpixel-based method

combines fuzzy C-means (FCM) clustering with a Markov Random Field (MRF) model. This approach integrates 3D neighborhood information to iteratively refine change maps. For leakage detection, a hierarchical methodology leverages both amplitude and instantaneous frequency analyses. Potential leakage regions are identified using 3D FCM clustering, followed by IFS to extract frequency attributes associated with water leakage, further narrowing regions of interest due to suspected leakage.

The individual research contributions in the above three perspectives were integrated to build a holistic forensic investigation in a real-world case study of leakage. The study began with an initial MCGPR survey over a 5.5 km road segment of known leakage, but it had no idea where it could be. By studying the anomaly of water pressure records, a 1.6 km subsection was identified as a preliminary area of suspected leakage, where three subsequent temporal MCGPR surveys were conducted. The developed error mitigation methods and change detection algorithms were applied to data collected during two distinct periods (Period 1 and Period 2), divided by the three temporal surveys. The suspected leakage regions identified during Period 2 aligned closely with the actual pipe burst location on 3rd November 2024, validating the effectiveness of the developed methods.

MCGPR, as the most recent development in the GPR community, is not merely an antenna array stacked and towed behind a vehicle. In reality, its adoption and change-detection diagnosis of underground requires a series of fundamental understanding of antennae and signals, and algorithm research and development. This research advances MCGPR-based underground hazard characterization by developing and validating a series of scientific approaches. The contributions of this work are believed to benefit engineering, surveying, and infrastructure management authorities by offering innovative tools that pave the way for the non-invasive and data-driven detection, characterization, and prediction of underground hazards.

List of Publications

Journal Papers:

- Zhou, Y., Lai, W. W. L., & Zhu, X. (2023). Superpixel-based change detection for GPR time-lapse slices using fuzzy c-means and the Markov random field method. *Tunnelling and Underground Space Technology*, 141, 105369. <https://doi.org/10.1016/j.tust.2023.105369>
- Zhou, Y., & Lai, W. W. L. (2024). Characterization of leakage signatures in buried water pipes by ground penetrating radar (GPR) and instantaneous frequency analysis. *Tunnelling and Underground Space Technology*, 153, 105984. <https://doi.org/10.1016/j.tust.2024.105984>
- Zhou, Y., Wong, P. T. W., Li, Y., Lai, W. W. L., & Wang, J. (2026). Enhancing multichannel ground penetrating radar (MCGPR) positioning using cross-channel data. *Measurement*, 257, 118715. <https://doi.org/10.1016/j.measurement.2025.118715>

Conference Paper:

- Zhou, Y., & Lai, W. W. L. (2023). Leakage Detection Using Ground Penetrating Radar C-Scan Based on 3D Fuzzy C-Means Clustering. 2023 12th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), 1–4. <https://doi.org/10.1109/IWAGPR57138.2023.10329148>
- Zhou, Y., & Lai, W. W. L. (2025). Projecting and Positioning Unknown Trajectories of MCGPR Swaths. 2025 13th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), 1–5. <https://doi.org/10.1109/IWAGPR65621.2025.11109040>

Acknowledgment

I would like to express my deepest gratitude to my chief supervisor, Prof. Wallace Lai, for his unwavering support and guidance throughout my PhD journey. When I began this journey, I was completely new to the field of GPR. His vast professional knowledge and insightful instructions not only introduced me to this field but also inspired me to grow within it. His encouragement consistently motivated me to take steps forward, one at a time, and his friendly, approachable nature made this process all the more enjoyable. Prof. Lai is an exceptional mentor, one who truly breaks the mold of a typical professor. I never felt any distance working with him; instead, I felt a sense of camaraderie and partnership. From conducting experiments and performing analyses to discussing research papers, his guidance was always thought-provoking and allowed me the freedom to explore potential directions. Beyond academics, he also emphasized the importance of physical and mental well-being. During the challenging period of the pandemic, his support helped me manage my anxiety and stay grounded. Some of my most cherished memories include the times we spent conducting field surveys, brainstorming ideas, sharing drinks, and even running marathons together. These moments are more like partners and friends' relationships, and I will carry them with me for a lifetime. I cannot imagine a better supervisor to have guided me through this journey, and I am truly grateful for everything he has done for me.

I would also like to express my heartfelt gratitude to my co-supervisor, Dr. Xiaolin Zhu. His guidance and encouragement have been instrumental in shaping my scientific research journey. At a critical crossroads between work and further study, he inspired me to pursue research and provided me with the motivation to move forward. Dr. Zhu's expertise in the field of remote sensing has been a constant source of valuable ideas, offering fresh perspectives that have enriched my understanding of GPR. The influence of his knowledge has been deeply

embedded in my research, and the "genes" of remote sensing that I inherited from him have become an essential component of my work. I am truly grateful for his professional insights, thoughtful guidance, and invaluable contributions, all of which have greatly impacted the direction and quality of my research.

I would also like to extend my heartfelt thanks to my groupmates: Janet, Ray, Bella, Lydia, Samuel, Sahib, Wenchao, Alex, Tom, Rene, Ho, and Marco. Thank you for your support during fieldwork and for your valuable comments. It has been such a wonderful experience working with all of you. I would also like to express my gratitude to Prof. Antonis, Dr. Sonia, Lisandro, and Jan for our insightful conversations. Your words and ideas have truly inspired me and broadened my perspective. My sincere thanks go to Kin, Dexter, Jack, Jacky, Mr. Ho, and Mr. Shing for their assistance throughout my studies. I am also deeply grateful to Yaxin, Zhifei, Xiaonan, Liujian, Lifang, and Huamei for their kind support and help during this journey. None of this would have been possible without your generosity and encouragement. I would like to thank Prof. George Liu, chair of the thesis examination board, as well as Dr. Craig Warren and Prof. Anja Klotzsche, the external examiners of this thesis, for their constructive comments and suggestions.

I would also like to express my deepest gratitude to my parents. Their unconditional love has been the source of my strength throughout this entire journey. To my friends, I am truly grateful for the relationships we have built, which have greatly enriched my life. Finally, I am grateful to my girlfriend, Yanbi Chen, for her understanding, patience, and unwavering support—your companionship has meant the world to me, and I cannot fully express how much I appreciate all you have done. Thank you all for the kindness, encouragement, and positivity you have sent my way—I have received them. I have felt and treasured them. These “wiggles” will not attenuate with time.

Table of Content

Abstract.....	ii
List of Publications.....	v
Acknowledgement.....	vi
Table of Content.....	viii
List of Figures.....	xi
List of Tables	xix
Table of Abbreviations.....	xx
Chapter 1. Introduction.....	1
1.1 Background-----	1
1.2 Research Objectives and Roadmap -----	4
Chapter 2. Literature Reviews.....	11
2.1 Mechanism of Subsurface Hazards -----	11
2.1.1 Factors Contributing to Pipeline Degradation and Leakage.....	13
2.1.2 Causes and Mechanisms of Subsurface Voids Formation	16
2.2 Theory of GPR -----	18
2.2.1 GPR Antenna and Operations.....	19
2.2.2 EM Wave Properties.....	21
2.2.3 Interactions between GPR and Materials	24
2.3 GPR Processing and Imaging -----	29
2.3.1 GPR Data Representations	30
2.3.2 Current Standard GPR Post-processing.....	31
2.3.3 GPR Imaging.....	37
2.4 Pattern Recognition and Change Detection-----	40
2.4.1 One-shot Measurement and Pattern Recognition	40
2.4.2 Temporal Measurements and Change Detection	45
2.5 Summary of the Current Researches and Research Gaps-----	50
Chapter 3. Research Resources	52

3.1 GPR Equipment-----	52
3.2 Research Experiment Sites-----	54
3.3 Software-----	57
Chapter 4. Leakage Signature Database and Leakage Characterization.....	59
4.1 Leakage Signature Database -----	60
4.1.1 Experiment Sites and Designed Leakage Experiments	60
4.1.2 Dataset One: Temporal B-scans of Buried Pipes Leakage	68
4.1.3 Dataset Two: Temporal C-scans of Buried Pipes Leakage.....	74
4.1.4 Dataset Three: Temporal C-scans of Buried Pipes with Various Leakage Volume.....	79
4.2 Characterization of Leakage Signatures using Instantaneous Frequency -----	83
4.2.1 Workflow of IFS Generation based on GPR Data.....	85
4.2.2 Experiment Results and Discussion	93
4.3 Chapter Summary and Contribution -----	109
Chapter 5. Underground Hazards Detection using Temporal GPR Data.....	111
5.1 Underground Void Detection based on GPR Time-lapse Slices Integrating Fuzzy C-Means and Markov Random Field Method -	112
5.1.1 Underground Void Detection Methodology	113
5.1.2 Experiment Results and Discussion	123
5.2 Enhancing Water Leakage Detection in Urban Infrastructure by Integrating AS and IFS -----	136
5.2.1 Leakage Detection Method	137
5.2.2 Experiment Results and Discussion	140
5.3 Chapter Summary and Contribution -----	149
Chapter 6. Time-series MCGPR Registration: From geolocation to Signal Domain.....	151
6.1 Enhancing MCGPR Positioning Using Cross-Channel Data -----	152
6.1.1 Position Solutions of MCGPR Survey	152
6.1.2 MCGPR Positioning Enhancement Method.....	155

6.1.3 Experiment Results and Discussion	169
6.2 Harmonizing MCGPR Signals	187
6.2.1 MCGPR Harmonizing Method	188
6.2.2 Experiments and Results	191
6.3 Chapter Summary and Contribution	199
Chapter 7. A Forensic Investigation in a Real-world Leakage Case.....	201
7.1 Site and MCGPR Temporal Measurement's Introduction	203
7.2 Position Enhancement and MCGPR Signal Harmonization	207
7.3 Change Detection Results	211
7.3.1 Amplitude Slices.....	211
7.3.2 Instantaneous Frequency Slices.....	214
7.3.3 Change Map Combining AS and IFS	216
7.4 Discussion and Conclusion.....	222
7.5 Chapter Summary and Contribution	229
Chapter 8. Conclusion	231
8.1 Summary of Major Findings and Contributions to Knowledge.....	231
8.2 Challenges and Recommendation on Future Work	234
Reference.....	237

List of Figures

Figure 1-1 Thesis roadmap: from data acquisition to hazard detection.....	5
Figure 1-2 Illustration of data acquisition in this thesis.....	6
Figure 1-3 Illustration of Objective 1	7
Figure 1-4 Illustration of Objective 2	8
Figure 2-1 Fault tree diagram of the corrosion causes for water pipelines. A comprehensive summary of the corrosion causes in water pipes based on bibliometric and systematic analysis (Hussein et al., 2023)	13
Figure 2-2 Hyperbolic reflection from common offset GPR, t_1 and t_2 are the GPR wave travel time at antenna position 1 and 2, respectively; while x is the moving distance of the antenna (Lai et al., 2016a).....	19
Figure 2-3 Simplified diagram of the impulse radar technique (Jol, 2009a)	20
Figure 2-4 Permittivity relaxation phenomena associated with an idealized, lossy dielectric material. Region A represents the polarisations associated with free-charge and material interactions whilst region B represents molecular and atomic polarisations associated with bound charges (Adapted from (Daniels, 1996))	23
Figure 2-5 Variation in velocity and attenuation in a simple medium with nondispersive physical properties, c and Z_0 are the velocity and impedance of free space (i.e., a vacuum) (Jol, 2009a).....	25
Figure 2-6 The illustration of scattering with object sizes a (Annan, 2004), σ is the ratio of GPR wavelength to the size of the scatterer, and λ denotes the wavelength	26
Figure 2-7 Illustration of the different GPR data representations (A-scan, B-scan / radargram, C-scan / slice).....	30
Figure 2-8 Typical processing sequence of 2D GPR data (Jol, 2009a)	32
Figure 2-9 An example of the de-wow on a GPR trace (Rowell et al., 2010): (a) before and (b) after the de-wow	33
Figure 2-10 3D GPR imaging workflow. PS: profile spacing; ST: slice thickness; SR: search radius; λ : wavelength of the EM wave propagation underground (Luo et al., 2019a)	39
Figure 2-11 C-scan intensity normalization scheme that continuous the 3D GPR imaging workflow (Luo et al., 2023)	39
Figure 2-12 Illustration of the development of tumors in different stages (Joseph, 2023).....	46

Figure 2-13 Illustrations of road subsidence development from delamination to subsidence	47
Figure 2-14 Static pattern recognition vs. dynamic pattern recognition for underground hazard detection.....	48
Figure 3-1 Photos of GPR systems (a) IDS-Hi mode, (b) DS2000, (c) StreamUP, and (d) Raptor45.....	53
Figure 3-2 Illustration of the channel configurations of StreamUp and Raptor45	54
Figure 3-3 Photos of the experiment site at Utilities Lab at PolyU.....	55
Figure 3-4 Photos of the Q-Leak center and leak detection ground	57
Figure 4-1 Overview of the leakage detection ground showing five different pavement types and corresponding photos	61
Figure 4-2 Locations of five leakage experiment cells and detailed views of the leakage pipes	62
Figure 4-3 GPR B-scans across buried pipes under different pavement types using a 600 MHz IDS Hi-Mode antenna (dry conditions)	66
Figure 4-4 (a) Top-view photo of Cell 5 pipe configurations; (b) GPR B-scan from the left survey line; (c) GPR B-scan from the right survey line.....	67
Figure 4-5 Experimental timeline for Dataset One and survey line layout overlaid on the cell schematic	69
Figure 4-6 Experimental timeline for Dataset Two and survey grid overlaid with the cell schematic	75
Figure 4-7 Experimental timeline for Dataset Three and survey grid overlaid with the cell schematic	80
Figure 4-8 The workflow of the AS and IFS generation	85
Figure 4-9 Schematic diagram of the process for generating instantaneous frequency B-scan via VMD from processed radargrams	89
Figure 4-10 Illustration of dynamically adjusting slice thickness based on hyperbolic-based velocity estimation.....	90
Figure 4-11 Illustration of the leakage region, background region, and manhole region at the experiment site.....	91
Figure 4-12 Illustration of center IF drift and IOU.....	93
Figure 4-13 Site photos of the designed leakage cell	94
Figure 4-14 Sketch of the leakage pipes with marked leak points and vertical cross-section	

views of the experimental site.....	94
Figure 4-15 Experimental timeline for leakage scenarios and survey grid overlaid with the cell schematic (same as Figure 4-6).....	95
Figure 4-16 (a) IFS of baseline ; (b) Scenario 1 at depth 0.9 – 1.05 m; (c) Histograms of background and target regions	96
Figure 4-17 Depiction of vertical water leakage patterns in Scenario 1, as illustrated by variation in centre IF drift, IOU, and Pearson correlation coefficient across different IFS	98
Figure 4-18 Depiction of vertical water leakage patterns in Scenario 2, as illustrated by variation in centre IF drift, IOU, and Pearson correlation coefficient across different IFS	99
Figure 4-19 Depiction of vertical water leakage patterns in Scenario 3, as illustrated by variation in centre IF drift, IOU, and Pearson correlation coefficient across different IFS	99
Figure 4-20 Depiction of vertical water leakage patterns in Scenario 1, as illustrated by variation in centre amplitude drift, IOU, and Pearson correlation coefficient across different AS	102
Figure 4-21 Depiction of vertical water leakage patterns in Scenario 2, as illustrated by variation in centre amplitude drift, IOU, and Pearson correlation coefficient across different AS	103
Figure 4-22 Depiction of vertical water leakage patterns in Scenario 3, as illustrated by variation in centre amplitude drift, IOU, and Pearson correlation coefficient across different AS	103
Figure 4-23 Difference image of IFS at a depth range from 1.215m to 1.365m in Scenario1, highlighting areas significantly affected by water leakage (marked by yellow dashed line)	105
Figure 4-24 Average IF and amplitude curve of 700MHz antenna for areas significantly affected by water leakage in Scenario 1	106
Figure 4-25 Average IF and amplitude curve of 600MHz antenna for areas significantly affected by water leakage in Scenario 1	106
Figure 4-26 Average IF and amplitude curve of 250MHz antenna for areas significantly affected by water leakage in Scenario 1	107
Figure 5-1 Workflow of the proposed superpixel-based change detection method.....	114
Figure 5-2 3D GPR imaging workflow. PS: profile spacing; ST: slice thickness; SR: search	

radius; λ : wavelength of the EM wave propagation underground (The underlined marks in the figure denote the parameter choices used in our method) (Luo et al., 2019b).....	116
Figure 5-3 Illustrations of pixel-based DI map and multiscale DI maps (a) pixel-based DI map; superpixel-based DI maps: (b) level-1; (c) level-2; (d) level-3; (e) level-4.....	118
Figure 5-4 Soft label clustering illustration: (a) hard classification: 0 indicates unchanged and 1 indicates changed; (b) soft classification: the values refer to the probability of change	119
Figure 5-5 Superpixel-adjacent neighbourhoods in three dimensions.....	121
Figure 5-6 Photos of the experiment site and targets. (a) the pipe and manhole distribution at the test site; (b) the test site and the IDS hi-mod antenna; (c) surveying grid of the test site; (d) a buried spherical fitness ball; (e) a buried peanut-shaped ball; (f) a void between pavement and soil	124
Figure 5-7 Example of the CLAHE image enhancement function for small intensity differences in DI. (a) and (d) are raw DI, (b) and (e) are enhanced DI, and (c) and (f) are histograms of C-scans. Red rectangles indicate the relatively small intensity change in raw and enhanced DI	126
Figure 5-8 Illustration of superpixel segmentation module for local noise resistance. (a) and (b) C-scans with and without void, (c) pixel-based DI, (d) superpixel-based DI....	128
Figure 5-9 Illustration of MRF module for noise cluster resistance. (a) and (b) C-scans with and without void, (c) superpixel-based DI, (d) adjacent superpixel-based DI of (c)	128
Figure 5-10 Influence of parameter β associated with four scenarios	129
Figure 5-11 Pairs of C-scans (no void case and void case) and change maps created using the proposed method with corresponding references in the four scenarios (White regions are detected as changed areas and red masks refer to ground truth in change maps).....	131
Figure 5-12 Change maps generated by different methods in Scenario A; (a) FCM +MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM +MRF-1 (superpixel-based). (d) FCM +MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based).....	133
Figure 5-13 Change maps generated by different methods in Scenario B; (a) FCM +MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM	

+MRF-1 (superpixel-based); (d) FCM +MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based).....	133
Figure 5-14 Change maps generated by different methods in Scenario C; (a) FCM +MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM +MRF-1 (superpixel-based); (d) FCM +MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based).....	134
Figure 5-15 Change maps generated by different methods in Scenario D; (a) FCM +MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM +MRF-1 (superpixel-based); (d) FCM +MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based).....	134
Figure 5-16 The whole workflow of the leakage detection method integrating amplitude and instantaneous frequency slices.....	137
Figure 5-17 (a) Illustration of leakage experiment design; (b) Illustration of background region, leakage region, and manhole region; (c) Vertical cross-section (align with the red line in (b)) views of the experimental site.....	142
Figure 5-18 AS without leakage, AS with leakage, and superpixel-based difference images generated from AS with and without leakage. All slices correspond to depth ranges from 0.675 m to 1.875 m at 0.15 m intervals. Remark: Arrows ‘a’ mark the orthogonal patterns that refer to artefacts during interpolation.....	145
Figure 5-19 IFS without leakage, IFS with leakage, and superpixel-based difference images generated from IFS with and without leakage. All slices correspond to depth ranges from 0.675 m to 1.875 m at 0.15 m intervals.....	147
Figure 5-20 Results of four leakage scenarios.....	148
Figure 6-1 illustration of GNSS positioning performance in the urban region; (a) GNSS recorded trajectories; (b) referenced trajectories by mobile mapping system ...	153
Figure 6-2 The whole workflow of the proposed positioning enhancement method	156
Figure 6-3 Workflow for trajectory reconstruction using cross-channel MCGPR image pairing (extended from Figure 6-2)	159
Figure 6-4 Illustration of underground feature extraction by HOG and 2DFFT	161
Figure 6-5 Illustration of underground feature pairing function.....	163
Figure 6-6 Illustration of the DP optimization processing.....	165

Figure 6-7 Illustration of MCGPR road survey and the components of MCGPR system: MCGPR antennas and position system.....	170
Figure 6-8 (a) Overview of the three validation sites used in the study; (b) Illustration of the overlapping region between Swath 01 and Swath 02, highlighting their spatial coverage; (c) Trajectories of the three sites, showing the reference paths and the masked sections for validation.....	170
Figure 6-9 The error distributions of three test sites with corresponding MAE and RMSE	174
Figure 6-10 RMSE of three testing sites with various frame sizes.....	176
Figure 6-11 Errors distribution of testing Site2 with sketch of trajectories of paired swathes	178
Figure 6-12 Illustration of the influence of turning angles.....	179
Figure 6-13 Impact of swath overlapping on trajectory reconstruction accuracy (Remark: Arrow ‘a’ marks the erroneous computations over the blue dotted lines in the last two cases).....	182
Figure 6-14 Impact of reliable positioning measurements on trajectory reconstruction accuracy	185
Figure 6-15 Example of typical strip-like noise comes from the variations of the MCGPR system in the shallow slice.....	188
Figure 6-16 Photos of the raptor MCGPR system and the antenna arrays (a) and (b); (c) The illustration of the transmitters, receivers, and channel location of the raptor MCGPR system	189
Figure 6-17 The whole workflow of the signal harmonization method	190
Figure 6-18 Photos of the experiments (a) Illustration of the time mode; (b) and distance mode; (c) data collection.....	193
Figure 6-19 Peak to peaks amplitude and time-zero variations across channels.....	194
Figure 6-20 Wavelet similarity matrices across channels for GPR data, with different rows representing materials (brick, concrete, and asphalt) and different columns representing correction stages (raw, after Stage 1 correction, and after two-stage correction).....	196
Figure 6-21 The signal harmonization performance of different stages in different pavement materials (brick, concrete, and asphalt)	197
Figure 6-22 Comparison of amplitude slices generated with; (a) and without; (b) the proposed harmonization workflow.....	198

Figure 7-1 Workflow of water leakage detection	202
Figure 7-2 Illustration of the research regions and the temporal MCGPR surveys have been conducted.....	204
Figure 7-3 (a) Utility record plant overlaid on a satellite image; (b) Amplitude slice at 1.1-meter depth showing the gas pipe; (c) Amplitude slice at 1.9-meter depth showing the water main; (d) Radargram corresponding to the yellow rectangle in (b); (e) Radargram corresponding to the orange rectangle in (c)	206
Figure 7-4 News and on-site photo of the pipe bursts (Emily Hung et al., 2024)	207
Figure 7-5 On-site photo of the target region	208
Figure 7-6 (a) GNSS trajectories of the target region; Amplitude slices (b) before; and (c) after positioning enhancement	209
Figure 7-7 Amplitude slice before and after signal harmonization (same as Figure 6-22)..	210
Figure 7-8 Temporal analysis of normalized amplitude slices (AS) and difference images: (a) Normalized amplitude slice (November 6, 2023); (b) Normalized amplitude slice (January 18, 2024); (c) Normalized amplitude slice (June 11, 2024); (d) Amplitude difference image (Period 1); (e) Amplitude difference image (Period 2); (f) Histograms of amplitude differences for Period 1 and Period 2	213
Figure 7-9 Temporal analysis of IFS and difference images: (a) IFS (November 6, 2023); (b) IFS (January 18, 2024); (c) IFS (June 11, 2024); (d) IF difference image (Period 1); (e) IF difference image (Period 2); (f) Histograms of IF differences for Period 1 and Period 2	215
Figure 7-10 Illustration of change clusters identified for Period 1 (a) Change clusters in amplitude difference image slices; (b) Change clusters in IF difference image slices; (c) The intersection of change clusters in amplitude and IF difference images.	218
Figure 7-11 (a) Histogram of the amplitude difference volume within cluster#4; (b) Histogram of the amplitude difference volume within cluster#5; (c) Histogram of the IF difference volume within cluster#4; (d) Histogram of the IF difference volume within cluster#5.....	219
Figure 7-12 Illustration of change clusters identified for Period 2 (a) Change clusters in amplitude difference image slices; (b) Change clusters in IF difference image slices; (c) The intersection of change clusters in amplitude and IF difference images.	221
Figure 7-13 (a) Histogram of the amplitude difference volume within cluster#1; (b) Histogram of the IF difference volume within cluster#1	221
Figure 7-14 The location of confirmed leak points overlayed with the suspected leakage	

clusters	222
Figure 7-15 (a) Example IF slice without background removal filter. (b) Example of IF slice with background removal filter.....	224
Figure 7-16 Wavelet similarity matrix of each channel pairs	225
Figure 7-17 (a) IF values of different channels target the same objects without a background removal filter; (b) IF values of different channels target the same objects with a background removal filter.....	226
Figure 7-18 Average of the MCGPR over-estimated IF values derived from background removal filter.....	227

List of Tables

Table 2-1 Typical values of relative permittivity (real component) and static conductivity for common subsurface materials at an antenna frequency of 100 MHz	22
Table 3-1 Summary of used GPR antennas	52
Table 4-1 Summary of the leakage points and their attributes.....	62
Table 4-2 Description of defect types	65
Table 4-3 Representative data samples from Dataset One (Raptor45)	70
Table 4-4 Representative data samples from Dataset Two (IDS Hi-Mode 600MHz).....	76
Table 4-5 Configurations of GPR survey and parameters of C-scan generation.	77
Table 4-6 Representative data sample from Dataset Three (Raptor45)	81
Table 4-7 Configurations of GPR survey and parameters of C-scan generation.	85
Table 4-8 Detailed description of simulated leakage scenarios	95
Table 5-1 Distribution and sizes of buried targets of the four scenarios.....	125
Table 5-2 TP, TN, FP, FN, F1-score, and Kappa Coefficient in the four scenarios	131
Table 5-3 Comparison of the change detection results obtained by different methods	134
Table 5-4 Detailed description of emulated leakage scenarios	147
Table 6-1 MAE and RMSE of three test sites.....	173
Table 6-2 Results comparisons between different overlapping rates.....	182
Table 6-3 Results comparisons between different reliable position measurements.....	186

Table of Abbreviations

Abbreviation	Full Form
2D FFT	2D fast Fourier transform
AGC	Automatic gain control
CLAHE	Contrast limited adaptive histogram equalization
DC-shift	Direct current shift
DI	Difference image
DMA	District metering area
EM	Electromagnetic
FCM	Fuzzy C-means clustering
FFT	Fast Fourier transforms
GPR	Ground penetrating radar
HOG	Histogram of oriented gradients
IFS	Instantaneous frequency slice
IMF	Intrinsic mode functions
IOU	Intersection over unit
JTFA	Joint time-frequency analysis
LSGI	Land survey and geo-informatic department of PolyU
MCGPR	Multichannel ground penetrating radar
MMS	Mobile mapping systems
MRF	Markov random field
NDT	Nondestructive testing
PCC	Pearson correlation coefficient
PolyU	The Hong Kong Polytechnic University
SLAM	Simultaneous localization and mapping
SLIC	Simple linear iterative clustering
STFT	Short-time Fourier transform
Utilities Lab	Underground utility survey laboratory

UUIS	Underground utilities information system
VMD	Variational mode decomposition
WSD	Water Supplies Department of Hong Kong Government

Chapter 1. Introduction

1.1 Background

Urbanization is a significant global trend, with approximately 56% of the world's population—around 4.4 billion people—currently living in urban areas. Projections suggest that by 2050, this proportion will rise to nearly 70%. The rapid pace of urbanization presents substantial challenges. Cities urgently need to provide affordable housing, efficient infrastructure, and essential services (World Bank, 2024). Underground utilities, including water supply pipelines, sewer systems, energy infrastructure, and telecommunication networks, are vital for modern metropolises, acting as the backbone of urban infrastructure. As cities continue to expand, the underground environment is becoming increasingly complex (Broere, 2016; Cui et al., 2021; Su, Wang, Wang, et al., 2023). Ageing and poorly maintained infrastructure exacerbate these challenges, leading to urban disasters such as road collapses caused by voids formed from water leaks. These incidents result in fatalities and significant economic losses (Forero-Ortiz & Martínez-Gomariz, 2020; Tanoli et al., 2019; Wang & Yin, 2022), underscoring the critical need for proactive subsurface hazard monitoring to ensure urban resilience.

Nondestructive testing (NDT) technologies are essential tools for assessing hidden structures without causing damage. Ground Penetrating Radar (GPR) is a widely used NDT method that operates by emitting electromagnetic pulses into the ground and recording reflections from subsurface features. Analogous to medical imaging, which diagnoses patients without invasive procedures, GPR enables detailed imaging of underground structures while preserving their integrity. GPR has emerged as the most suitable NDT technology for underground imaging due to its advantages in:

- (1). Rapid data collection, because of the non-contact nature of electromagnetic (EM) wave physics.
- (2). The ultra-wideband and diverse frequency range of GPR wave, allows enables detection of subsurface features at varying depths depending on the different fields of applications.
- (3). Centimeter-scale resolution can be achieved, making it particularly well-suited for detailed infrastructure assessments (Lai et al., 2018).

These attributes make GPR a versatile tool for civil engineering applications, such as locating rebars in concrete (Liu et al., 2020), mapping underground utilities (Su, Wang, Li, et al., 2023), reconstructing buried stratigraphy, and investigating archaeological sites (Kwan & Lai, 2024). In addition, GPR's high sensitivity to air space that forms voids (Luo et al., 2020a; Luo & Lai, 2020) and material wetness, which are the major sources of leakage (Benedetto, 2010a; Klotzsche et al., 2018), makes it particularly effective in detecting voids and leaks.

Single-channel GPR surveys, though effective, are time-consuming and labour-intensive for extensive coverage. To address these limitations for large-scale subsurface surveys, multichannel GPR (MCGPR) systems have been developed, enabling efficient data acquisition over larger areas without disrupting urban activities. MCGPR has proven valuable in mapping underground utilities, such as manholes (Yamaguchi & Mizutani, 2021), pavement layers (Liu et al., 2016), and underground pipes (Gabryś & Ortyl, 2020; Yamaguchi et al., 2021). It has also been used for subsurface investigations, including identifying uncharted utilities and "no-man's land" for trial pits. Additionally, recent studies have demonstrated the potential of MCGPR for detecting subsurface hazards, such as cavities (Kang et al., 2019a; Yamaguchi et al., 2022a), and roadbed damage (Xu et al., 2014).

Despite its advancements, traditional GPR-based hazard detection relies heavily on static and one-off pattern recognition, which has significant limitations. This approach

identifies characteristic reflection patterns in GPR data to detect voids and leaks, but it suffers from three critical flaws:

- (1). Ambiguity in patterns: Similar reflection patterns can arise from distinct phenomena, leading to false alarms.
- (2). Limited temporal insights: Single surveys cannot capture hazard progression with time, increasing the likelihood of missed detections.
- (3). Environmental variability: Variations in subsurface conditions, particularly in host materials, reduce the consistency and accuracy of pattern recognition.

To overcome these challenges, an alternative strategy has emerged: change detection in temporal GPR data. This approach analyzes repeated GPR measurements of the same region over time, identifying dynamic changes to detect evolving hazards. Change detection offers several key advantages:

- (1). It captures the progression of hazards over time, providing a comprehensive understanding of their evolution.
- (2). By comparing data across time, it isolates dynamic subsurface changes, filtering out static features and enhancing detection accuracy.

Despite its advantages, its adoption remains limited by unresolved challenges:

- (1). Temporal data often exhibit inconsistencies due to spatial misalignment and equipment noise, complicating direct comparisons.
- (2). Distinguishing true hazards from environmental changes or equipment drift requires robust algorithms to avoid misinterpretation.

Change detection offers a complementary perspective by capturing dynamic patterns and trends over time. However, very few studies have explored change detection using multi-

temporal measurements to identify urban subsurface hazards systematically. This thesis bridges this gap by developing a systematic framework for change detection in horizontal time slices of MCGPR data. By integrating temporal analysis with unsupervised machine learning, we advance urban hazard monitoring from reactive pattern-matching to predictive, large-scale diagnostics.

1.2 Research Objectives

As established in Chapter 1.1, this research addresses critical gaps in urban subsurface hazard detection by integrating temporal GPR data with change detection algorithms. The work is structured around three interconnected objectives:

(1). **Objective 1:** Establish a Leakage **Characterization** Framework.

Design experimental protocols to systematically capture and analyze diverse leakage patterns across pipe materials and pavement types.

(2). **Objective 2:** Develop MCGPR-Specific **Error Mitigation** Workflows.

Develop solutions to address two systemic MCGPR challenges: (1) geolocation errors that distort temporal alignment, and (2) inter-channel signal variations that obscure true subsurface changes.

(3). **Objective 3:** Develop Change Detection **Algorithms** for Urban Underground Hazards.

Develop urban-scale hazard detection methods that utilize findings from Objectives 1 and 2, focusing on dynamic changes revealed by temporal GPR measurements.

1.3 Thesis Structure

The overall research workflow of this thesis is illustrated in Figure 1-1, where the three key objectives are systematically integrated into four stages of the research process. Following an extensive literature review (Chapter 2), the ultimate goal of this research was defined: to

analyze the dynamic patterns of underground hazards over time and to develop specialized detection methods.

The workflow consists of four main stages: “A. Data Acquisition” -> “B. GPR Post-Processing” -> “C. Temporal Slice Generation” -> “D. Hazards Detection”. Together, a series of approaches form a framework for addressing the listed objectives.

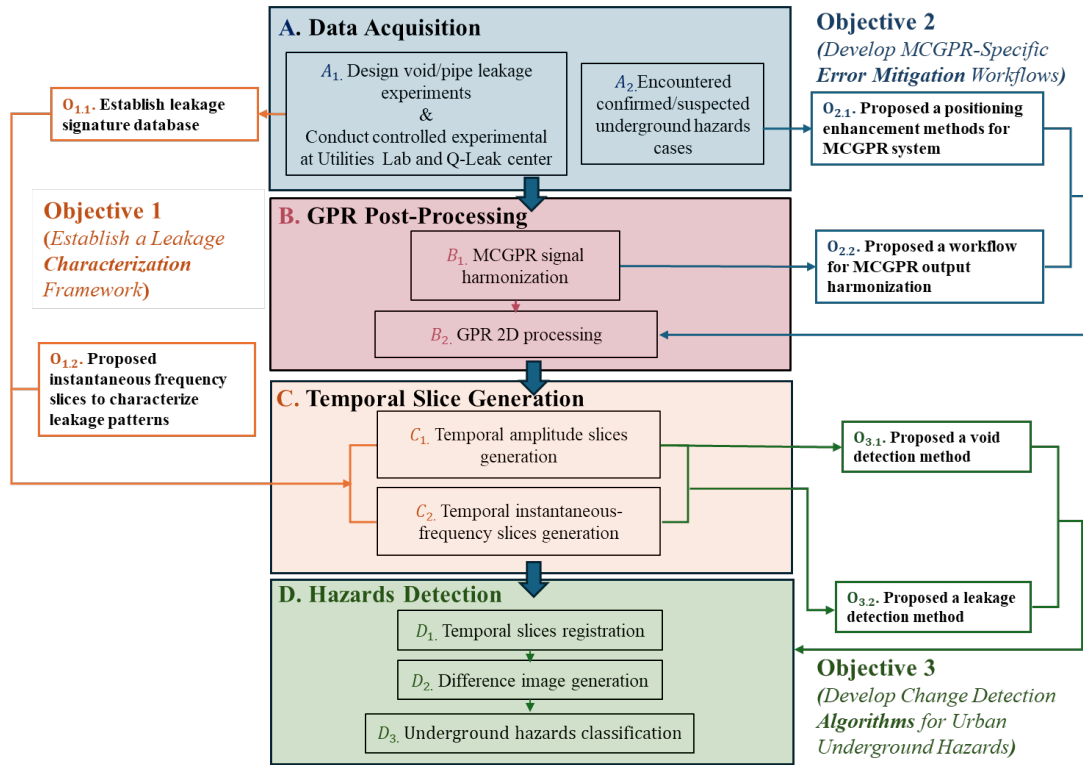


Figure 1-1 Thesis roadmap: from data acquisition to hazard detection

The initial stage of the research involved data acquisition, which was conducted using two primary sources: controlled experiments and field data collection. Controlled experiments were conducted at the Utilities Lab and the Q-Leak Center to simulate real-world underground conditions and meet specific research objectives, as illustrated in Figure 1-2. For void detection, critical parameters such as void sizes, depths, and locations were carefully predefined. For leakage scenarios, variables influencing GPR reflections—such as pipe properties, failure modes, and pavement types—were systematically designed. These controlled setups provided

comprehensive datasets that captured the complexity of underground environments. A variety of GPR systems and antennas were employed during these experiments to ensure the robustness and versatility of the collected data, which is detailed in Chapter 3.

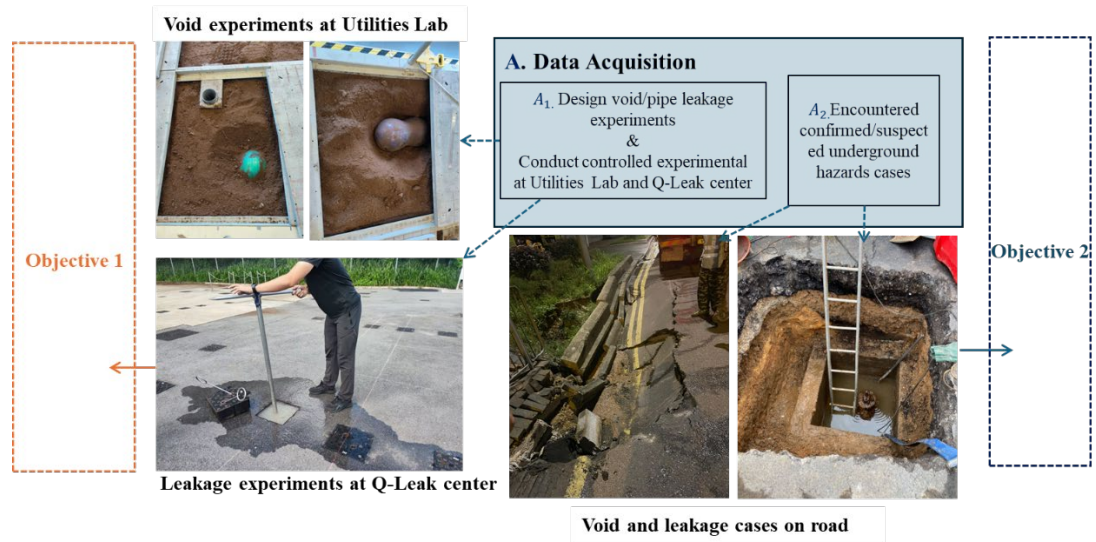


Figure 1-2 Illustration of data acquisition in this thesis

The controlled experiments also led to the development of a comprehensive leakage signature database, aimed at reducing reliance on simulated data (Figure 1-3). Leakage patterns were extended into the time-frequency domain using instantaneous frequency slices (IFS), aligned with Objective 1, which is detailed in Chapter 4.

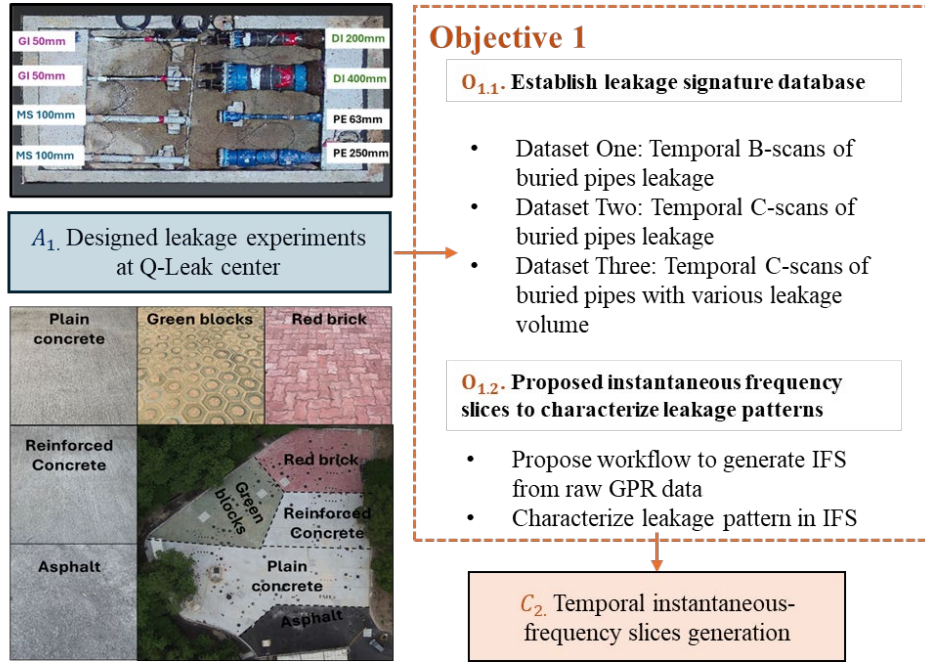


Figure 1-3 Illustration of Objective 1

Field surveys were conducted at real-world sites with confirmed cases of underground hazards, such as voids and water main leaks. These sites were identified through collaborations with government departments and other organizations. The authenticity of the field data was validated through open-up procedures conducted after the surveys. Given the environmental and logistical constraints of field sites, MCGPR systems were predominantly used, offering an efficient means of conducting large-scale surveys.

While MCGPR systems provide significant efficiency, they also present several challenges. One of the primary issues was positioning inaccuracies, which could distort the spatial alignment between GPR images and real-world geolocations. This misalignment introduced errors that complicated the interpretation of subsurface features. Additionally, signal variations inherent to MCGPR systems, such as antenna-specific gain inconsistencies, often resulted in artefacts during post-processing. These issues were critical to resolve, as they directly impacted the quality and stability of the data.

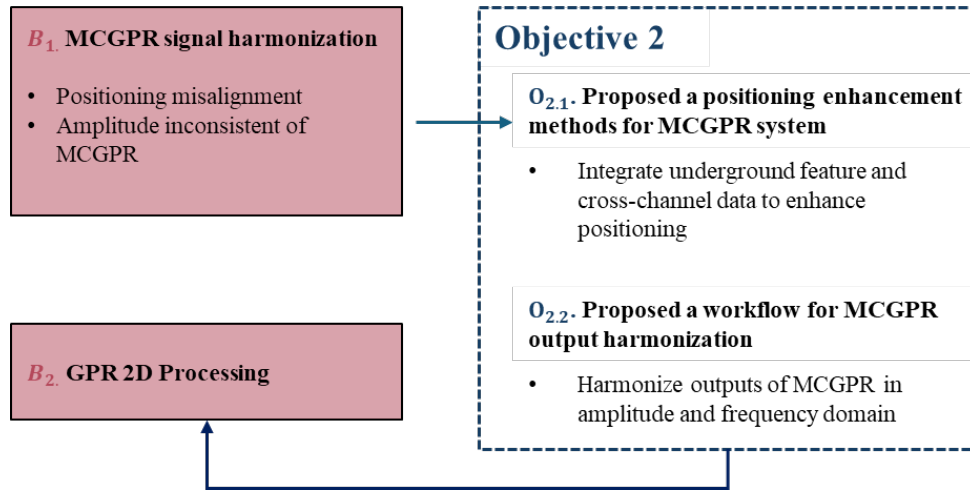


Figure 1-4 Illustration of Objective 2

To address the challenges associated with MCGPR systems, this research proposed two key solutions, illustrated in Figure 1-4. A positioning enhancement method to improve the spatial accuracy of MCGPR data by mitigating geolocation errors was developed. Also, a signal harmonization workflow was developed to address inter-channel inconsistencies and ensure consistent data quality across antenna arrays. These solutions, aligned with Objective 2, are discussed in detail in Chapter 6. Once the data was processed, temporal slices were generated to capture underground changes over time. Two types of temporal slices were created: amplitude slices, which represent changes in signal strength, and IFS, which provide frequency-based insights into subsurface patterns. These temporal slices formed the basis for subsequent change detection analyses.

The final stage of the workflow involved applying change-detection techniques to the temporal GPR slices, enabling the identification of underground hazards such as voids and leaks. By analyzing temporal changes, this approach provided a dynamic perspective on subsurface conditions, capturing the progression of hazards over time. Two novel detection methods were developed in this stage. The first method focused on void detection, leveraging temporal data to identify voids and track their development. The second method targeted the detection of pipe

leaks, integrating AS and IFS to accurately identify leakage patterns. By shifting the focus from one-off static pattern recognition to temporal analysis, these change detection techniques offer significant improvements in the accuracy and reliability of underground hazard detection. These methods, developed under Objective 3, are discussed in Chapter 5.

Chapter 7 presents a forensic investigation of a real-world leakage case, integrating findings from the three research objectives. The study began with an initial MCGPR survey over a 5.5 km road segment. Based on water pressure anomaly records, a 1.6 km subsection was identified as a suspected region of interest due to leakage for further analysis, where three additional temporal MCGPR surveys were conducted to monitor its changes. To ensure data accuracy, geolocation misalignments and signal variations were corrected using the positioning enhancement and signal harmonization workflows detailed in Chapter 6. These corrections addressed key MCGPR challenges, such as spatial misalignment and antenna-specific inconsistencies, ensuring reliable data for subsequent analysis. Processed data were used to generate temporal amplitude slices (AS) and IFS following the methodologies in Chapter 4. The amplitude slices captured changes in signal strength, while the IFS provided frequency-based insights into subsurface patterns. These outputs formed the basis for change detection analyses. The change detection methods developed in Chapter 4 and Chapter 5 were applied to identify leakage in the temporal data. By leveraging temporal analysis, the leakage signature database, and IFS, these methods effectively characterized leakage patterns and tracked the progression of leakage. This case study demonstrates the practical applicability of the proposed methodologies, showcasing their ability to enhance the accuracy and efficiency of GPR-based underground hazard detection in real-world conditions.

This research introduces several novel approaches to tackle the challenges of underground hazard detection using temporal GPR data. The main findings, contributions, and practical recommendations for addressing real-world complexities are summarized in Chapter

8, along with suggestions for future research directions.

Chapter 2. Literature Reviews

Chapter 2 established the theoretical and methodological foundation for detecting subsurface hazards using GPR. It begins with an in-depth examination of GPR's operational principles, electromagnetic wave properties, and interactions with subsurface materials—critical for understanding its capacity to identify anomalies like voids and leaks. The discussion then advances to GPR data processing and imaging techniques, detailing how raw signals are transformed into interpretable visualizations through post-processing workflows and advanced algorithms. Before exploring GPR theory and data processing, the chapter introduced the mechanisms underlying subsurface hazards, with a particular focus on the processes that lead to the formation of voids, road collapses, and pipe degradation or leakage. This section bridges the theoretical underpinnings of GPR with its practical applications.

The final section reviews current underground hazard detection methods, categorizing them into two primary strategies: one-shot pattern recognition and temporal measurement techniques. One-shot pattern recognition is widely used for detecting static hazards, such as voids and leaks, by identifying characteristic patterns in GPR data. In contrast, temporal measurement techniques monitor changes over time, enabling the detection of dynamic subsurface conditions. These two approaches highlight the dual role of GPR in both immediate hazard detection and long-term monitoring of subsurface environments. Through these discussions, Chapter 2 establishes a comprehensive foundation for understanding the capabilities, limitations, and applications of GPR in addressing the challenges associated with underground hazard detection. This foundational knowledge supports the development of the novel methodologies proposed in subsequent chapters.

2.1 Mechanism of Subsurface Hazards

Urban underground spaces are increasingly affected by human activities, such as the

construction of infrastructure, pipeline installation, and heavy traffic, leading to frequent subsidence incidents (Momeni et al., 2023). These ground collapse events are often sudden and unpredictable, posing significant risks to urban infrastructure. Recent studies have identified underground pipeline leakage, rainfall, and geological conditions as the primary contributors to urban ground collapse (Rahnema & Mirasi, 2012; Rasol et al., 2022). Wang and Xu analyzed collapse incidents in China and identified the primary causes of urban ground collapse as underground pipeline leakage, rainfall, and geological conditions (Wang & Xu, 2022). While rainfall and geological conditions are highly location-specific and influenced by regional geographic or climatic factors, pipeline leakage is a universal issue affecting urban environments worldwide. Research has shown that underground pipeline breakage is the most prevalent factor leading to urban ground collapses (Tan et al., 2022). For instance, Hu et al. conducted a statistical analysis of ground collapse cases in China and reported that 55% of all such incidents were caused by pipeline breakage. Unlike other factors, the concealed nature of underground pipelines makes it particularly challenging to observe the processes of water seepage and soil erosion, as well as their cumulative impacts on ground stability, after a pipeline failure. To better understand the mechanisms of ground collapse caused by pipeline breakage, researchers have focused on the sequential disaster-causing process:

“Pipeline thinning → pipeline breakage → water leakage → soil erosion → cavity evolution → ground collapse.”

This sequence underscores the critical role of pipeline breakage as the initial trigger that leads to water leakage, soil erosion, void formation, and ultimately ground instability. Given the complex interplay of these processes, it is crucial to investigate both the factors contributing to pipeline degradation and leakage, as well as the mechanisms of subsurface void formation. The following subsections explore these aspects in detail, shedding light on the underlying causes of urban ground collapses.

2.1.1 Factors Contributing to Pipeline Degradation and Leakage

Underground pipelines are essential infrastructure widely used for the transmission of water, gas, and sewage. In the metropolis of Hong Kong, which features a typical subtropical climate and a complex water mains network constructed over an extended period, the degradation of water pipes presents significant challenges. A comprehensive dataset on freshwater mains and their failures was acquired from the Water Supplies Department (WSD) of the Government of Hong Kong, SAR. This dataset includes a water mains network spanning 8,400 km across the entire city and documents 20,589 failure cases from 2000 to 2024 (Water Supplies Department, 2025). The age of these mains ranges from newly installed to over 60 years old. Furthermore, the conditions of the buried soil play a crucial role in the degradation process of these pipes. Based on the classification of the factors, the corrosion causes are categorized as environmental factors, pip-related factors, and operation factors, as shown in Figure 2-1.

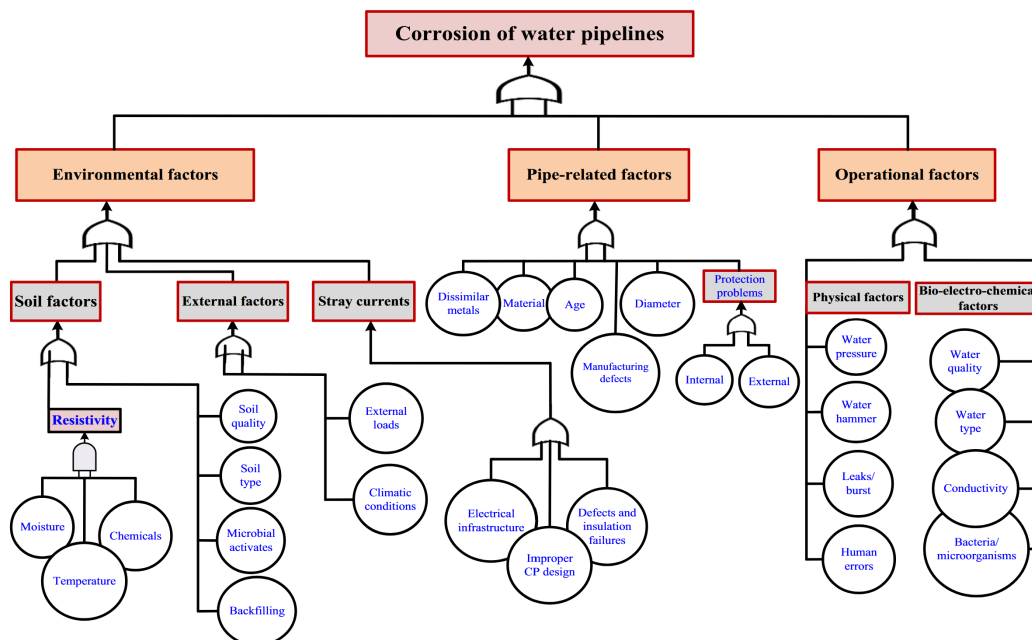


Figure 2-1 Fault tree diagram of the corrosion causes for water pipelines. A comprehensive summary of the corrosion causes in water pipes based on bibliometric and systematic analysis (Hussein et al., 2023)

(1). Environmental factors

Environmental factors are some of the most significant contributors to pipeline corrosion, particularly through soil properties, external loads, and stray currents.

Soil factors significantly influence the corrosion of underground water pipes. Different soil types, such as clay, sand, gravel, and loam, can impact corrosion due to their varying textures and grain sizes (Wang et al., 2019). For instance, sandy soils with high alkalinity may cause general corrosion, while soil particle size can affect corrosion behaviour, as seen in studies where increased particle size in sandy soils reduces corrosion rates (He et al., 2015). Soil quality, including properties like water content, resistivity, pH, redox potential, and chemical composition, also plays a crucial role. Soil pH, particularly, is a significant factor; highly acidic soils ($\text{pH} < 5$) can accelerate corrosion by preventing the formation of protective layers on steel. Soil acidity, influenced by environmental factors like acid rain and microbial activity, can cause localized pitting corrosion (Ismail & Elshamy, 2009). Differential aeration, where low-aeration soils act as anodes and high-aeration soils as cathodes, also affects corrosion, especially in moist conditions. Soil resistivity, inversely related to corrosion, is influenced by moisture content, ion concentrations, and temperature. High moisture content and chloride concentrations lower soil resistivity, increasing corrosion risk (Gupta & Gupta, 1979; Noor & Al-Moubaraki, 2014). Soil temperature, though less studied, can affect corrosion rates, with higher temperatures generally increasing corrosion (Nie et al., 2009). Chemical substances in soil, such as chlorides and sulfates, directly impact soil resistivity and metal dissolution reactions, posing significant corrosion risks (Żakowski & Sokółski, 1999). Microbial activities in soil can lead to microbiologically influenced corrosion, producing corrosive substances like carbon dioxide and hydrogen sulfide. Lastly, backfilling practices, which involve burying pipes in a mixture of excavated soil strata, can affect corrosion, with soil homogeneity around the pipe being crucial for

minimizing corrosion risks.

External loads, including heavy traffic and shifting rocks, induce stress on pipelines, creating localized electrochemical cells and increasing corrosion susceptibility (Seghier et al., 2020). Stray currents from nearby electrical infrastructure, such as railways and power lines, are another major factor. These currents enter pipelines at cathodic regions and exit at anodic regions, causing localized corrosion. Poorly designed cathodic protection systems or insulation failures significantly exacerbate this issue, requiring effective CP design to mitigate the risk (O'Day, 1989).

(2). Pipe-related factors

In Hong Kong, the mains are constructed from more than 10 types of materials, with seven primary materials officially regarded as standard: asbestos cement, ductile iron, unlined galvanized iron, lined galvanized iron, mild steel, polyethene, and unplasticized polyvinyl chloride (Shi et al., 2013). Each material has its advantages and drawbacks. Metal pipes, such as ductile iron and mild steel, offer high strength, durability, and pressure capability but are susceptible to corrosion. On the other hand, PVC pipes provide excellent corrosion and abrasion resistance but are prone to creep and a reduction in strength with increasing temperature. Additionally, the use of salt water and chlorine dioxide as disinfectants in drinking water has historically impacted the integrity of water pipes.

Pipe-related factors significantly influence the corrosion of buried water pipelines. These factors include pipeline age and exposure time, materials, dissimilar metals, manufacturing defects, pipeline diameter, and protection layers. The age and exposure time of a pipeline are critical, as corrosion susceptibility generally increases with age, though this varies by pipe type. For instance, cast iron pipes often experience more breaks and leaks as they age. The materials used in pipeline construction, including the type of metal and its

alloy composition, play a crucial role in corrosion resistance. High-alloy steels typically exhibit lower initial corrosion rates. Galvanic corrosion can occur when dissimilar metals are welded together, leading to accelerated corrosion in anodic areas (Wasim et al., 2018). Manufacturing defects, such as improper bending or coating flaws, can create environments conducive to unexpected corrosion. Pipeline diameter also affects corrosion; large-diameter pipes have more exposed surface area, while small-diameter pipes often face higher loads and less protection. Lastly, protection layers, both passive (e.g., coatings, linings) and active (e.g., cathodic protection), are essential for mitigating corrosion. The absence or failure of these protective measures can significantly increase corrosion rates, highlighting the importance of proper maintenance and protection strategies (Rajani & Kleiner, 2003).

(3). Operational factors

Operational factors significantly influence the corrosion of water pipelines and can be categorized into physical and bio-electro-chemical factors. Physical factors include fluctuating water pressure, which causes fatigue and micro-cracks; water hammer from sudden valve closures, inducing cracks; human errors during manufacturing, installation, or operation; and leaks/bursts that increase moisture and introduce corrosive elements like sulfate and chlorine. Bio-electro-chemical factors encompass water quality (pH, temperature, oxygen, chemical content), water type (e.g., desalinated water), conductivity, and bacteria/microorganisms. High water conductivity accelerates corrosion, while sulfate-reducing bacteria (SRB) can form biofilms and promote microbiologically influenced corrosion (Dao et al., 2021; Suban et al., 2010). Proper management and monitoring of these factors are crucial to mitigate corrosion risks in water pipelines.

2.1.2 Causes and Mechanisms of Subsurface Void Formation

Following pipeline degradation and leakage, water seepage and soil erosion lead to the

formation of subsurface voids, which destabilize the overlying ground and result in surface settlement or collapse. This section focuses on the processes occurring after pipeline breakage, specifically, water leakage leading to soil erosion, cavity formation, and eventually ground collapse.

(1). Soil erosion due to water leakage

Water leakage is a critical driver of internal erosion, which occurs when seepage forces dislodge soil particles, transporting them downstream and creating voids within the soil matrix. Internal erosion mechanisms include suffusion (washing away of fine particles), backward erosion (soil removal at pipe exits), concentrated erosion (along cracks), and contact erosion (at interfaces between soil layers) (Fell et al., 2003). Seepage flow exerts hydrodynamic forces on soil particles, leading to detachment when shear stress exceeds the critical threshold of the soil material (Briaud et al., 2001). As particles are transported, seepage-induced shear stress enlarges voids, forming pipes or channels that destabilize the soil (Wan & Fell, 2004). At the soil-structure interface, the mechanical disparity between soil and hydraulic structures makes these areas more susceptible to seepage channel formation. These channels expand over time, creating cavities and unstable zones (Chen et al., 2022; Kaoser et al., 2006). Experimental studies using tools like flume tests, jet erosion tests, and erosion function apparatus tests have demonstrated that soil erosion is influenced by soil particle size distribution, hydraulic gradients, flow velocity, and soil density (Foster & Fell, 2001; Moffat et al., 2011; Pachideh & Hosseini, 2019).

(2). Cavity evolution

The continuous erosion of soil leads to the formation of localized cavities or piping within the soil. These cavities grow due to a combination of hydraulic, mechanical, and geometric factors. For instance, water flow through pre-existing pores can mobilize fine

particles, altering the distribution and content of soil particles, and thereby weakening the soil's mechanical properties (Mehdizadeh et al., 2017; Salazar et al., 2015). Cavity formation is further influenced by factors such as soil cohesion, inclination, and stress conditions, which determine the stability of the surrounding soil matrix (Chang & Zhang, 2013). If left unmitigated, the expansion of these cavities can eventually lead to structural instability, uneven settlement, and ground collapse.

(3). Ground collapse

The final stage of void formation occurs when cavities become large enough to compromise the stability of the overlying ground. Ground collapse is often triggered when erosion reaches a critical point where the soil can no longer support its weight or the imposed load. This failure mechanism is common in embankments and other hydraulic structures, where uncontrolled seepage leads to piping, cavity expansion, and eventual collapse (Foster et al., 2000; Kaoser et al., 2006; Sterpi, 2003). The process is often sudden, as erosion progresses invisibly until a critical point is reached where the remaining soil can no longer support the load above (Ali & Choi, 2019; Weil, 1995). This underscores the importance of early detection to prevent severe damage.

2.2 Theory of GPR

The theory of GPR forms the foundation for understanding its operation, signal propagation, and interaction with subsurface materials. This section explores the principles of GPR, focusing on antenna operations, electromagnetic (EM) wave properties, and their interactions with subsurface media. These topics are essential for interpreting GPR data, extracting subsurface hazards' features, and addressing practical challenges. In practice, contrasts in dielectric permittivity, conductivity, and magnetic permeability manifest as amplitude variations and phase changes, enabling the estimation of layer boundaries, velocities,

thicknesses, and material properties.

2.2.1 GPR Antenna and Operations

In GPR surveys, a transmitting antenna emits a radar signal, which is essentially a pulse of EM waves. This signal propagates downward through the material until it encounters an interface with electrical properties differing from the surrounding medium. At this interface, part of the radar energy is reflected and captured by the receiving antenna, while the remaining energy penetrates further into the subsurface. Figure 2-2 illustrates a typical setup for a common-offset GPR survey, where hyperbolic reflections are formed due to the varying travel times (t_1 and t_2) as the antenna moves horizontally (x).

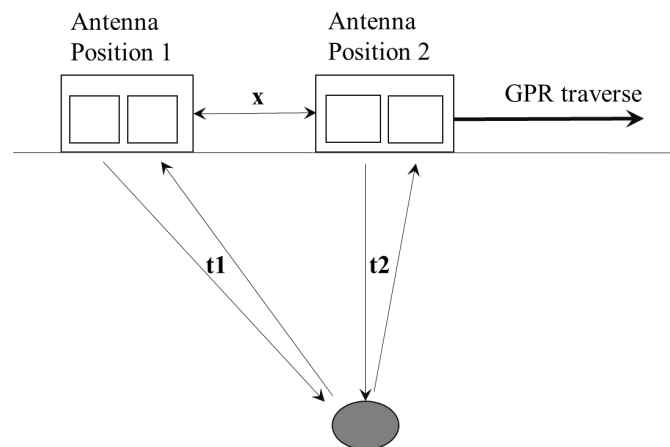


Figure 2-2 Hyperbolic reflection from common offset GPR, t_1 and t_2 are the GPR wave travel time at antenna positions 1 and 2, respectively, while x is the moving distance of the antenna (Lai et al., 2016a)

GPR systems comprise interconnected components that work together to generate and process radar pulses. Figure 2-3 provides a simplified diagram of the impulse radar technique. The core elements of an impulse radar system include a timing unit, transmitter electronics (e.g., a pulse generator), a transducer element (the antenna), and analog-to-digital conversion circuitry. The timing unit initiates the process by triggering the transmitter, which generates a short direct current pulse. This pulse is transmitted to the antenna, which emits the radar signal.

The resulting output signal is described as the convolution of the pulse's transfer function, $p(t)$, with the antenna's transfer function, $a(t)$, as expressed in Equation [2-1]:

$$\omega(t) = 2\pi f(t) = p(t)a(t). \quad [2 - 1]$$

The antenna's tuning properties act as a band-pass filter, shaping the emitted pulse. The specific characteristics of the pulse are influenced by the antenna's transfer function. In an ideal scenario, $f(t)$ would equal $\delta(t)$, representing the Dirac delta function or a perfect impulse. However, in real-world applications, the emitted energy is centred around a frequency defined by the antenna's design. The bandwidth of this energy and the duration of the transmitted pulse depend on both the applied pulse width and the properties of the antenna, as indicated by Equation [2-1]. Broadband transducer elements, such as bow-tie antennas or resistively loaded dipole antennas, are commonly used because of their dependable performance (Johnson and Jasik, 1984). The energy reflected from subsurface objects is captured by the receiving antenna and recorded or digitized over time for further analysis.

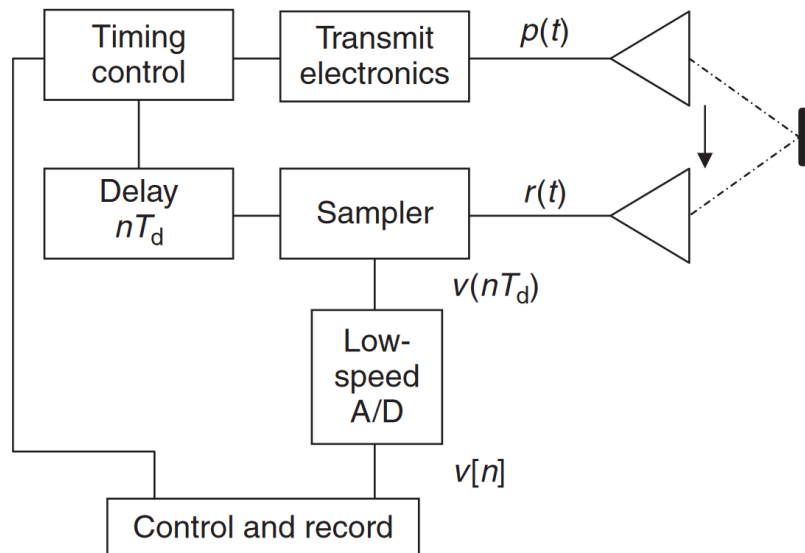


Figure 2-3 Simplified diagram of the impulse radar technique (Jol, 2009a)

2.2.2 EM Wave Properties

The governing principles and behaviour of this EM wave propagation are described by Maxwell's equations (Equations [2-2] – [2-5]). These equations outline the relationship between the time-varying electric field ($\vec{E}$) and magnetic fields ($\vec{B}$), which are interdependent in both space and time. This interdependence facilitates the propagation of EM waves in free space as well as in various media.

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad [2-2]$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}, \quad [2-3]$$

$$\vec{\nabla} \cdot \vec{D} = q, \quad [2-4]$$

$$\vec{\nabla} \cdot \vec{B} = 0. \quad [2-5]$$

In the above equations, $\vec{E}$ is the electric field strength vector; $\vec{B}$ is the magnetic flux density vector; $\vec{D}$ is the electric displacement vector; $\vec{H}$ is the magnetic field intensity; q is the electric charge density and $\vec{J}$ is the electric current density vector. The properties of different materials are characterized by the constitutive relations (Equations [2-6] – [2-8]), which quantitatively describe the material's dielectric permittivity ($\epsilon, F/m$), electrical conductivity ($\sigma, S/m$), and magnetic permeability ($\mu, H/m$).

$$\vec{D} = \epsilon \vec{E}, \quad [2-6]$$

$$\vec{J} = \sigma \vec{E}, \quad [2-7]$$

$$\vec{B} = \mu \vec{H}. \quad [2-8]$$

The dielectric properties of subsurface materials, including permittivity and conductivity, are critical in GPR applications. Permittivity describes a material's capacity to store and release electromagnetic energy, while conductivity reflects its ability to conduct electric charges. For practical purposes, permittivity is often simplified to its relative permittivity (κ), which is the

real component of permittivity normalized to the permittivity of free space (ϵ_0). Relative permittivity is typically used in GPR applications because it provides a straightforward way to estimate wave velocity and approximate radar signal behaviour in subsurface materials. Table 2-1 lists the relative permittivity values of common subsurface materials at 100 MHz, which are often assumed constant for practical calculations. These values allow for simplified estimations of wave velocity and wavelength, but may not fully capture the dynamic behaviour of permittivity at different frequencies.

Table 2-1 Typical values of relative permittivity (real component) and static conductivity for common subsurface materials at an antenna frequency of 100 MHz

Material	Static conductivity, σ_s (mS/m)	Relative permittivity, ϵ_{ave}
Air	0	1
Clay – dry	1–100	2–20
Clay – wet	100–1000	15–40
Concrete – dry	1–10	4–10
Concrete – wet	10–100	10–20
Freshwater	0.1–10	78 (25 °C)–88
Freshwater ice	1–0.000001	3
Seawater	4000	81–88
Seawater ice	10–100	4–8
Permafrost	0.1–10	2–8
Granite – dry	0.001–0.00001	5–8
Granite – fractured and wet	1–10	5–15
Limestone – dry	0.001–0.0000001	4–8
Limestone – wet	10–100	6–15
Sandstone – dry	0.001–0.0000001	4–7
Sandstone – wet	0.01–0.001	5–15
Shale – saturated	10–100	6–9
Sand – dry	0.0001–1	3–6
Sand – wet	0.1–10	10–30
Sand – coastal, dry	0.01–1	5–10
Soil – sandy, dry	0.1–100	4–6
Soil – sandy, wet	10–100	15–30
Soil – loamy, dry	0.1–1	4–6
Soil – loamy, wet	10–100	10–20
Soil – clayey, dry	0.1–100	4–6
Soil – clayey, wet	100–1000	10–15
Soil – average	5	16

The behaviour of permittivity is not static; it is influenced by relaxation phenomena,

which occur when the polarization of a material cannot keep pace with the oscillating electric field as frequency increases. This dynamic behaviour is described by the complex permittivity, which consists of two components: The real component (ϵ'), which represents the material's ability to store energy from an electric field. This is what relative permittivity typically refers to. The imaginary component (ϵ''), which accounts for energy dissipation as heat, contributes to signal attenuation. At lower frequencies, the real component of permittivity is higher because material polarization mechanisms (e.g., dipolar polarization) can align with the oscillating electric field. As frequency increases and approaches the relaxation frequency, polarization mechanisms fail to keep up, causing a reduction in the real component and an increase in energy loss, represented by the imaginary component. This behaviour is illustrated in Figure 2-4, which shows how different polarization processes contribute to the overall permittivity response across the electromagnetic spectrum.

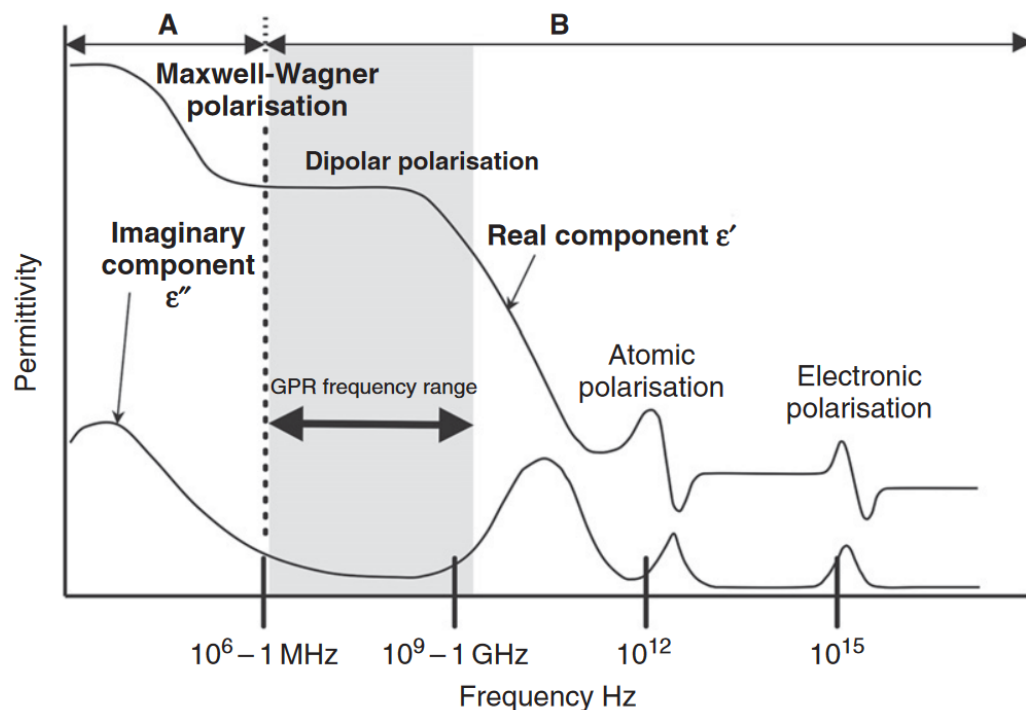


Figure 2-4 Permittivity relaxation phenomena associated with an idealized, lossy dielectric material. Region A represents the polarisations associated with free-charge and material interactions whilst

region B represents molecular and atomic polarisations associated with bound charges (Adapted from (Daniels, 1996))

Water content has a profound effect on the permittivity of subsurface materials, making it one of the most critical factors influencing GPR surveys. The high dielectric permittivity of water ($\kappa \approx 80$ at low frequencies) compared to most dry materials (e.g., soil, rock, or air, with $\kappa \approx 3 - 10$) significantly increases the overall permittivity of water-bearing materials. In soils and porous media, the presence of both free water (e.g., water in pore spaces) and bound water (e.g., water adsorbed onto particle surfaces) contributes to the observed permittivity. Free water, due to its mobility, exhibits strong dipolar polarization, which dominates the dielectric response, especially at lower frequencies. Bound water, on the other hand, is less mobile and contributes to permittivity with weaker polarization effects.

2.2.3 Interactions between GPR and Materials

EM wave interactions with materials are governed by three key properties: propagation velocity (v), attenuation rate (α), and electromagnetic impedance (Z). These properties are determined by the material's permittivity, conductivity, and permeability, which directly influence how the EM waves travel through and interact with the medium. For a simple medium with fixed material properties, these wave characteristics are most easily expressed under the assumption of a sinusoidal time variation. At low frequencies, wave properties exhibit diffusive behaviour, where they depend on the square root of angular frequency ($\sqrt{\omega}$). As frequency increases, the behaviour transitions to wave propagation at a critical point called the transition frequency ($f_t = \frac{\sigma}{2\pi\epsilon}$). This frequency marks the point at which wave properties change from being frequency-dependent to frequency-independent. Figure 2-5 illustrates this transition by showing how velocity and attenuation vary across different frequency ranges for a simple medium with non-dispersive physical properties.

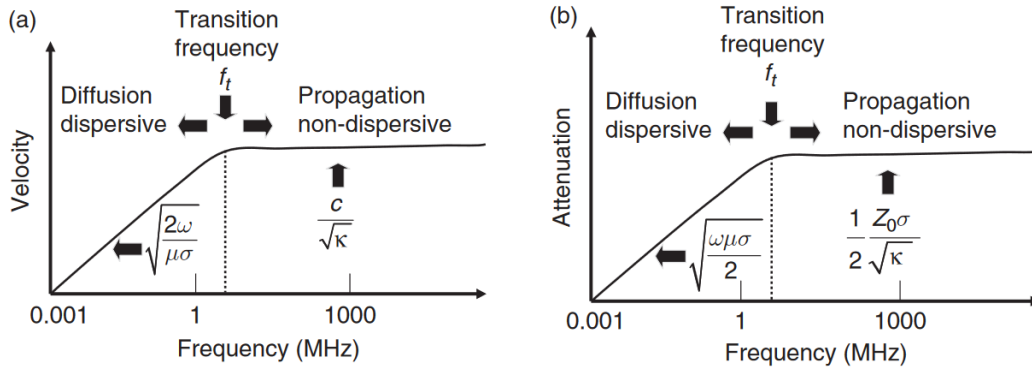


Figure 2-5 Variation in velocity and attenuation in a simple medium with nondispersive physical properties, c and Z_0 are the velocity and impedance of free space (i.e., a vacuum) (Jol, 2009a)

2.2.3.1 Dispersion

Figure 2-5 demonstrates that at low frequencies, velocity and attenuation exhibit a strong frequency dependence, with both parameters varying as a function of $(\sqrt{\omega})$. However, at the transition frequency (f_t), the behaviour stabilizes, forming a plateau where velocity and attenuation become constant. This plateau region is particularly relevant to GPR operations, as it corresponds to frequencies where radar waves propagate without dispersion. In this high-frequency regime, all frequency components travel at the same velocity and experience the same attenuation, ensuring that impulsive signals maintain their shape during propagation (Annan, 1996). This stable propagation behaviour is essential for accurate imaging and interpretation of subsurface features.

The relationships between velocity, attenuation, and impedance in this plateau region can be expressed using the following equations:

$$v = \frac{1}{\sqrt{\epsilon \cdot \mu}} = \frac{c}{\sqrt{\kappa}}, \quad [2 - 9]$$

$$\alpha = \sqrt{\frac{\mu}{\epsilon}} \cdot \frac{\sigma}{2} = Z_0 \frac{\sigma}{2 \cdot \sqrt{\kappa}}, \quad [2 - 10]$$

$$Z = \sqrt{\frac{\mu}{\epsilon}} = \frac{Z_0}{\sqrt{\kappa}}, \quad [2 - 11]$$

where c is the speed of light, Z_0 is the impedance of free space, and κ is the relative

permittivity (dielectric constant). These equations demonstrate how wave velocity and attenuation are influenced by the material's electrical properties. Notably, the real part of permittivity (ϵ), which is mainly determined by the material's water content, plays a key role in wave velocity estimation. Permittivity and conductivity show strong frequency dependence. As frequency increases, permittivity generally decreases, while conductivity increases (Lai et al., 2014).

2.2.3.2 Scattering

In heterogeneous subsurface environments, such as soils with voids or leakage zones, scattering becomes a significant factor. Scattering occurs when radar waves encounter interfaces or inclusions that have contrasting dielectric properties. The behaviour of scattering depends on the ratio of the wavelength (λ) to the size of the scatterer (σ), as shown in Figure 2-6.

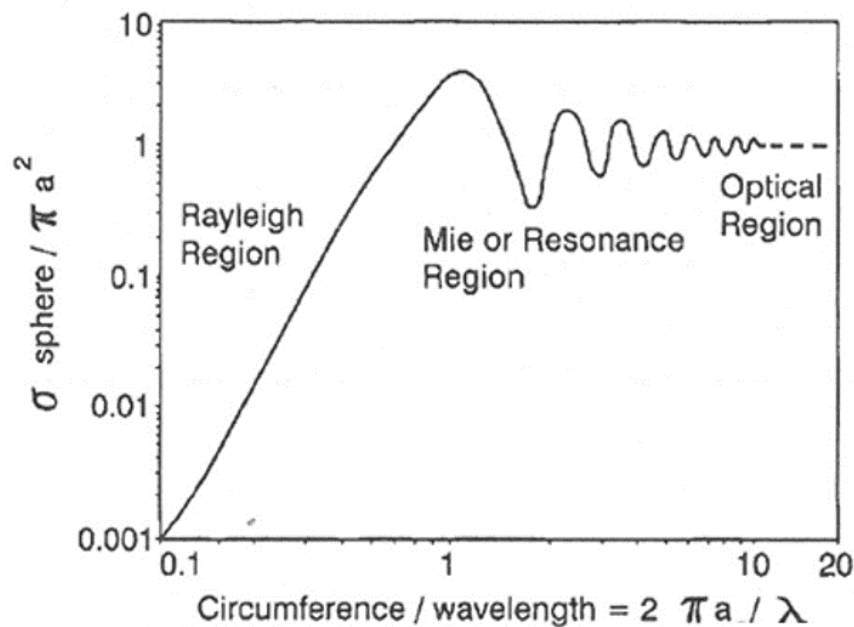


Figure 2-6 The illustration of scattering with object sizes a (Annan, 2004), σ is the ratio of GPR wavelength to the size of the scatterer, and λ denotes the wavelength

Figure 2-6 illustrates how the wavelength-to-scatterer-size ratio governs the nature of

scattering. When scatterers are much smaller than the wavelength ($\sigma \ll \lambda$), Rayleigh scattering dominates. This results in weak, diffuse reflections and lower detectability of small-scale subsurface features. Mie scattering occurs when the scatterer size is comparable to the wavelength ($\sigma \approx \lambda$), producing stronger and more directional reflections. Finally, in the optical scattering regime, where scatterers are significantly larger than the wavelength ($\sigma \gg \lambda$), reflections are optimal, and signal losses are minimal. These scattering mechanisms are particularly important for GPR applications. For example, air-filled voids often fall into the optical scattering regime, producing distinct and detectable reflections due to their large size relative to the radar wavelength. Leakage zones, on the other hand, may exhibit Mie or Rayleigh scattering, depending on the size and distribution of water-saturated areas or material heterogeneities. Leakage zones significantly influence scattering behaviour in GPR surveys due to changes in moisture content, but the nature of this influence varies depending on the surrounding material and environmental conditions.

In urban environments, where water-impermeable layers (e.g., concrete, asphalt, or compacted clay liners) are often present, leakage tends to increase subsurface heterogeneities. As water accumulates above the impermeable layer, it saturates localized regions of the host material, creating sharp contrasts in dielectric properties between saturated and unsaturated zones. These contrasts enhance scattering effects, with the specific scattering mechanism (Rayleigh, Mie, or optical) determined by the size of the heterogeneities relative to the GPR wavelength. For example, small-scale pockets of water saturation may produce Rayleigh scattering, while larger saturated regions near impermeable structures may cause Mie or optical scattering, resulting in stronger and more directional reflections. In contrast, in natural materials like soil or sand, increased water content due to leakage may homogenize the host material, as the water diffuses more uniformly into the porous matrix. This homogenization reduces dielectric contrasts between adjacent regions, diminishing scattering effects and making

leakage zones less distinguishable in GPR data. However, localized variations in moisture content, such as those caused by differences in grain size, porosity, or permeability, can still create scattering, especially in layered soils with variable composition. The interplay between moisture-induced homogenization and heterogeneity thus plays a critical role in determining the scattering behaviour and the detectability of leakage zones. Understanding these contrasting effects is essential for interpreting GPR data in varied environments and optimizing survey parameters for leakage detection.

2.2.3.3 Reflection, Deflection, and Transmission

GPR methods primarily rely on the detection of reflected or scattered signals. The reflection of EM waves occurs at interfaces where the dielectric properties of materials differ. This phenomenon is governed by Snell's Law, as expressed in Equation [2-12]:

$$\frac{\sin \theta_1}{v_1} = \frac{\sin \theta_2}{v_2}, \quad [2 - 12]$$

where θ_1 and θ_2 are respectively the angle of incidence and refraction, while v_1 and v_2 are the propagation velocities in the respective materials. Changes in the wave velocity across these interfaces cause refraction and determine the reflection strength, enabling the detection of buried objects.

One practical application of wave velocity estimation is the determination of object depths using two-way travel time (see Figure 2-2). As an EM wave travels downward, it is reflected by a subsurface object and returns to the surface. The total travel distance is twice the depth of the object, as described in Equation [2-13]:

$$d = \frac{vt}{2}, \quad [2 - 13]$$

where d is the depth of the buried reflector, v is the wave velocity, and t denotes the travel time. Wave velocity estimation is influenced by several factors, including the homogeneity of

the host material, the selection of the velocity model, the orientation of the GPR profile, and the geometry of reflectors.

Recent advancements have addressed these challenges, improving the accuracy of velocity estimation. Notably, (Sham & Lai, 2016) and (Lai et al., 2016) incorporated object diameter into wave velocity estimation models. (Xie et al., 2018) introduced the effect of distance and angle between the GPR device and a linear object in the wave velocity estimation. These developments have significantly enhanced the reliability of wave velocity estimation, enabling more accurate interpretations of GPR data and improving the detection of subsurface features.

2.3 GPR Processing and Imaging

GPR is a powerful NDT technology widely used for subsurface investigations. Its ability to capture high-resolution electromagnetic signals enables precise mapping of subsurface features, such as buried objects, layers, and voids. However, the raw data acquired from GPR surveys often contain noise, distortions, and artefacts that obscure meaningful patterns. Proper data processing and representation are therefore critical for maximizing the interpretability and reliability of GPR results.

This section explores the key stages of GPR signal processing and imaging, beginning with the fundamental data representations (A-scans, B-scans, and C-scans). Each representation serves as the foundation for visualizing subsurface features in increasing levels of detail, from raw point measurements to comprehensive 3D representations. Next, the standard post-processing techniques used to enhance data quality are discussed, including filtering, time-varying gain, and migration. Finally, advanced imaging techniques, such as C-scan generation and 3D data visualization, are introduced as tools to overcome the limitations of traditional 2D radargrams.

2.3.1 GPR Data Representations

The accurate interpretation of GPR data depends on a thorough understanding of how the signals interact with subsurface material properties and on the ability to represent these signals in formats that faithfully capture the true subsurface conditions. GPR data is typically represented in three formats—A-scan, B-scan, and C-scan—corresponding to one-dimensional, two-dimensional, and three-dimensional data formats, respectively (see Figure 2-7). Each format offers unique insights into the subsurface environment and serves specific purposes in GPR imaging and analysis.

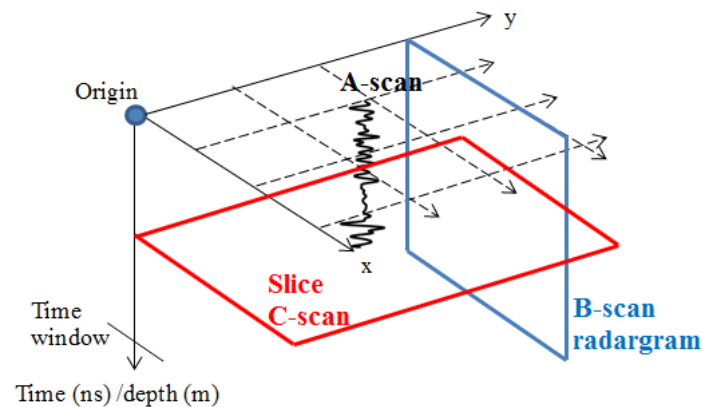


Figure 2-7 Illustration of the different GPR data representations (A-scan, B-scan / radargram, C-scan / slice)

A-scans represent the most basic 1D data format in GPR and display vertical reflections at a single survey point. These reflections provide information about waveform amplitude, signal phase, and estimated propagation velocity, all of which are influenced by the material properties of the subsurface. While simple, A-scans are essential for analyzing raw signal characteristics and can be further processed using forward and inverse modelling techniques to reconstruct the underground environment.

B-scans, also known as radargrams, offer 2D cross-sectional views of the subsurface by combining sequential A-scans collected along a survey line. These vertical profiles are widely used for visualizing subsurface features such as layers, voids, or buried objects. B-scans

are especially useful for identifying hyperbolic patterns caused by point reflectors, which can be further analyzed to estimate the depth and geometry of underground features. However, B-scans are limited to vertical representations and may not fully capture the lateral extent of subsurface anomalies.

C-scans provide 3D representations of the subsurface by capturing horizontal sections, or "depth slices," at specified depths. These depth slices offer layered, intuitive visualizations of the underground environment, making them especially valuable for interpretation. Unlike A-scans and B-scans, which focus on vertical profiles, C-scans summarize GPR properties in horizontal slices, enabling a more comprehensive understanding of subsurface conditions. Recent advancements in 3D GPR data representation techniques, such as isosurfaces, time-frequency analyses, energy-based semantic imaging, and feature-enhanced visualization methods (Böniger & Tronicke, 2010a; Böniger & Tronicke, 2010b; Leckebusch, 2003) have significantly improved the accuracy and interpretability of C-scans. These methods allow for the analysis of large-scale subsurface volumes with greater reliability, facilitating the identification and characterization of underground features.

The raw data collected during GPR surveys often contains noise, distortions, and artefacts that obscure meaningful patterns. To interpret these representations accurately, data must undergo systematic post-processing to enhance signal clarity and extract essential subsurface features.

2.3.2 Current Standard GPR Post-processing

Daniels et al. emphasized that the purpose of GPR data processing is to improve the data's appearance without altering its fundamental interpretation (Daniels, 1996). Before interpreting GPR radargrams, proper processing is essential to extract valuable information while filtering out irrelevant noise. Over decades of development, GPR signal processing

techniques have matured, providing reliable tools to enhance data quality and extract meaningful subsurface details. Jol outlined a typical GPR data processing workflow, as illustrated in Figure 2-8 (Jol, 2009a). The goal is to transform raw radar data into interpretable radargrams, where recognizable patterns—such as hyperbolas, layers, and other textures colloquially referred to as "happy patterns"—become visible.

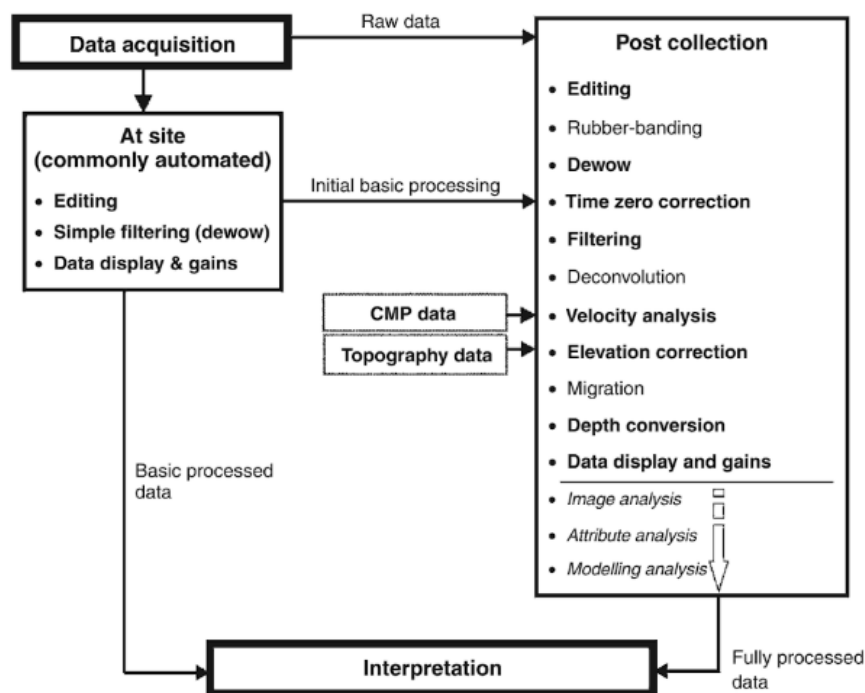


Figure 2-8 Typical processing sequence of 2D GPR data (Jol, 2009a)

(1). De-wow and adjust direct current shift (DC-shift)

De-wow correction is a temporal filtering method aimed at removing very low-frequency components from the data, which are often caused by inductive processes or equipment limitations. As shown in Figure 2-9, de-wow correction shifts the later part of a GPR trace to its correct position, ensuring the signal is centred and free from low-frequency distortions. In addition to de-wow correction, adjusting the DC shift is a critical step in preprocessing. The DC shift refers to the offset in the baseline of the GPR signal caused by inconsistencies in the transmitter-receiver system or environmental electromagnetic

interference. This offset manifests as a non-zero average value in the recorded signal amplitude. If left uncorrected, the DC shift can distort the radargram and obscure meaningful reflections. Adjusting the DC shift involves centring the signal around zero amplitude by subtracting the mean value of the recorded signal from each trace. This ensures that the baseline of the signal is accurately aligned, improving the clarity and interpretability of the radargram.

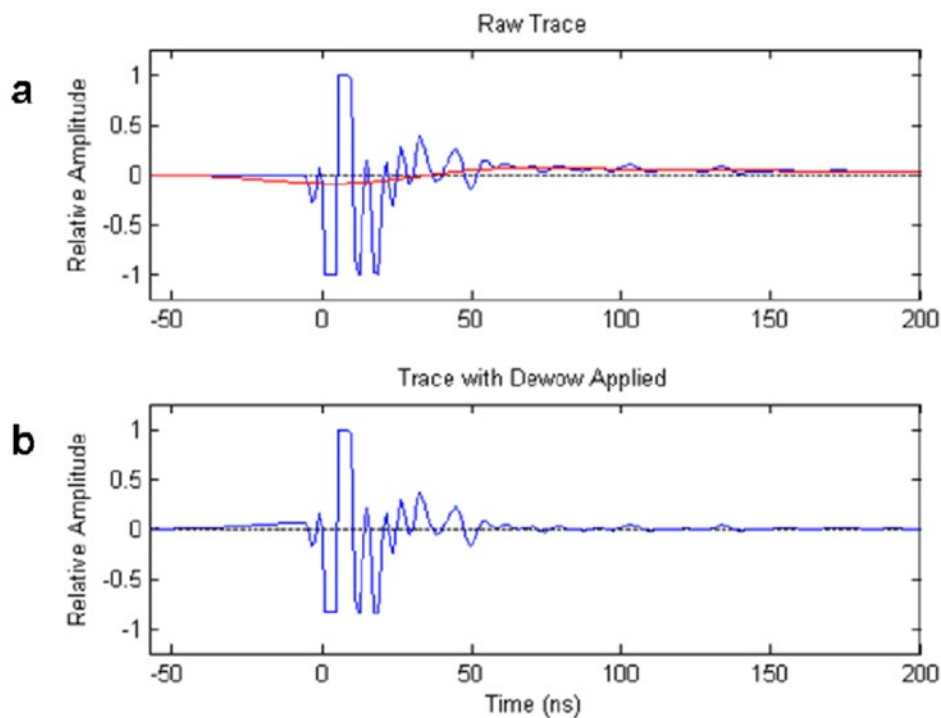


Figure 2-9 An example of the de-wow on a GPR trace (Rowell et al., 2010): (a) before and (b) after the de-wow

(3). Time-varying gain

As radar waves propagate into the ground, the signal strength diminishes due to attenuation, making reflections from deeper depths difficult to detect. Time-varying gain compensates for this attenuation by amplifying weaker signals. Techniques such as Automatic Gain Control (AGC) and user-defined gain functions are commonly used to enhance visibility. However, excessive gain may amplify noise, requiring careful adjustment

to balance signal fidelity and interpretability (Goodman et al., 2013).

(3). Filtering

Filtering can be processed before/after time-varying gains, as long as the gain's result is recognized. Frequency-domain filtering involves isolating specific frequency ranges to suppress undesired frequencies and enhance the target signal. This is commonly achieved using bandpass filters, which are one-dimensional filters applied in the frequency domain. Bandpass filtering retains a particular frequency range while removing the frequencies outside this range, thereby improving the interpretability of the radargram. Fast Fourier transforms (FFT) are often used to convert waveforms from the time domain to the frequency domain, enabling precise filtering of spectral components (Daniels, 2004; Malagodi et al., 1996). After filtering, the amplitudes of the retained frequencies can be adjusted by amplifying or diminishing them to optimize signal clarity (Goodman et al., 2013). High-pass and low-pass spatial filters are two well-known filters: the high-pass maintains dipping responses while removing flatter events, while the low-pass filter does the opposite.

Spatial-domain filtering targets noise and artefacts that are repetitive or systematic across the radargram. One common technique is background removal, which eliminates constant clutter noise caused by systematic responses in every trace. Background removal is performed by calculating the average pulse across the entire radargram and subtracting it from each trace, thereby removing repetitive (Bernabini et al., 1995; Malagodi et al., 1996).

(4). Time zero correction

The zero position ('time zero') changes depending on the ground material. The definition of the first measured arrival as time zero has been much debated. When an antenna is placed on the ground surface, the direct wave is altered in shape and shifts later in time by up to several tenths of a nanosecond, due to the dielectric loading of the ground material

in the near field of the antenna (Yelf, 2004). Time zero correction refers to moving the zero position of each trace to eliminate the lag from the electronic zero to ground zero. GPR users are advised to estimate the time zero in each application through their experiments gained from sites.

(5). Velocity analysis

As discussed in Section 2.2, the mathematical form of the hyperbolic reflection relates spatial position (x) to travel time (T). A handy aid to interpretation when estimating GPR wave velocity is to visually fit a model hyperbolic shape to the GPR data, using the two-way travel time as described by Equation [2-13].

(6). Elevation correction with topographic data

For surveys on undulating ground, GPR scans must be shifted based on elevation changes. Standard topographic adjustments require the velocity to be estimated from the radargrams, and then the scans can be shifted vertically to account for the change in topography. The tilt of the scan is estimated using a ground slope calculation based on the measurement of the local topography of the GPR scan. This slope is used to project the ray signal emitted from the tilted antenna.

(7). Migration

Most GPR antennas emit a broad beam of radio energy into the ground, which results in hyperbolic reflections being recorded from objects buried in the ground. Migration is a key GPR processing technique used to enhance resolution and create spatially accurate subsurface images. By utilizing a velocity model of the subsurface, migration aims to reconstruct GPR data in a more realistic spatial form. The process is particularly effective at collapsing diffraction hyperbolae into single point sources and repositioning dipping reflections to their correct locations, resulting in clearer and more accurate interpretations.

The Kirchhoff migration in the time domain and the Stolt migration in the frequency domain are two popular migration algorithms that yield similar results in most cases.

(8). Envelop to erase phase information

The enveloping of the recorded radar signal is known as a Hilbert Transform. It is a mathematical transformation process that uses a Fourier transform. The negative reflections are swapped 90 degrees and processed with an inverse Fourier transform. The recorded pulse is then ‘enveloped’ to produce a reflection that resembles the geometry of a circular object. In this process, the phase information is lost.

(9). Depth conversion

After estimating time zero and the GPR wave velocity, the depth of an object can be calculated by measuring the two-way travel time. Note also that the estimated depth after migration and envelope appears deeper than that of the hyperbolic reflections.

While standard 2D GPR post-processing techniques significantly enhance the visibility of radargrams by filtering noise, amplifying signals, and correcting distortions, their outputs, such as B-scans, still have limitations. B-scans provide cross-sectional views of the subsurface that are useful for interpreting localized features like hyperbolas, layers, and voids. However, these interpretations are confined to discrete survey lines, making it challenging to understand the spatial relationships between adjacent features or to form a comprehensive view of the subsurface. Without a broader spatial perspective, patterns across neighbouring regions may remain undetected, limiting the ability to interpret subsurface structures holistically. To address these limitations, advanced 3D processing and C-scan generation techniques have been developed, enabling the creation of detailed, spatially coherent subsurface images. These techniques bridge the gap between localized interpretations and a comprehensive understanding of the spatially related patterns, as explored in the next section.

2.3.3 GPR Imaging

As discussed in the previous section, 2D GPR post-processing techniques, while effective at improving the visibility of radargrams, are limited to localized interpretations of subsurface features. To address these limitations, C-scans—a type of horizontal cross-sectional representation derived from 3D GPR data volumes—have become a popular tool in the GPR field. C-scans offer several advantages over traditional 2D radargrams, making them particularly useful for large-scale subsurface investigations.

First, C-scans provide straightforward, intuitive visualizations, summarizing batches of radargrams into a single horizontal slice at a specific depth. This makes it easier to represent and interpret the spatial relationships between subsurface features, offering a broader perspective that is often missing in localized 2D interpretations. Second, C-scans are particularly effective for identifying spatially related patterns and visualizing the distribution of continuous objects, such as layers or buried structures, across large areas. This ability to highlight spatial relationships simplifies the task of narrowing down regions of interest for further analysis. Finally, C-scans are well-suited for large-scale interpretations, where they help reduce the complexity of massive datasets by providing overviews of subsurface conditions. This makes them an excellent tool for rough filtering, allowing practitioners to focus on specific regions that require closer inspection. Due to these advantages, C-scans have been widely used in the GPR field for a variety of applications. Numerous studies have demonstrated their utility in mapping subsurface features (Birken & Oristaglio, 2022; Metwaly, 2015; Ristic et al., 2009), monitoring changes in material properties (Kowalsky et al., 2005; Léger et al., 2014; Leucci et al., 2003), and inspecting structural conditions (Alani et al., 2013; Baker et al., 1997; Lai et al., 2012; Lai et al., 2013).

The popularity of C-scans has also driven efforts to standardize their generation

methods. Tess Luo et al. (Luo et al., 2019a) proposed a structured workflow for optimizing C-scan generation by defining key parameters such as profile spacing (PS), slice thickness (ST), and interpolation radius (SR), as illustrated in Figure 2-10. This workflow provides recommended parameter ranges depending on the availability of background information. If the target size is known, parameters are adjusted to size-related values. If the target size is unknown, parameters are set as ratios of the wavelength of the electromagnetic wave (λ). These parameters were designed to balance resolution with computational efficiency while ensuring consistency and accuracy in representing subsurface features.

In addition to parameter optimization, Tess Luo et al. introduced a standardized grayscale normalization scheme for visualizing C-scans, as shown in Figure 2-11 (Luo et al., 2023). This scheme enhances the contrast and clarity of C-scans, improving their interpretability for different application goals. Together, these advancements have significantly improved the reliability and consistency of C-scan representations, reinforcing their role as an indispensable tool in GPR imaging.

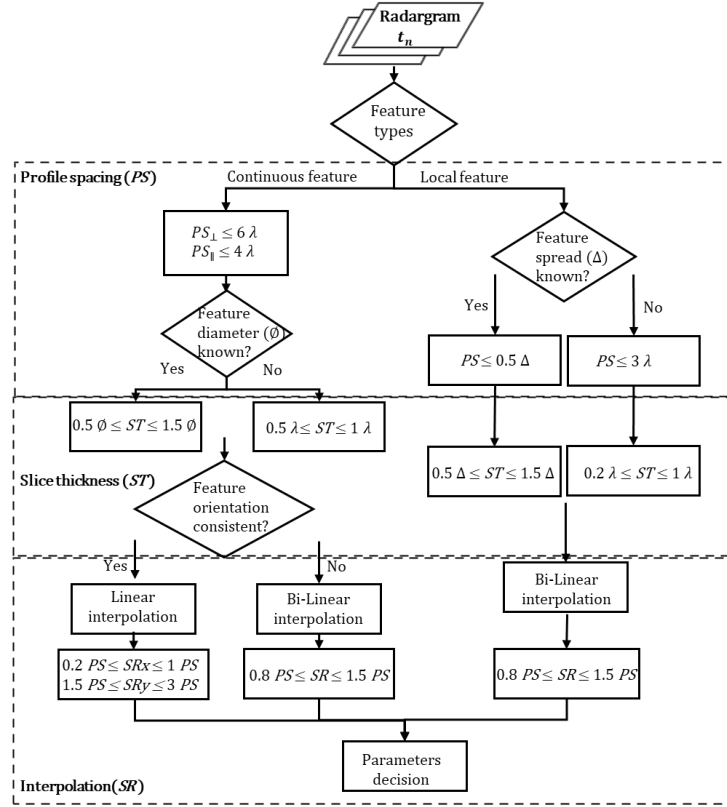


Figure 2-10 3D GPR imaging workflow. PS: profile spacing; ST: slice thickness; SR: search radius; λ : wavelength of the EM wave propagation underground (Luo et al., 2019a)

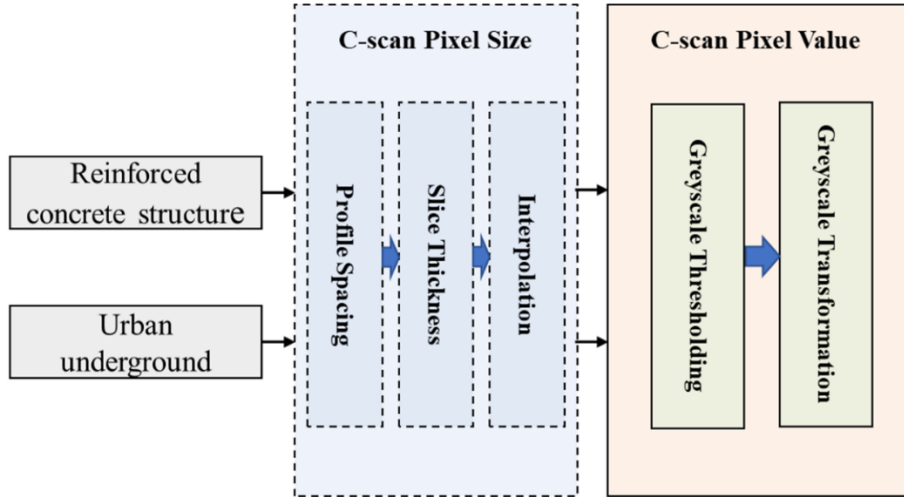


Figure 2-11 C-scan intensity normalization scheme that continuous the 3D GPR imaging workflow (Luo et al., 2023)

In this thesis, C-scans are selected as the primary GPR representation for developing a change detection method, based on several practical considerations. First, during MCGPR road surveys, it is often impossible to perfectly align survey lines across repeated measurements in

survey paths. C-scans, being horizontally interpolated representations, naturally accommodate such inconsistencies by summarizing data across the surveyed area. Second, the mature processing workflow for C-scan generation ensures that high-quality representations can be consistently produced, providing a stable foundation for detecting changes in subsurface conditions over time. These characteristics make C-scans the ideal representation for this study, enabling reliable and efficient change detection in large-scale GPR surveys.

2.4 Pattern Recognition and Change Detection

Relying on the unique electromagnetic properties of subsurface materials to identify anomalies, two primary strategies are employed in GPR-based hazard detection: pattern recognition from single measurement and change detection using temporal data from repeated surveys. This section explores these methods, focusing on their principles, advantages, and challenges.

2.4.1 One-shot Measurement and Pattern Recognition

One-shot detection involves analyzing a single GPR survey to locate potential underground hazards. This approach uses the distinct electromagnetic properties of materials, such as air-filled voids and water-saturated zones, which produce recognizable patterns in GPR data. Voids, often filled with air, exhibit distinct electromagnetic properties that make them easily identifiable in GPR surveys. The low permittivity and negligible conductivity of air cause GPR waves to propagate at higher velocities and with minimal attenuation. This sharp contrast in velocity and impedance at the boundaries between voids and surrounding materials produces strong reflections, which appear as high-amplitude signals in GPR data. Additionally, voids typically display limited frequency dispersion, allowing for clean wave propagation without significant distortion.

In contrast, regions with high water content, such as areas of water leakage, exhibit a

significantly different set of electromagnetic properties. Water's high permittivity dramatically reduces wave velocity in water-saturated zones, while its conductivity contributes to increased attenuation. This results in weaker GPR signals and limited penetration depth. Furthermore, water's dielectric relaxation effects cause frequency-dependent dispersion, with energy absorption and scattering increasing at higher frequencies.

For one-shot measurement, the underground hazards detection focuses on pattern recognition based on different representations of GPR measurements. When the GPR representations (A, B, and C-scans) have been prepared, further interpretation can be conducted. Various pieces of research have been developed that aim to detect the underground void or leakage by leveraging various features from different representations.

2.4.1.1 Pattern Recognition for Void Detection

The strong permittivity contrast between the soil surrounding the void and the air in the void leads to localized GPR reflections stronger than the background. Many researchers have investigated GPR responses of voids. At an early stage, Lai et al. (Lai, Chang, et al., 2017) concluded that voids would generate reverberation-like GPR responses. Casas et al. (1996) illustrated that if the void was much smaller than the GPR wavelength, diffracting hyperbolas occur in GPR B-scans. On the contrary, if the voids' diameters are much larger than the wavelength, irregular signals with chaotic reflections and a decrease in the received frequency are obtained. Kofman et al. (Kofman et al., 2006) simulated GPR responses with gprMax (Giannopoulos, 2005; Warren et al., 2016, 2019) and indicated that oscillations happened only when the void size was significantly larger than the GPR wavelength.

These studies demonstrate two major patterns for void detection: reverberation and hyperbolic reflections. In detailed forward modelling, the response of air voids represents hyperbola, cross, bowl, and plain reverberation, considering the four factors: physical spread,

void depth, antenna frequency, and wave velocity (Luo et al., 2020).

In the time domain, some researchers focus on reverberation recognition and feature recognition by template matching. In the frequency domain, the phase gradient of a phase-frequency spectrum of signals reflected from underground layers and buried objects was related to their electrical properties, showing a way to detect the underground void (Sugak & Sugak, 2010). The instantaneous phase is another indicator of underground material changes. By detecting the in-phase and out-phase signals, relative permittivity relationships can be retrieved to identify voids (Park et al., 2018). Also, GPR data can be applied as additional data to support other resource data, including time domain reflectometry (Hong & Lee, 2018) and electrical resistivity tomography (Orlando, 2013), for void detection tasks. Overall, these feature-based methods aim to bridge the features extracted from GPR responses with underground objects based on expert knowledge.

Given the promised performance of pattern recognition in other fields, learning-based methods are becoming more and more popular in void detection tasks. Techniques such as UcNet, which combines CNNs with phase analysis of super-resolution GPR images, reduce misclassification of non-cavity features (Kang et al., 2019b). Similarly, few-shot learning approaches, such as RelationNet, offer solutions for cavity morphology classification with minimal data, addressing overfitting issues in datasets with limited samples (Hou et al., 2022). In other approaches, 3D convolutional neural networks applied to voxel data achieve high classification accuracy, effectively distinguishing features like pipes, cavities, and manholes, with minimal false alarms (Kang et al., 2020; Khudoyarov et al., 2020). Incorporating multi-view and multidimensional data has also advanced GPR interpretation. Triplanar methods using B-, C-, and D-scan images outperform traditional B-scan-based techniques, providing richer spatial information and improving object classification (Kim et al., 2019). Similarly, MV-GPRNet leverages multi-view data fusion from C- and Top-scan features to enhance defect

detection in noisy GPR data, demonstrating superior performance for various subsurface defects (Li et al., 2022). A complementary method uses simulated 3D reflection patterns in 3D CNNs to improve void detection accuracy and visualization, confirming the advantages of integrating 3D sensing (Yamaguchi et al., 2022b). The use of data augmentation and generative methods has further enabled detection in data-scarce scenarios. For example, B-scan image augmentation using SinGAN paired with YOLOv5 facilitates subsurface void detection even with limited real-world samples (Chen et al., 2021). Similarly, clutter reduction techniques such as RCE-GAN improve void detection in challenging environments, like tunnel linings obstructed by rebar reflections, demonstrating robustness across synthetic and real-world datasets (Wang et al., 2022). Finally, applications in infrastructure safety, such as airport runway inspections, highlight the practical utility of these advancements. Methods like GPR-RCNN and ResNet-YOLOv2 leverage deep learning to automate defect detection and reduce noise, achieving high accuracy and reliability in identifying voids, cracks, and other structural anomalies (Li et al., 2021; Zhang et al., 2022). These innovations collectively enhance the reliability and scalability of GPR-based inspections, offering promising solutions for urban safety and infrastructure maintenance.

2.4.1.2 Pattern Recognition for Leakage Detection

Like the special patterns that indicate the air void, GPR antennas are a useful tool for detecting water sewages for the following three reasons: Firstly, amongst non-metallic materials, water is the single most important factor altering the GPR signal characteristics because it plays a major role on both real and imaginary part of material permittivity, resulting in changes of wave velocity and spectral energy absorption (Huisman et al., 2003; Klotzsche et al., 2018; Klysz & Balayssac, 2007), respectively. Secondly, by GPR, the internal condition of the subsurface in multiple dimensions can be unfolded efficiently and in very high resolution in millimeters. It is because, unlike acoustic methods such as acoustic emission or leak noise

correlation methods, or electromagnetic induction methods (or pipe cable locating), GPR imaging does not require physical contact with any objects connecting the pipe, like valves. Compared with the seismic and low-frequency vibro-acoustics (Muggleton et al., 2011; Muggleton & Papandreou, 2014; Wang & Yin, 2022), GPR does not require physical excitation of the ground, and its centimetre resolution allows high-resolution imaging of the sub-surface in urban, congested underground. Lastly, GPR's wide frequency range matches different physical sizes/scales of structure thicknesses.

The detection of localized water leakages using GPR can present complex patterns. Several pieces of research highlight the propagation velocity of signals derived from water content changes based on radargrams (Cataldo et al., 2014; Cheung & Lai, 2019; Crocco et al., 2009; Lau et al., 2021). However, its application is somewhat limited to precise velocity estimations. For instance, finding high-quality hyperbolas, as many hyperbolic patterns can be distorted in a complex underground. Liu et.al model the various reflections introduced by water leakage, resulting in oscillating patterns. Leakage point location methods based on wavelet entropy were proposed (Hu et al., 2020; Liu et al., 2022, 2023). One of the challenges with the entropy-based method is that certain underground anomalies can also increase entropy. Additionally, using wavelet entropy to localize leakage points is effective only when the soil remains in an unsaturated state.

Besides the radargram-based leakage detection method, Lai et al. proposed a leakage detection method based on wave attenuation, revealed by energy absorption in amplitude slices (Lai et al., 2016a; Lai et al., 2017). For better leakage pattern recognition, Ocaña-Levario et al. proposed a variation filter to highlight leakage signals, and Zhu et al. developed a framework using hydromechanical modelling to predict electromagnetic responses of water leakage (Ocaña-Levario et al., 2018; Zhu et al., 2024).

In addition to time-domain analysis, water content also manifests distinct characteristics in the frequency domain. Water that is trapped and bound within dielectric materials—such as soil, asphalt, and concrete—displays an effect akin to free water but at a significantly lower relaxation frequency (De Loor, 1983), such as 500 MHz in general (Jol, 2009b), or 1000 MHz in asphalt (Shang & Umana, 1999a) and 1200 MHz in concrete (Lai et al., 2011; Xiao et al., 2017). The relaxation frequency of a material is determined by the volumetric fractions and forms of free water present in the material, and it can vary based on the state of the water—whether fully bound, partially bound, or free—and the volumetric content of pore water across a wide range of porosities in soil, concrete, and asphalt. In summary, the permittivity (dielectric constant) of the subsurface mixture increases with water content. Accompanied by an increase in material conductivity, the energy loss also escalates. Specifically, at frequencies around the relaxation frequency, the absorption of energy by water becomes the principal loss mechanism in soil (Doolittle & Butnor, 2009).

However, this approach faces challenges. Underground hazards exhibit varying patterns at different stages of development, requiring detection methods to be carefully designed to account for these variations. Additionally, similar imaging patterns in GPR data can correspond to different subsurface conditions, complicating interpretation and reducing the performance of detection algorithms.

2.4.2 Temporal Measurements and Change Detection

Conducting a GPR survey is comparable to performing a CT scan, as both techniques capture detailed internal images to assess changes or abnormalities. Each GPR survey provides a timestamped snapshot of subsurface conditions, enabling a dynamic understanding of underground changes over time. This approach transitions from static, one-shot detection methods to a temporal framework, where multiple measurements are compared to monitor the

progression of underground anomalies.

To illustrate the importance of temporal data, a parallel can be drawn to medical imaging. Figure 2-12 shows the five stages of tumour development, demonstrating how analyzing changes over time provides deeper insights into a dynamic process. While a single medical image can identify abnormalities based on specific patterns, such as irregular shapes or tissue densities, tracking a series of images over time enables the detection of growth or shrinkage, offering a more comprehensive understanding of the condition. Similarly, in GPR-based subsurface monitoring, temporal measurements allow for dynamic analysis, helping to detect progressive changes such as void expansion or water infiltration. This capability is crucial for early hazard detection and prevention.

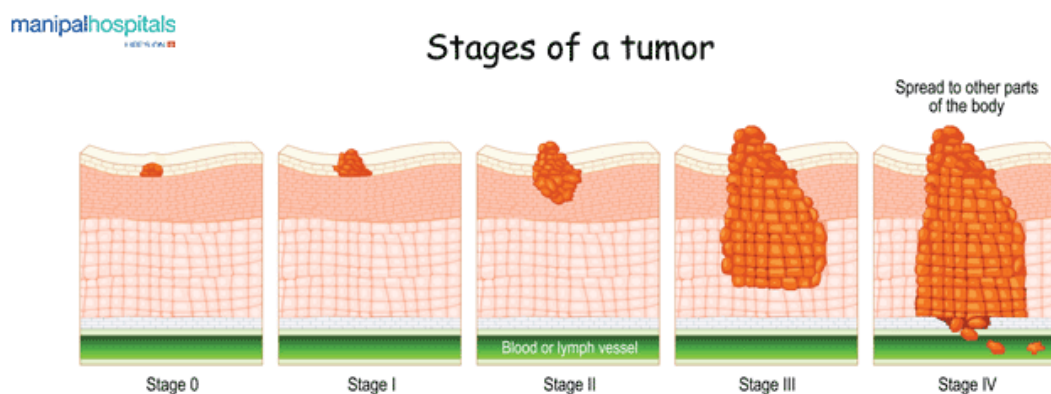


Figure 2-12 Illustration of the development of tumors in different stages (Joseph, 2023)

Underground hazards, such as road subsidence or water leakage, typically evolve in stages, making early detection vital to avoid accidents. Road subsidence, as depicted in Figure 2-13, often begins with delamination at material interfaces caused by soil erosion or water infiltration. Over time, the delamination progresses into a cavity, eventually forming a void that compromises pavement stability. Monitoring these changes requires long-term observations and systematic analysis.

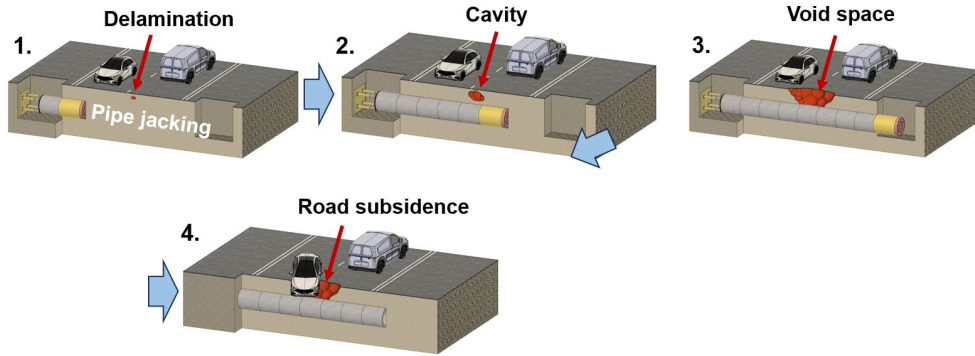


Figure 2-13 Illustrations of road subsidence development from delamination to subsidence

Long-term monitoring and change detection involve comparing multiple GPR observations over time to identify subsurface changes. This approach provides a dynamic perspective by classifying regions as either unchanged (stable) or changed (potentially hazardous). By tracking temporal changes, this method highlights abnormal regions while filtering out noise that might appear significant in a single observation. It is particularly effective for detecting unstable areas that signal the development of underground hazards. Figure 2-14 illustrates the difference between static pattern recognition and dynamic pattern recognition for underground hazard detection. In subplot (a), the orange circle represents the existing hazards at T_n and the green circle represents the detected hazards by the static pattern recognition-based method. Region 1, region 2, and region 3 represent true negatives, true positives, and false positives. In subplot (b), existing hazards represented as the orange circle are divided into two parts: the existing hazards have changes, and the existing hazards have no changes between T_0 and T_n . Dynamic detection focuses only on hazards that exhibit changes, which are represented by true positives in the overlapping area (Region 3). However, hazards that remain static (unchanged) are undetectable by dynamic recognition (Region 1), as they fall outside the scope of this method. Additionally, hazards with changes that are missed by the detection process are categorized as false negatives (Region 2), while areas incorrectly identified as hazardous changes are considered false positives (Region 4). Dynamic pattern

recognition improves detection accuracy by identifying evolving hazards, which is a new perspective for detecting underground hazards.

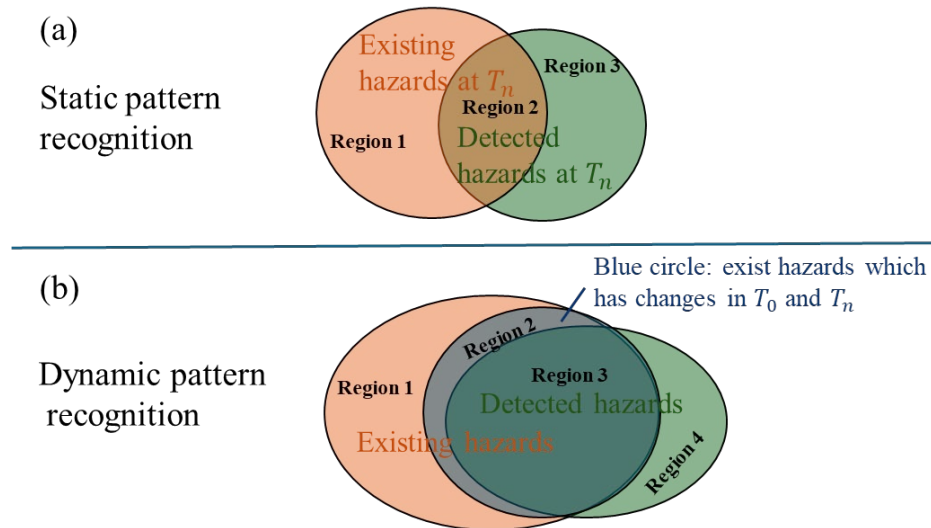


Figure 2-14 Static pattern recognition vs. dynamic pattern recognition for underground hazard detection

The transition from one-shot measurements to time-series observations is a natural progression in the development of monitoring technologies. Remote sensing, for example, evolved from static single-date mapping to dynamic multi-date image analysis and systematic change detection. Early applications focused on characterizing static landscapes, addressing questions such as “Where are the cities, and how much area do they occupy?” (Hansen & Loveland, 2012; Zhu & Woodcock, 2014). This single-date approach, often referred to as “mapping,” assumed that a snapshot of the landscape at a specific time could yield valuable insights (Singh, 1989). As advancements in Earth observation technologies increased the temporal resolution of data, and computing power and storage capacities expanded, the field shifted toward analyzing dynamic processes. Researchers began asking questions like “How have cities grown over the past decade?” (Hermosilla et al., 2015). This change detection approach focused on identifying differences between two points in time, leveraging multi-temporal imagery to highlight areas of transformation. Examples include monitoring deforestation, tracking urban expansion, and assessing vegetation health under stress, with the

nature of "change" tailored to specific study objectives.

In the context of GPR, early studies on change detection employed similar multi-temporal strategies. Initial efforts addressed clutter-limited environments, using a constellation matching approach to identify buried object changes across successive scans, effectively improving detection–false alarm rate performance even in misaligned paths (Paglieroni et al., 2015). Laboratory experiments followed, focusing on signal perturbation patterns caused by water leaks, which provided diagnostic signatures to identify leak points and wetting fronts through dielectric contrast and reflection mechanisms (Lai et al., 2016b).

Long-term monitoring later gained prominence, exemplified by a ten-year study on reinforcement corrosion in concrete. This work used image registration and intensity normalization to track gradual changes over time, demonstrating the potential for periodic health assessments of infrastructure (Hong et al., 2017). Time-lapse GPR data analysis advanced further with workflows combining image registration and K-means clustering, enabling semi-automatic extraction of structural changes while minimizing noise. These methods proved effective in urban subsurface studies, as validated through excavation and pipe replacement scenarios (Luo & Lai, 2020a). The latest developments in GPR change detection include GPR_CDNet, a deep learning framework designed for shape change detection in B-scans under dynamic environmental conditions. By employing a pseudo-Siamese network, multiscale convolutions, and feature transformation modules, this method improves the accuracy of reconstructing defect shapes despite variations in permittivity and shielding effects (Xu et al., 2024).

Despite its advantages, long-term monitoring has its own set of challenges. Consistency between observations is critical, requiring accurate registration of geolocation and signal properties across multiple surveys. Additionally, changes detected in the subsurface must be

carefully analyzed to distinguish those caused by underground hazards from those resulting from environmental factors, such as variations in moisture content or temperature. Addressing these challenges is essential to enhancing the reliability of this method for detecting and monitoring underground anomalies. While traditional GPR applications primarily relied on single-point surveys to map subsurface features at a given moment, advancements in data acquisition and processing now enable systematic time-series analysis. This shift mirrors broader trends in remote sensing and expands GPR's utility for monitoring subsurface dynamics over time. By integrating temporal dimensions, researchers can uncover trends, diagnose the progression of voids or leaks, and develop predictive models for structural degradation. These innovations underscore the transformative potential of GPR in long-term monitoring and change detection.

2.5 Summary of the Current Research and Research Gaps

This chapter provided a comprehensive review of the theoretical foundation and practical applications of GPR in underground hazard detection. The discussion began with A review of subsurface hazard mechanisms, detailing the degradation and leakage of underground pipes (Section 2.1.1) and the formation of voids and road collapses (Section 2.1.2). These insights established the foundation for exploring detection strategies. Following mechanisms, the fundamentals of GPR theory, including the design and operation of GPR antennas, EM wave properties, and their interactions with subsurface materials (Sections 2.2.1–2.2.3). The section then examined the standard approaches to processing and visualizing GPR data (Section 2.3). This included an overview of different data representations, such as raw radargrams, time-slices, and volumetric imaging (Section 2.3.1), as well as widely adopted post-processing techniques like filtering, migration, and amplitude normalization (Section 2.3.2). Additionally, advanced imaging methods and colourization techniques were introduced, showcasing their

potential to enhance the interpretability of GPR data (Section 2.3.3). Two main approaches were highlighted: pattern recognition based on single observations (one-shot measurement, Section 2.4.1) and change detection leveraging temporal data from repeated measurements (Section 2.4.2). Through this review, three critical research gaps were identified:

- (1). Lack of realistic leakage signature databases hinders the development of reliable leak detection methods, as most studies rely on oversimplified simulations.
- (2). Limited MCGPR-tailored post-processing methods fail to address spatial misalignments and inter-channel signal variations, reducing multi-survey comparability.
- (3). Subsurface hazard detection from a change detection perspective remains underexplored, with most studies prioritizing static anomaly identification over dynamic monitoring of evolving voids/leaks.

Chapter 3. Research Resources

During my doctoral studies, the Underground Utility Survey Laboratory (Utilities Lab) at The Hong Kong Polytechnic University (PolyU) and the Q-Leak Center collaborated with the Land Survey and Geo-Informatics Department (LSGI) and WSD provided invaluable support for my research. This support included access to a wide range of advanced experimental equipment, purpose-built research sites designed to emulate real-world conditions, and collaboration with other departments to collect data for addressing various urban hazard scenarios. These resources were instrumental in conducting comprehensive underground inspections and achieving the research objectives.

3.1 GPR Equipment

The GPR equipment used in this research consisted of both single-channel and MCGPR systems, selected for their ability to meet a wide range of survey requirements. The characteristics and configurations of these systems are summarized in Table 3-1. Two single-channel systems, the IDS Hi-Mode and DS2000, were used for detailed subsurface imaging. Both systems are equipped with dual-frequency antennas aligned along the heading direction, allowing simultaneous recording of underground reflections. Operating within a frequency range of 200 MHz to 700 MHz, these systems provide GPR measurements from various frequencies. In addition to single-channel systems, two MCGPR systems, StreamUP and Raptor45, were integral to this research, particularly for large-scale surveys (see Figure 3-1).

Table 3-1 Summary of used GPR antennas

Name	Manufacturer	Antennas type	Working frequency	Number of channels	Channel spacing	Scan width
IDS Hi-mode	IDS GeoRadar	Impulse	200MHz/600 MHz	2	/	/
DS 2000	Leica	Impulse	250 MHz/700	2	/	/

			MHz			
StreamUP	IDS GeoRadar	Impulse	200MHz/600MHz	19 +10	8.3cm and 16.7cm	1.58m
Raptor45	Impulse Radar	Impulse	450MHz	18	8.4cm	1.4m

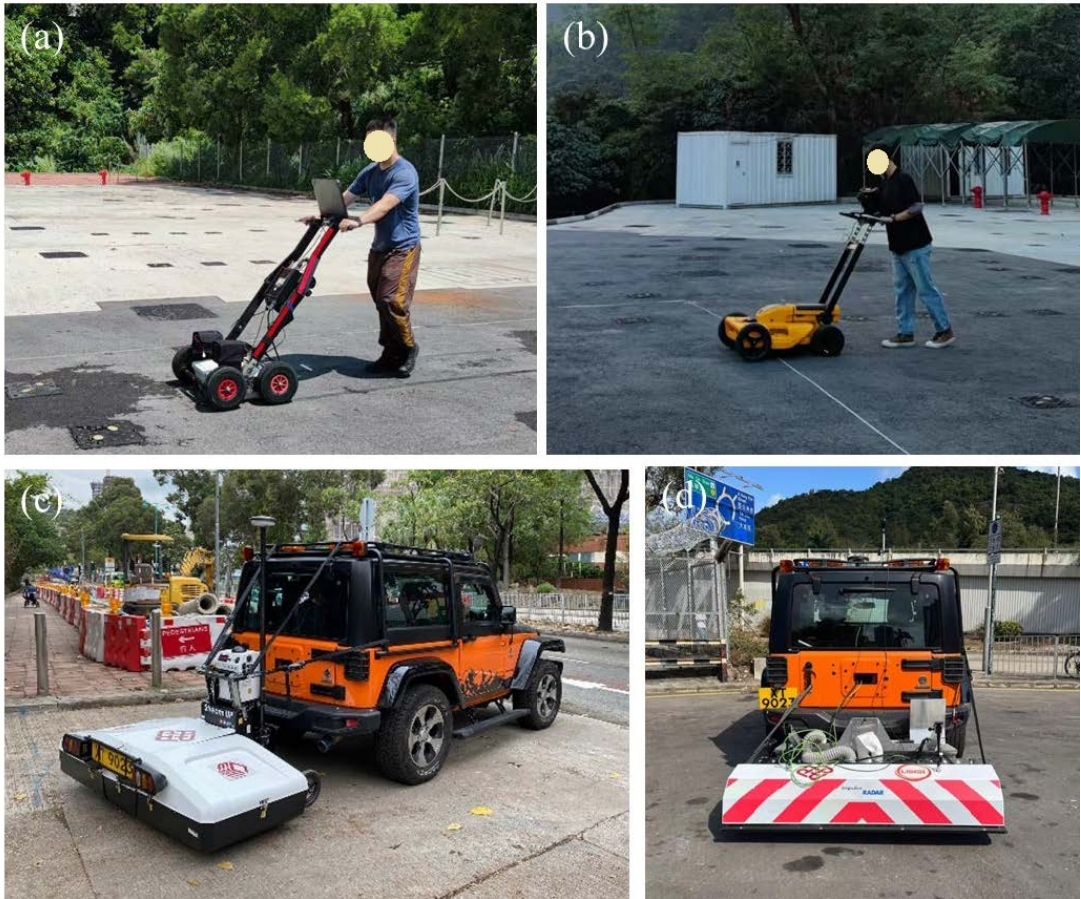


Figure 3-1 Photos of GPR systems (a) IDS-Hi mode, (b) DS2000, (c) StreamUP, and (d) Raptor45

MCGPR systems differ significantly based on their design parameters, especially the configuration of antenna channels. The StreamUP system features two arrays of antennas operating at 200 MHz and 600 MHz, with each frequency having distinct polarization modes due to the different orientations of the dipoles, as shown in Figure 3-2. The Raptor45 system operates with a single working frequency of 450 MHz, offering a balanced trade-off between depth penetration and resolution. Its antenna configuration consists of one row of transmitters and one row of receivers, providing a single polarization mode, as seen in Figure 3-2. Other

parameters, such as channel numbers, channel spacing, and scan width, describe the data acquisition abilities.

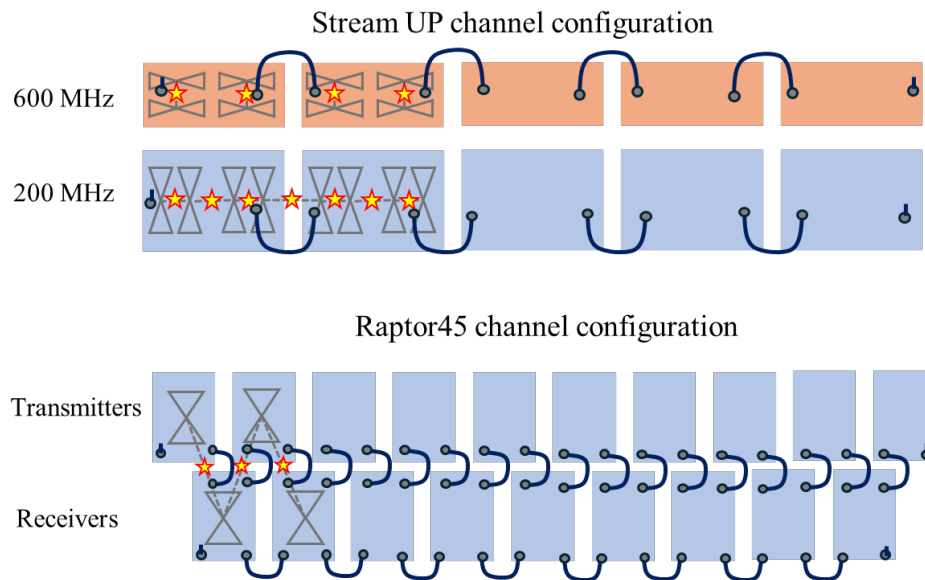


Figure 3-2 Illustration of the channel configurations of StreamUp and Raptor45

To ensure accurate geolocation of MCGPR data, a GNSS-RTK receiver (Leica GG04) was integrated into the survey setup. This receiver provided high-precision positioning in open-sky environments, ensuring reliable spatial positioning of GPR data during field surveys. The diversity in frequency ranges, antenna configurations, and polarization modes of the GPR systems used in this research provided the versatility necessary for comprehensive underground inspections. These systems supported both small-scale, high-resolution surveys and large-scale, city-wide investigations, ensuring the research objectives were met with precision and efficiency.

3.2 Research Experiment Sites

The success of this research relied heavily on access to well-equipped experimental facilities and carefully selected field sites. Laboratory experiments were conducted at the Utilities Lab and the Q-Leak Center, both of which provided controlled environments and

specialized resources. Field surveys were conducted at real-world road sites, selected based on project permissions, research objectives, and confirmed or suspected occurrences of urban underground hazards.

The Utilities Lab at PolyU houses a specialized testing platform designed to replicate scaled-down subsurface urban environments. The platform measures 5 meters in length, 3 meters in width, and 1 meter in depth, as outlined in yellow dashed lines in Figure 3-3. It contains plant soil and various buried utilities, including freshwater pipes, saltwater pipes, gas pipes, and power cables, creating a controlled environment for underground investigations. The platform includes removable glass fibre-reinforced polymeric panels marked with orthogonal red lines for accurate mapping and experimental adjustments. Photos taken before soil filling, shown as the right photo of Figure 3-3, are essential in defining the ground-truth condition of the testbed layout and utility placement. The platform supported experiments related to Objective 3 by enabling precise testing of methodologies for detecting underground voids.

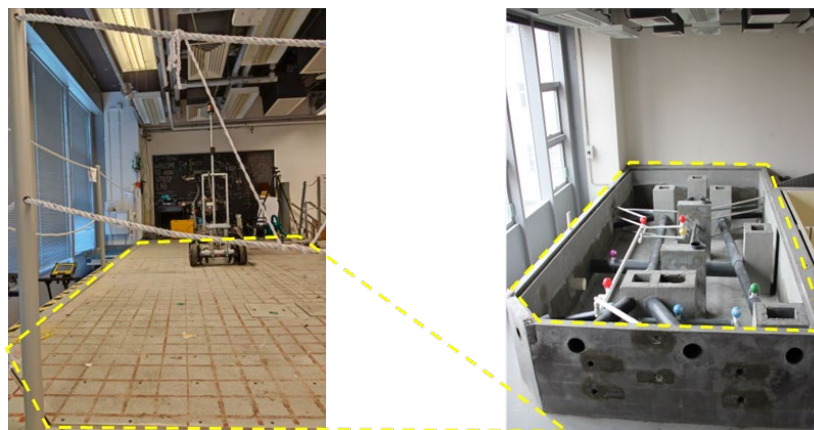


Figure 3-3 Photos of the experiment site at Utilities Lab at PolyU

Q-Leak Center, located in Tsing Yi, Hong Kong SAR, is an advanced training and research facility for underground water mains leak detection, as shown in Figure 3-4. Spanning approximately 2,400 square meters, which significantly upsets the Utilities lab to the full scale, it incorporates a sophisticated emulated water supply network with varied pipe materials, sizes, and failure modes.

To simulate leakage scenarios effectively, five dedicated leakage cells were constructed beneath different types of pavements. Each leakage cell contains eight pipelines with varying materials, diameters, and defect types (e.g., cracks, pinholes, joint failures). The design of these pipelines, combined with a network of connected control valves, allows researchers to precisely control water flow and pressure, enabling the simulation of a wide range of leakage scenarios. These scenarios replicate common failure modes observed in real-world water mains, supporting experiments that investigate the behavior of leaks under different conditions. The facility is also equipped with a comprehensive array of monitoring equipment, such as control valves, flow meters, and pressure sensors, to replicate the operational parameters of real water networks. These tools provide precise control over experimental conditions, ensuring consistency and repeatability across experiments. To detect and analyze the emulated leaks, various advanced technologies can be applied, including ground microphones, noise loggers, acoustic leakage correlators, and GPR. This multi-instrument approach allows researchers to establish cross-technology observations of leakage phenomena, providing insights into how different detection methods respond to various leakage conditions.

This facility played a crucial role in research related to Objective 1, where various leakage scenarios were emulated to develop a comprehensive leakage signature database. More details of the leakage signature database establishment will be introduced in Chapter 4. Additionally, controlled experiments to harmonize MCGPR outputs were conducted at the centre, contributing to Objective 2.



Figure 3-4 Photos of the Q-Leak center and leak detection ground

Complementing the controlled environments of the lab and Q-Leak Center, real-world road sites provided more complex and challenging conditions for data collection. These sites, designed by PolyU LSGI, included locations with confirmed cases of road collapses and water main leaks. Data collected from these real-world environments validated the methodologies developed in controlled settings, ensuring their applicability to urban-scale surveys.

3.3 Software

This research relied on a combination of proprietary and open-source software for data acquisition, processing, and analysis. In this data acquisition stage, Different GPR systems utilized their respective proprietary software for data acquisition, including Onevision for IDS hi-mode, Unext for Leica DS2000, Umap for StreamUP system, and Talon for Raptor45. The

majority of data processing and algorithm development was conducted using Python scripts developed during this research. These scripts incorporated classic GPR processing functions, drawing inspiration from tools such as RGPR (Huber & Hans, 2018), GPR-SLICE, Geolitix, and Reflexw. For generating time slices, the logic of GPR-SLICE was particularly influential, providing foundational ideas for processing workflows. In the core part of this thesis, change detection methods were developed using open-source Python libraries, including OpenCV, scikit-image (Walt et al., 2014), GDAL (GDAL/OGR contributors, 2025), and others. Also, I would like to acknowledge the powerful simulation tool, gprMax (Giannopoulos, 2005; Warren et al., 2016, 2019). Special thanks are extended to the open-source community for providing these tools, which were instrumental in achieving the research objectives.

Chapter 4. Leakage Signature Database and Leakage Characterization

This chapter introduces the establishment of a comprehensive leakage signature database, created by simulating 78 leakage scenarios at a carefully designed experimental site. Given the nature of leakage investigations, GPR surveys are typically conducted after the leakage has occurred and visible signs are present. This reactive approach makes it challenging to observe the complete development process of water leakage in real-world conditions. The signature database addresses this limitation by providing an extensive dataset that bridges the gap between simulated and real-world leakage scenarios. It serves as a valuable resource for developing detection methods targeted at the dynamic patterns of pipe leakages. The database comprises three sub-datasets. These datasets are tailored for different GPR data representations, enabling robust pattern recognition in temporal GPR measurements. By offering a structured and diverse collection of leakage scenarios, the database supports the development and validation of advanced detection methods that can adapt to real-world complexities.

To enhance leakage characterization, IFS was introduced as a novel method for analyzing water leakage patterns. By extending the analysis into the time-frequency domain, IFS uncovers two key mechanisms influencing leakage signatures. Reflections at the dry-wet interface and absorption caused by water distributed in the soil. Validated using the signature leakage database, IFS demonstrates its effectiveness in detecting water leakages in buried pipes, addressing the limitations of traditional amplitude slice analysis. This approach provides a more reliable and nuanced characterization of leakage patterns, making it a significant advancement in GPR-based leakage detection.

4.1 Leakage Signature Database

This section details the design and implementation of the leakage signature database. It includes a description of the experimental site, the methodology used for simulating leakage scenarios, and the structure of the three sub-datasets, each tailored to support different research objectives. By capturing diverse leakage patterns under controlled conditions, the database serves as a foundation for advancing GPR-based methods for leakage detection and characterization.

4.1.1 Experiment Sites and Designed Leakage Experiments

The leakage experiments were conducted at the Q-Leak Center, a controlled facility designed to replicate real-world water supply network conditions. These experiments were carefully designed to account for various factors essential to building a comprehensive and representative pipe leakage signature database. The inclusion of diverse pavement materials, pipe configurations, and leakage defects was intended to ensure that the database captures the wide range of conditions encountered in real-world leakage scenarios, making it suitable for validating and advancing the proposed methods. The detection ground was divided into five regions, each featuring distinct pavement types: red brick, green blocks, plain concrete, reinforced concrete, and asphalt (Figure 4-1).



Figure 4-1 Overview of the leakage detection ground showing five different pavement types and corresponding photos

Beneath these surfaces, a comprehensive network of water and drainage pipes was constructed, mimicking real-world utility layouts. These pipes were interconnected with joints and valves to facilitate controlled leakage simulations, creating an ideal environment for studying leakage phenomena.

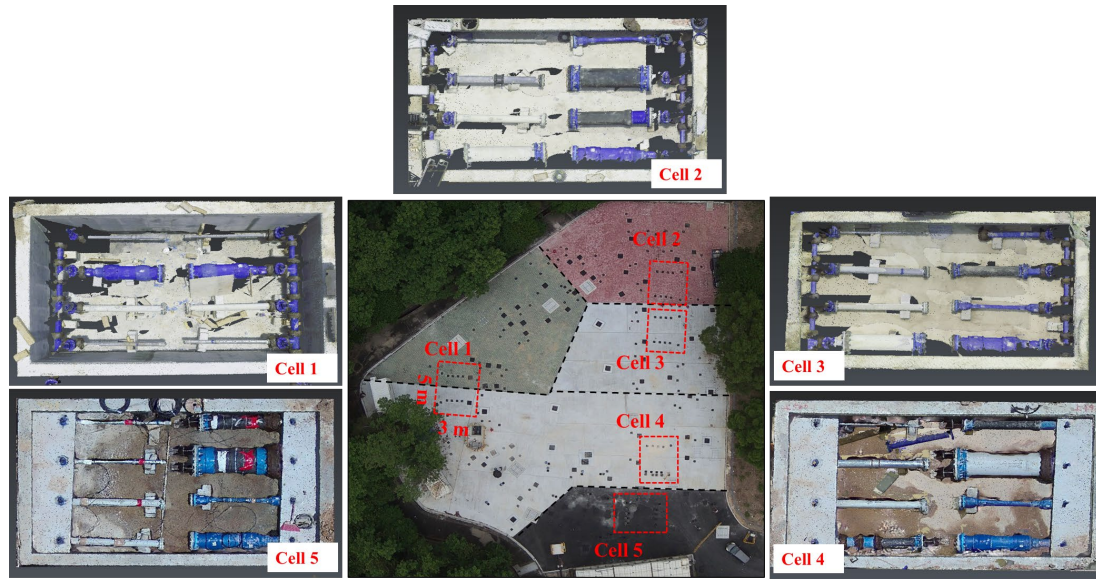


Figure 4-2 Locations of five leakage experiment cells and detailed views of the leakage pipes

To simulate leakage events under various conditions, five specialized leakage cells were constructed within the experimental area, marked by dashed red rectangles in Figure 4-2. Each cell was a closed concrete tank measuring 500 cm in length, 300 cm in width, and 150 cm in depth. These tanks housed eight types of commonly used pipes varying in material, diameter, and depth of burial, covered by different pavement types. Each pipe was equipped with gate valves to precisely control water flow and induce leakage. This configuration allowed for controlled experiments investigating leakage across combinations of pipe materials, defect types, and surface conditions. The Q-Leak centre supported up to 40 leakage scenarios, encompassing a wide range of realistic conditions encountered in actual water supply networks.

Table 4-1 Summary of the leakage points and their attributes

Cell	Overlaid material	Pipe material	Pipe size	Defect type	Size defect
1	Green blocks	GI	50	FJ	
		PE	OD250	WFJ	
		MS	100	LH	5mm in diameter
		GI	50	LC	2mm

					wide 150mm long
	Reinforced Concrete	GI	50	CC	2mm wide
		PE	OD250	WFJ	
		MS	100	WFJ	
		GI	50	SC	2mm wide
2	Red brick	GI	25	ITJ	
		GI	100	FJ	
		MS	100	WFJ	
		MS	300	LH	5mm in diameter
		PE	OD63	CC	5mm in diameter
		DI	300	LC	2mm wide 150mm long
		DI	200	LH	2mm wide
		PE	OD250	WFJ	
3	Plain concrete	GI	25	LH	5mm in diameter
		GI	100	LC	2mm wide 150mm long
		MS	100	WFJ	
		MS	300	FJ	
		DI	100	CC	2mm

		DI	150	WFJ	wide
		PE	OD63	EFJ	
		PE	OD250	LH	5mm in diameter
4	Reinforced Concrete	GI	25	CC	2mm wide
		GI	100	SC	2mm wide
		MS	100	WFJ	
		DI	150	FJ	
		DI	100	WPIJ	
		MS	400	WMJ	
		PE	OD63	LC	2mm wide 150mm long
		PE	OD250	CC	2mm wide
5	Asphalt	GI	50	ITJ	
		GI	50	LH	5mm in diameter
		MS	100	WFJ	
		MS	100	LH	5mm in diameter
		DI	200	WFJ	
		DI	400	WPIJ	
		PE	OD63	LC	2mm wide 150mm long
		PE	OD250	EFJ	

Table 4-1 provides a detailed summary of the leakage points, categorized by Cell and pavement type. Attributes such as overlaid material, pipe material, pipe size, defect type, and defect dimensions are listed. For instance, Cell 1 included green block pavement with galvanized iron (GI), polyethylene (PE), and mild steel (MS) pipes, showcasing defect types such as longitudinal cracks (LC) and weak flange joints (WFJ).

Table 4-2 *Description of defect types*

Defect Types		
Hole/crack	LH -	Leak hole
	LC -	Longitudinal crack
	CC -	Circumferential crack
	SC -	Shear crack in GI pipe threaded joint
Joint	WFJ -	Weak flange joint
	WPIJ -	Weak push-in joint's rubber gasket failure
	EFJ -	In-complete electro-fusion joint to PE pipe
	WMJ -	Weak mechanical joint for DI pipe
	FJ	Flexible joint (coupling for both MS and GI, flange adaptor for MS pipe)
	ITJ -	Incomplete threaded joint

The defect types introduced in the experiments are described in Table 4-2, which categorizes them into two main groups: cracks or holes and joint failures. Crack-related defects in pipes included leak holes (LH), longitudinal cracks (LC), circumferential cracks (CC), and shear cracks (SC) in GI pipe threaded joints. Joint-related defects included weak flange joints (WFJ), weak push-in joints (WPIJ), incomplete electro-fusion joints (EFJ), weak mechanical joints (WMJ), flexible joints (FJ), and incomplete threaded joints (ITJ). These defect types reflect common issues in real-world water supply systems, enabling the study of their GPR signatures under controlled conditions.

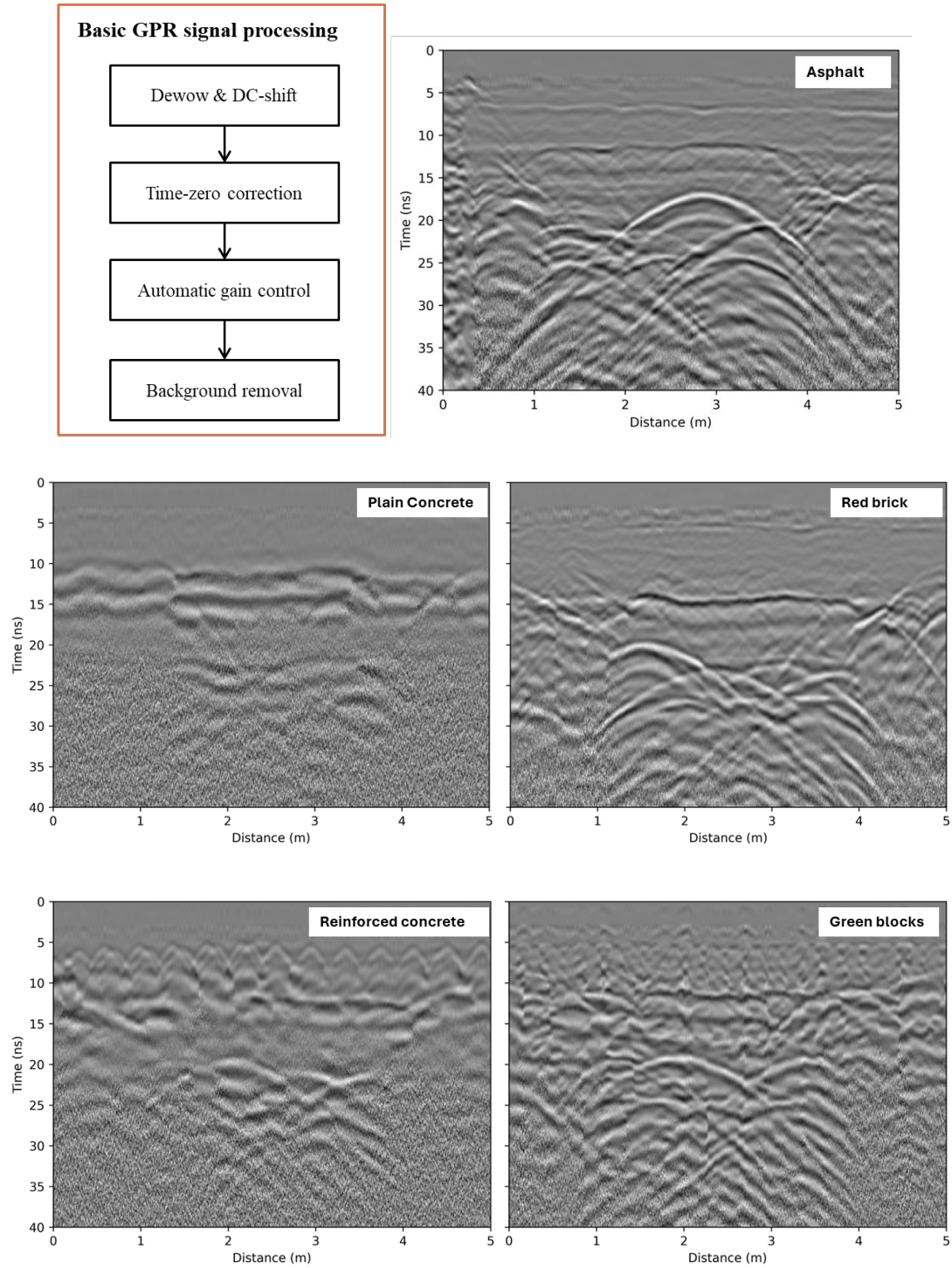


Figure 4-3 GPR B-scans across buried pipes under different pavement types using a 600 MHz IDS Hi-Mode antenna (dry conditions)

Figure 4-3 presents the processed GPR B-scans collected across buried pipes under different pavement types. The data was acquired using a 600 MHz IDS Hi-Mode antenna in dry conditions (with no man-made leakage introduced) and processed following the basic post-

processing workflow. Each survey line crosses four pipes with different pavements. Pavement type significantly influenced the quality of GPR signals. Asphalt and red brick pavements provided the clearest reflections, with well-defined pipe hyperbolas and stratigraphic layers. Reinforced concrete, however, introduced interference from reinforcing bars, causing discontinuous hyperbolas and obscuring deeper signals. Green block pavements, with their uneven surfaces and voids for plant growth, caused reverberations and vertical stripe noise, distorting deeper hyperbolas. Plain concrete posed the greatest challenges, exhibiting weak contrast between pipes and the background, severe attenuation, and scattering noise, which blurred hyperbolic reflections and stratigraphy. In summary, asphalt and red brick pavements were the most favourable materials for GPR surveys, followed by reinforced concrete and green blocks, while plain concrete offered the least favourable conditions for signal clarity.

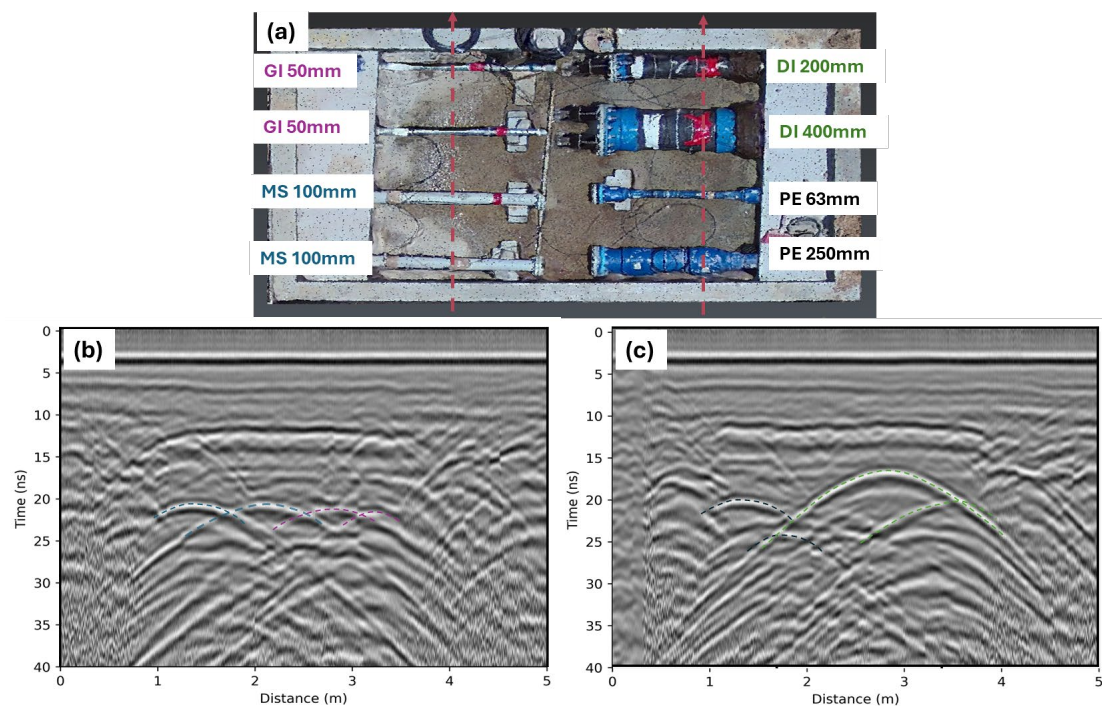


Figure 4-4 (a) Top-view photo of Cell 5 pipe configurations; (b) GPR B-scan from the left survey line; (c) GPR B-scan from the right survey line

Figure 4-4 (a) shows a top-view photo of the pipe configurations in Cell 5, which included four types of pipe materials (GI, MS, DI, and PE) with varying diameters. The

corresponding GPR B-scans (Figures 4-4 (b) and (c)) illustrate the influence of pipe size on hyperbolic reflections. Smaller pipes, such as GI 50 mm and PE 63 mm, produced discontinuous and incomplete hyperbolas due to interference from closely spaced pipes. Larger pipes, on the other hand, generated more distinct reflections, making them easier to identify. This analysis highlights the challenges posed by pipe size and material on GPR signal interpretation.

Based on the Q-Leak Center's configuration, three experimental datasets were developed to address specific research objectives:

- Dataset One: Focused on the pattern recognition of temporal B-scans across various pavement types and leakage scenarios.
- Dataset Two: Centered on temporal C-scan analysis, considering variations in pavement and leakage types.
- Dataset Three: Designed to investigate temporal C-scan patterns, with a particular focus on different leakage volumes.

4.1.2 Dataset One: Temporal B-scans of Buried Pipes Leakage

Underground hazards, such as pipe leakage, typically exhibit delayed or hysteretic behaviour. In real-world situations, GPR surveys are often initiated only after observable signs — such as abnormal readings from water pressure gauges or flow meters, or visible water leakage — have been detected. This delayed response introduces two significant challenges for GPR-based detection: First, leakages are often detected at later stages of their development, which means that earlier-stage leakage data, which could provide critical insights for early warning systems, is unavailable. Second, background GPR measurements—those taken from the same region but without leakage—are often missing. These background measurements are essential for filtering out non-leakage-related reflections, thereby improving the accuracy of

leakage detection.

The absence of both early-stage leakage data and background GPR measurements poses a challenge to effective pattern recognition and detection. To address this gap, experimental datasets have increasingly relied on simulations and small-scale laboratory experiments. However, these simulated datasets often fail to accurately replicate the complexity of real-world environments, limiting their applicability to practical scenarios. In response to these challenges, we designed the leakage experiments in the Q-Leak centre to capture temporal B-scan patterns, comparing GPR measurements taken before and after the occurrence of leakage.

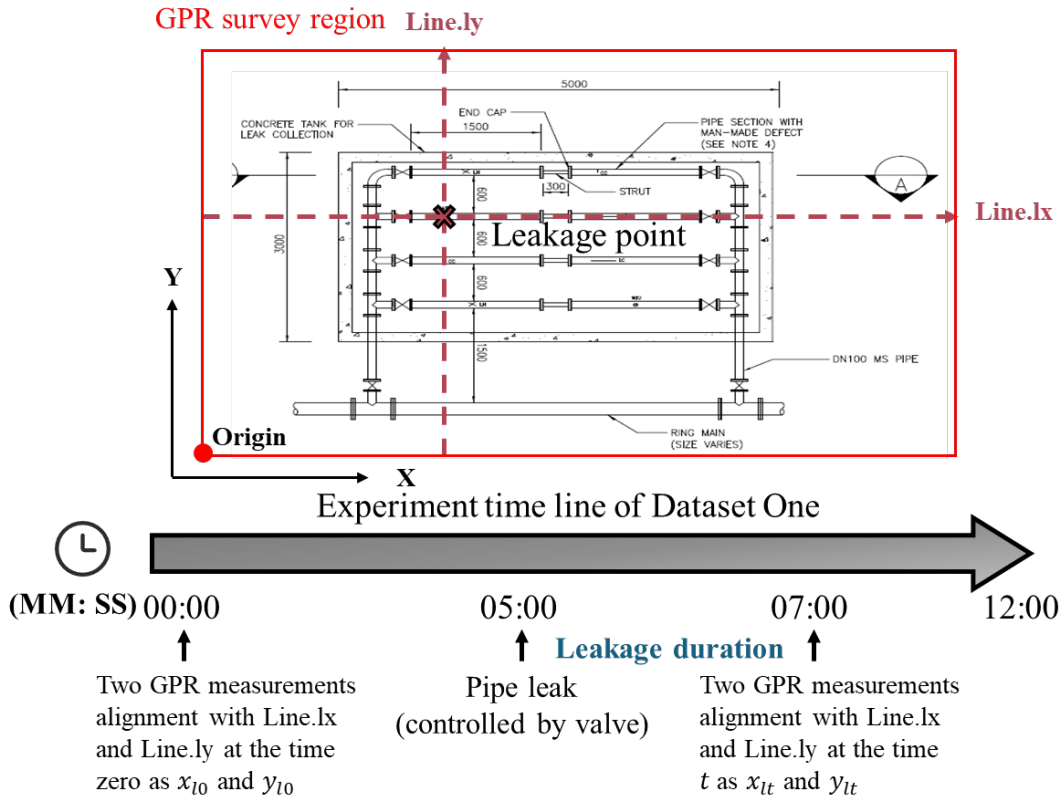


Figure 4-5 Experimental timeline for Dataset One and survey line layout overlaid on the cell schematic

The experimental process for each leakage point is outlined in Figure 4-5. Initially, two GPR measurements were taken along the survey lines (denoted as Line.lx and Line.ly) at the beginning of the experiment, represented as x_{l0} and y_{l0} . Subsequently, the leakage was

simulated by switching on the control valve connected to the buried pipe at the identified leakage point. After allowing a leakage period of two minutes, two additional GPR measurements were collected along the same survey lines (x_{lt} and y_{lt}). For each leakage point, two pairs of GPR measurements were taken—one pair before the leakage event and one pair after. This experimental setup allows for a direct comparison of GPR responses in the presence and absence of leakage, providing valuable data for pattern recognition and improving the understanding of GPR signal behaviour in leakage detection.

In total, 34 leakage scenarios were conducted, each consisting of four GPR measurements captured along two orthogonal survey lines. Table 4-3 summarizes data from six representative scenarios in Dataset One, illustrating the variety of pipe attributes, defect types, and pavement materials considered. The table begins by detailing essential attributes of the buried pipes, such as material, size, and burial depth. It also includes descriptions of two orthogonal GPR survey lines for each leakage scenario, explicitly marking the leakage point location. The final columns provide B-scans recorded before and after leakage events, with arrows indicating the leakage point in the B-scan views. These B-scans were processed using standard post-processing techniques, including DC shift removal, Dewow, and time-zero correction. Following these steps, AGC gain was applied to improve the clarity of texture patterns in the data. To preserve critical layer patterns, such as those corresponding to the water table, a background removal filter was intentionally omitted. While this approach retains essential subsurface features, it also results in some ringing noise, particularly in scenarios involving concrete pavement.

Table 4-3 Representative data samples from Dataset One (Raptor45)

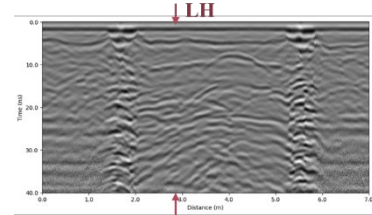
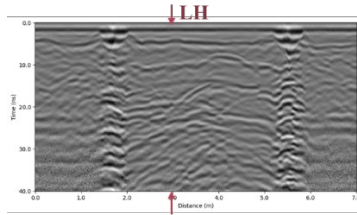
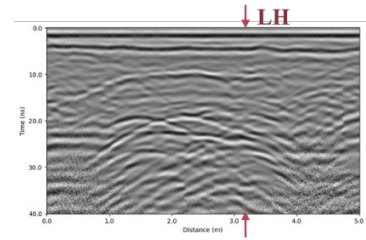
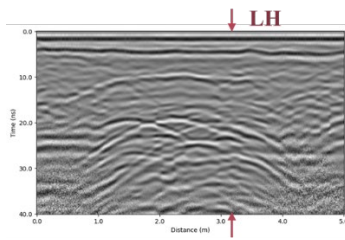
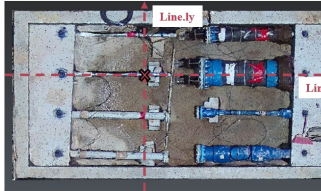
Pipes' attributes & leak point locations	Without leakage	With leakage
	(Line.ly & Line.lx)	(Line.ly & Line.lx)
Scenario 1		

Defect type: Leak hole (LH)

Pipe size: 50mm

Pipe materials: GI

Pavement: Asphalt



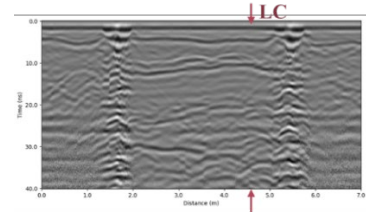
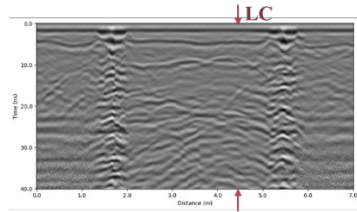
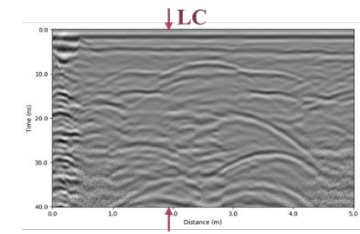
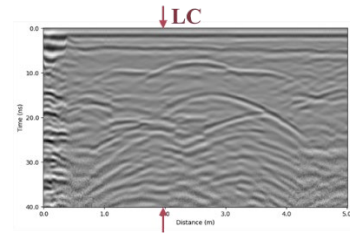
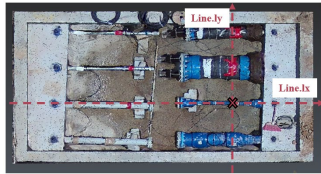
Scenario 2

Defect type: Longitudinal crack (LC)

Pipe size: 63mm

Pipe materials: PE

Pavement: Asphalt



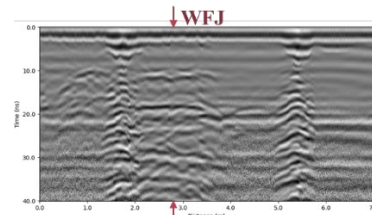
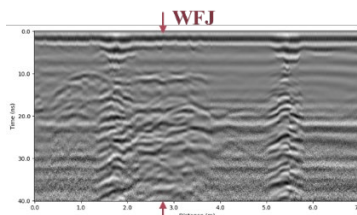
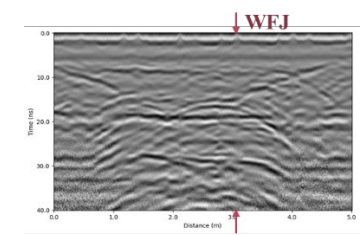
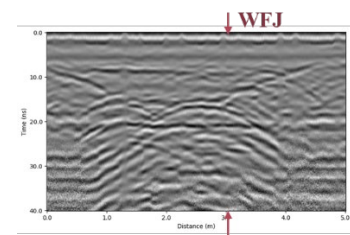
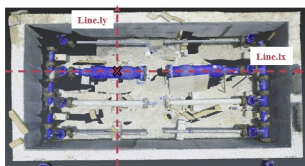
Scenario 3

Defect type: Weak flange joint (WFJ)

Pipe size: 250mm

Pipe materials: PE

Pavement: Green blocks



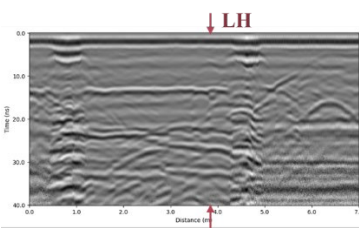
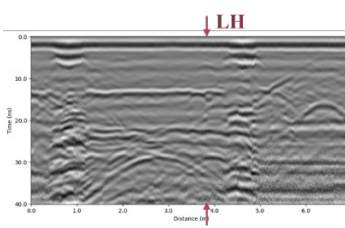
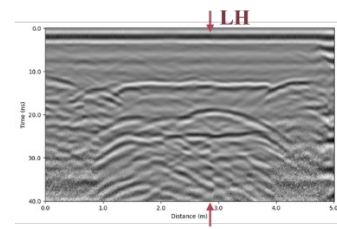
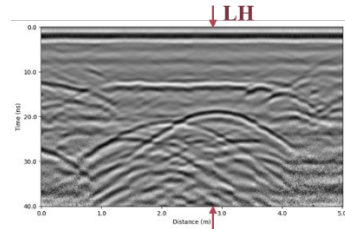
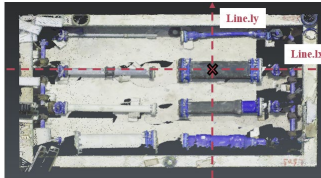
Scenario 4

Defect type: Leak hole

Pipe size: 300mm

Pipe materials: PE

Pavement: Red brick



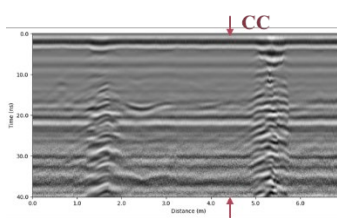
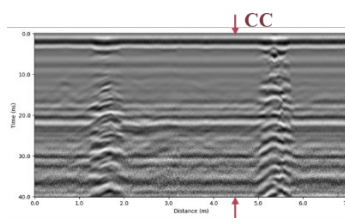
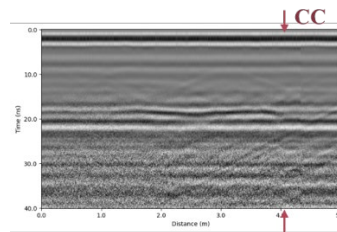
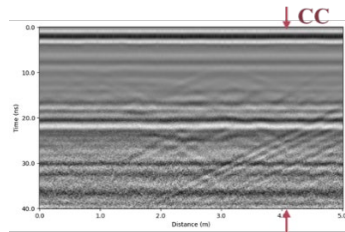
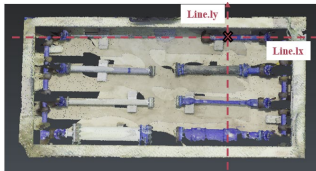
Scenario 5

Defect type: Circumferential crack (CC)

Pipe size: 100mm

Pipe materials: DI

Pavement: Plain concrete



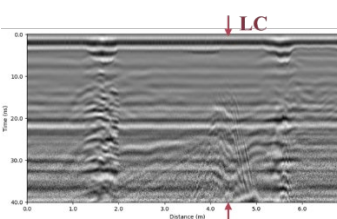
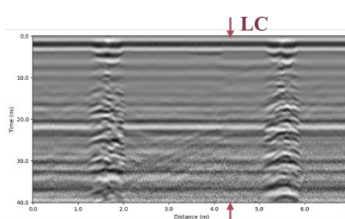
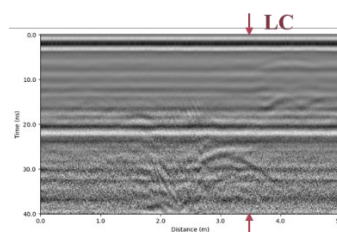
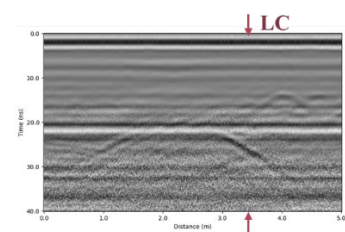
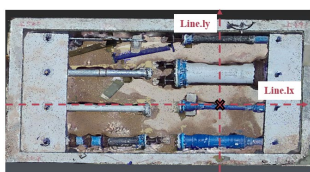
Scenario 6

Defect type: Longitudinal crack (LC)

Pipe size: 63mm

Pipe materials: PE

Pavement: Reinforced concrete



Despite the structured observations, recognizing leakage patterns directly from single-shot B-scans remains challenging due to the complexity and variability of GPR signals. Without background reference measurements, distinguishing true leakage features from artefacts becomes ambiguous. For instance, common indicators of leakage, such as "layer patterns," can also result from dielectric contrasts between non-leakage materials. Similarly, antenna ringing can produce horizontal stripes that mimic leakage features. These limitations highlight the difficulty of relying solely on single-shot B-scans for accurate leakage detection.

To address this, GPR measurements before and after leakage offer a new perspective by extending leakage pattern recognition to include temporal features. Our dataset extends leakage pattern recognition by incorporating temporal features through paired GPR measurements before and after leakage events. This pairing provides critical background reference, enabling a more comprehensive analysis of leakage patterns. In relatively GPR-friendly pavements like asphalt, red brick, and green blocks, paired data reveal additional horizontal layers ("water tables") and downward hyperbola shifts due to reduced wave velocity caused by increased moisture content. For example, in Scenarios 2, 3, and 4, leakage forms a reflective interface at the water level, while the increased moisture content alters the wave velocity, creating identifiable features in the B-scans. Conversely, in less favourable conditions such as concrete and reinforced concrete pavements (e.g., Scenarios 5 and 6), recognizable patterns are harder to discern even with paired measurements. Here, leakage effects are often limited to signal distortions or increased noise, providing only subtle clues. However, these paired observations still allow for better context and enhance the ability to distinguish true leakage from noise or artefacts.

In summary, Dataset One provides a robust foundation for advancing leakage detection by addressing the limitations of single-shot B-scans. By including real-world GPR measurements with and without leakage, alongside explicitly marked leakage locations, it

enables the development of more reliable recognition techniques tailored to the complex variability of underground environments.

4.1.3 Dataset Two: Temporal C-scans of Buried Pipes Leakage

While Dataset One effectively captures a variety of real-life leakage scenarios and provides valuable insights into leakage patterns and reflections in B-scan data, it has significant limitations when applied to change detection technologies in temporal GPR measurements. One major challenge arises from the misalignment of GPR survey lines in open environments. During road surveys, it is practically impossible for operators to ensure that the GPR system scans the same lines as in previous surveys, making consistent B-scan alignment unattainable. This operational limitation hinders the application of change detection strategies based on temporal B-scans, particularly in large-scale road surveys.

In contrast, C-scans overcome these alignment issues through the integration of intersecting temporal GPR measurements, allowing for the effective application of change detection methods. As discussed in Chapter 2, C-scans offer additional advantages over B-scans, particularly for large-scale leakage detection. They are the preferred representation for implementing change detection strategies in extensive road surveys, making them indispensable for practical applications. To address these challenges, Dataset Two was specifically designed to capture leakage features in C-scans and enable robust change detection.

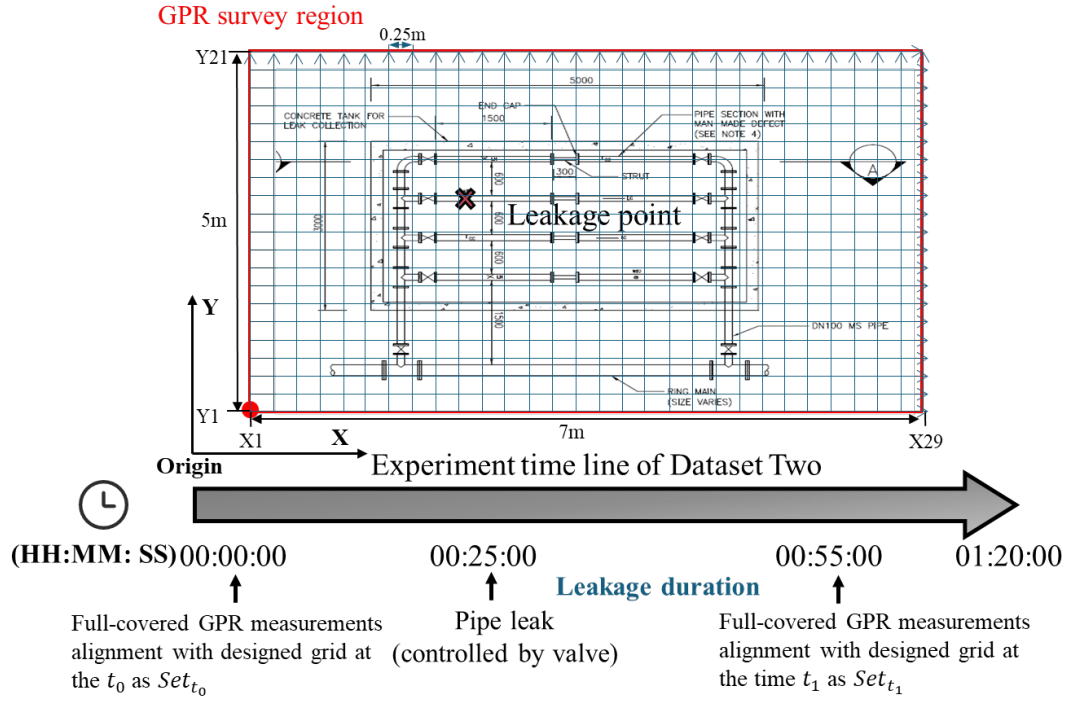
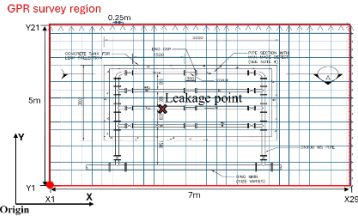
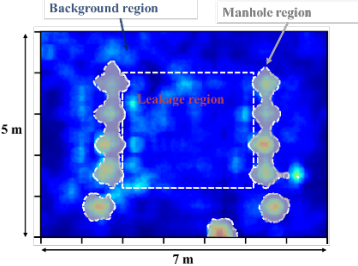
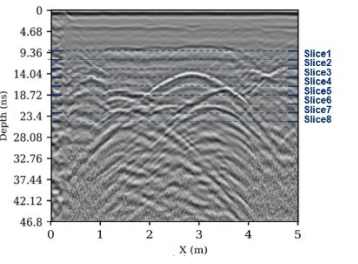
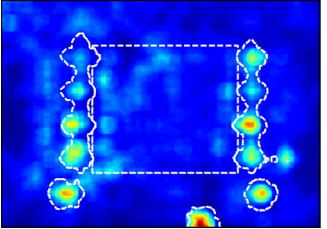
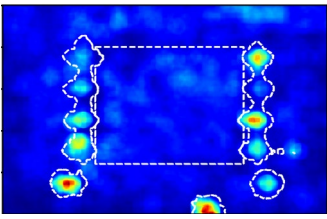
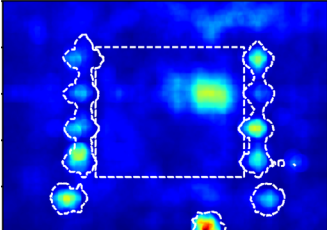
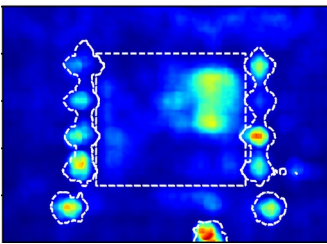
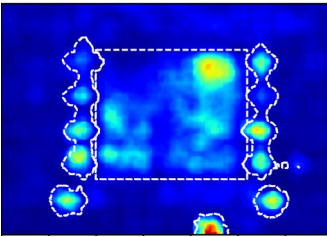
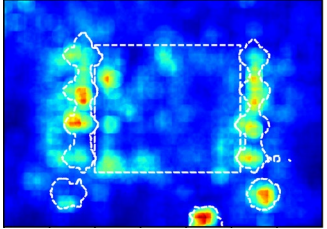
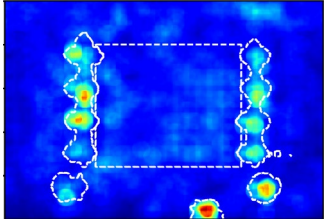
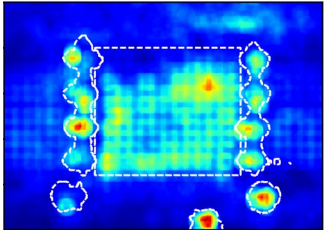
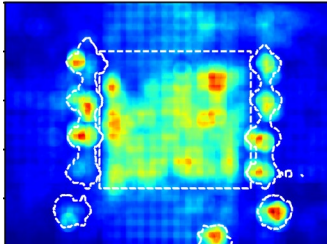
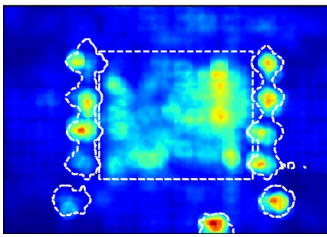


Figure 4-6 Experimental timeline for Dataset Two and survey grid overlaid with the cell schematic

The experimental process is illustrated in Figure 4-6. Initially, comprehensive GPR measurements were collected over a predefined survey grid, with orthogonal scan directions and a spacing of 0.25 meters, before any leakage event (denoted as Set_{t_0}). Leakage was then simulated by opening the control valve associated with the buried pipe at the identified leakage point. After a designated leakage period, another set of full-coverage GPR measurements was collected using the same survey grid configuration at the time t_1 , denoted as Set_{t_1} . For each leakage point, this setup yielded two complete sets of C-scan data: one before the leakage event and one after. The direct comparison of these datasets provides critical information for pattern recognition and enhances the understanding of leakage features in C-scans. This approach is particularly suited for large-scale applications, addressing the operational limitations of B-scans. Dataset Two comprises 24 leakage scenarios, encompassing varying conditions and utilizing GPR systems equipped with antennas operating within a frequency range of 200 MHz to 700 MHz (IDS Hi-Mode, Raptor45, DS 2000, and StreamUP). For each scenario, two GPR measurements were conducted over the survey grid, one before and one after the simulated

leakage event. Table 4-4 presents a representative sample from Dataset Two.

Table 4-4 Representative data samples from Dataset Two (IDS Hi-Mode 600MHz)

Slice parameters and on-site photo	Amplitude slices without leakage (Slice 1-8)	Amplitude slices with leakage (Slice 1-8)
Pavement: Asphalt Profile spacing: 0.25 m Slice thickness: 2.4 ns Search radius: 0.3 m Interpolation methods: invest distance   	    	    

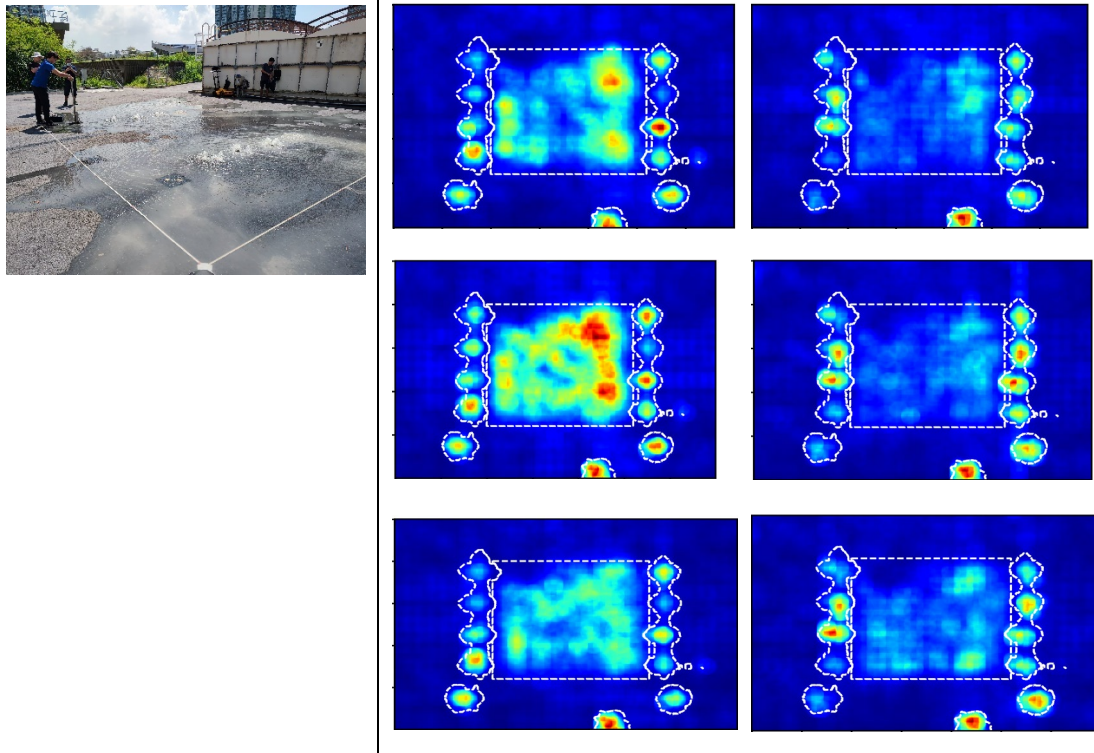


Table 4-4 provides an example of Dataset Two, showcasing GPR data from a leakage scenario conducted over Cell 5 with asphalt pavement. The survey included two full-grid GPR measurements—one taken before and one after a simulated leakage event. The on-site photo in the first column illustrates the setup. During the experiment, the leakage water filled a concrete tank and overflowed onto the surface. Profile spacing is 0.25 meters, and slice thickness is 2.4 ns. The slices were generated using the inverse distance interpolation method with a power factor of 2 and a search radius of 0.4 m, according to Luo et al (Luo et al., 2019b). A total of eight horizontal slices were produced, spanning depths from 56 cm to 136 cm at 10 cm intervals, detailed in Table 4-5. This depth range excluded the pavement and underlying soil, focusing instead on imaging the top and bottom boundaries of the buried concrete tank.

Table 4-5 Configurations of GPR survey and parameters of C-scan generation.

GPR survey		C-scan generation	
Antenna	Leica DS2000/IDS Hi-mode	Grid spacing (X & Y)	0.25 m
Frequency	700 MHz/600 MHz/250 MHz	Profile spacing	0.25 m
Time window	50 ns	Slice thickness	2.4 ns

Sampling interval	0.0114 m/0.0089 m/0.0114 m	Interpolation method	Inverse distance
Time samples	512	Search radius	0.4 m

For analysis, the survey area was divided into three regions: the manhole region, the leakage region, and the background region. Initially, the manhole region is omitted due to its irregular signal reflection, characterized by significant reverberation—a result of the high-energy surface reflection of manholes. In Table 4-4, this region is represented by a grey mask derived using a manhole filter. The leakage region is designed as a subterranean concrete enclosure to prevent the spread of leakage water into the background region. Consequently, variations in the background region, unaffected by leakage water, represent natural environmental changes across different scenarios.

The horizontal slices were generated by summing amplitude values from batches of B-scans, providing amplitude distribution maps of underground features at different depths. In the slices without leakage, high-reflection clusters associated with buried pipes and concrete tanks are visible. For instance, in Slice 3, the largest pipe produces a high-amplitude signal, while smaller pipes are partly visible due to interference caused by their close spacing. This interference results in unclear linear patterns in the slices, as hyperbolas from adjacent pipes overlap. Additionally, high reflections from manholes are prominent in all slices, appearing as two parallel lines. In the slices after leakage, signal distortions caused by water content changes manifest as altered amplitude patterns. Comparing different regions, the background region shows minimal changes, which is expected since the concrete tank contains the leakage water, preventing its spread. The manhole region, masked due to unstable signals, remains excluded from analysis. In the leakage region, significant amplitude variations were observed. From Slice 1 to Slice 4, an increase in amplitude indicates enhanced reflections caused by leakage-induced changes. Conversely, from Slice 5 to Slice 8, the amplitude decreases, likely due to the attenuation of electromagnetic waves as water saturates the region. These patterns suggest that

water leakage can increase reflection at shallow depths but also attenuate signals at greater depths. A comparison of slices with and without leakage highlights the leakage region more distinctly, aligning well with the known leakage location.

This example highlights the potential of amplitude slices in effectively identifying leakage regions, even under challenging conditions. Dataset Two encompasses 12 leakage scenarios with varying leakage rates and conditions, as detailed in the appendix. While some scenarios involve substantial leakage, others present subtle changes, making detection more demanding. Notably, C-scan generation serves as a method to distil and summarize information from the volumetric data of B-scans. Although amplitude slices are the most commonly used type, other attributes derived from GPR signals also hold significant potential. These attribute-based slices can emphasize specific subsurface phenomena by reflecting how the attributes correspond to those features. In the following Section 4.2, we propose a workflow for generating IFS from GPR data volumes. The IFS approach demonstrates a strong capability to characterize buried pipe leakage, as validated by the diverse scenarios in Dataset Two. This diversity offers a robust foundation for exploring advanced leakage detection patterns and methodologies, enabling a more comprehensive understanding and application of GPR imaging techniques.

4.1.4 Dataset Three: Temporal C-scans of Buried Pipes with Various Leakage Volume

Building on the experimental framework established in Dataset Two, Dataset Three follows similar procedures but extends the scope to capture the entire progression of leakage events. In this dataset, additional GPR measurements were periodically conducted during the leakage process, continuing until the leakage water fully saturated the cell and began to overflow onto the surface, as illustrated in Figure 4-7. This extended observation period enables

a detailed characterization of leakage dynamics, capturing GPR responses across various stages—from the initial onset of leakage to severe leakage with surface manifestations.

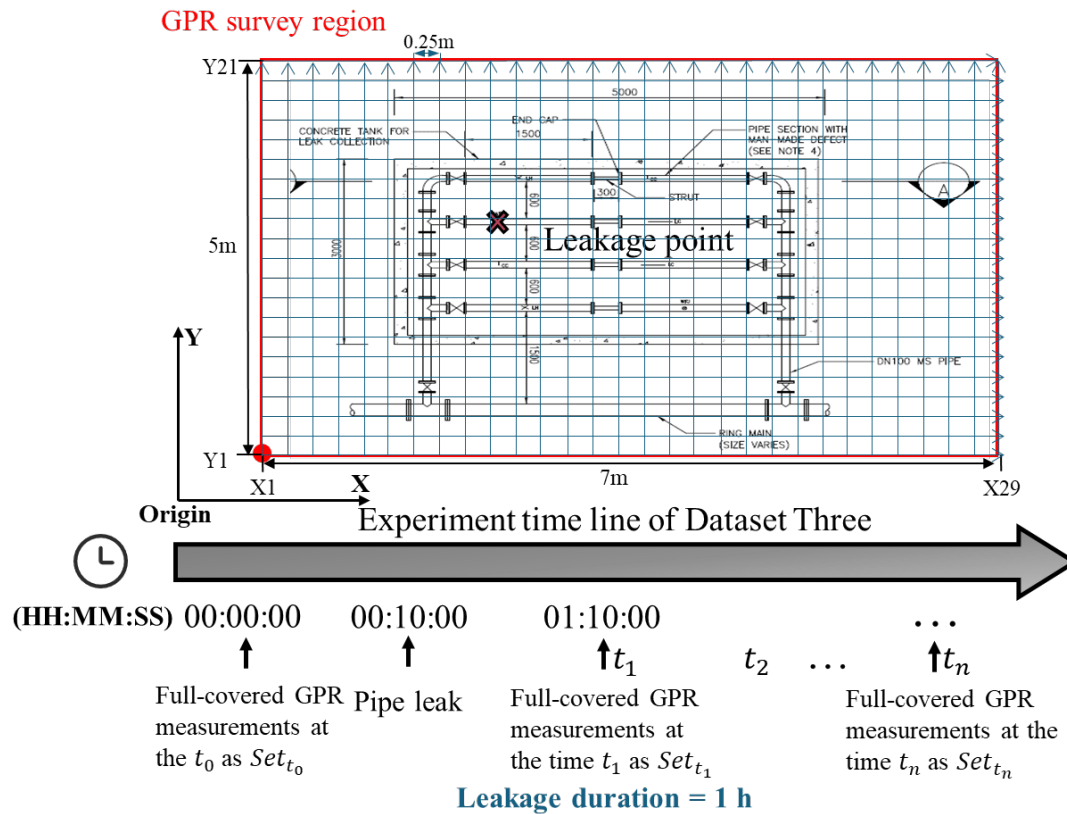
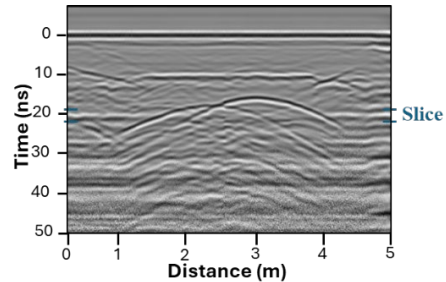
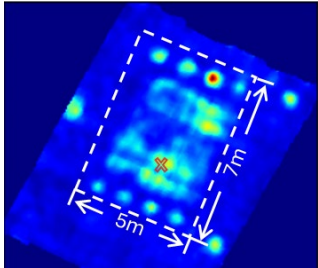
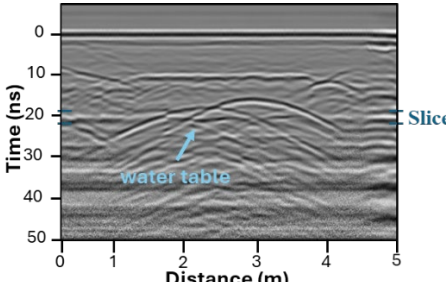
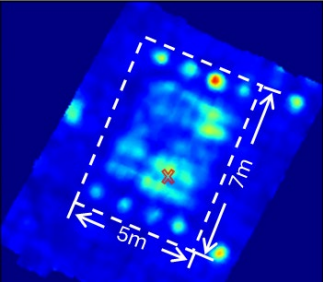
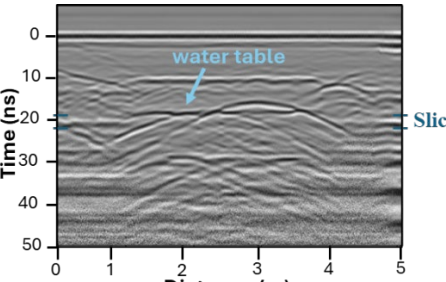
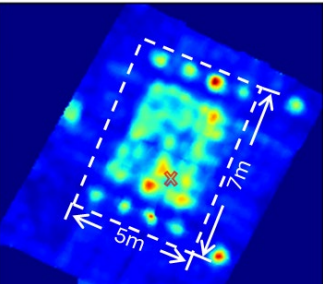
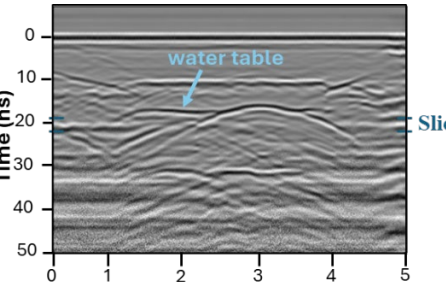
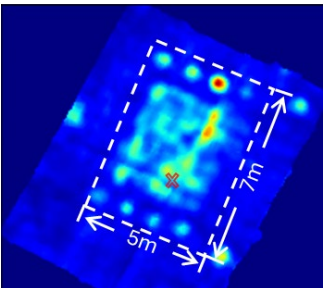


Figure 4-7 Experimental timeline for Dataset Three and survey grid overlaid with the cell schematic

Dataset Three was specifically designed to provide a comprehensive understanding of the temporal evolution of leakage, offering rich data for advanced analysis. Each leakage scenario includes 2–5 phases of GPR measurements, depending on the leakage rate, with observations concluding when the leakage water fills the tank. In total, 20 leakage scenarios were conducted, utilizing IDS Hi-Mode, DS 2000, and Raptor45 systems. This dataset supports advanced studies of the temporal progression of leakage patterns, enhancing the applicability of change detection methods for scenarios ranging from early-stage leakage to severe underground water loss.

Table 4-6 Representative data sample from Dataset Three (Raptor45)

Observation time	GPR measurement cross leak pipes	Slices (Depth: 18.5 ns – 21.5 ns)
Baseline (without leakage)		
Phase 1 (after 1 hour)		
Phase 2 (after 2 hours)		
Phase 3 (after 3 hours)		

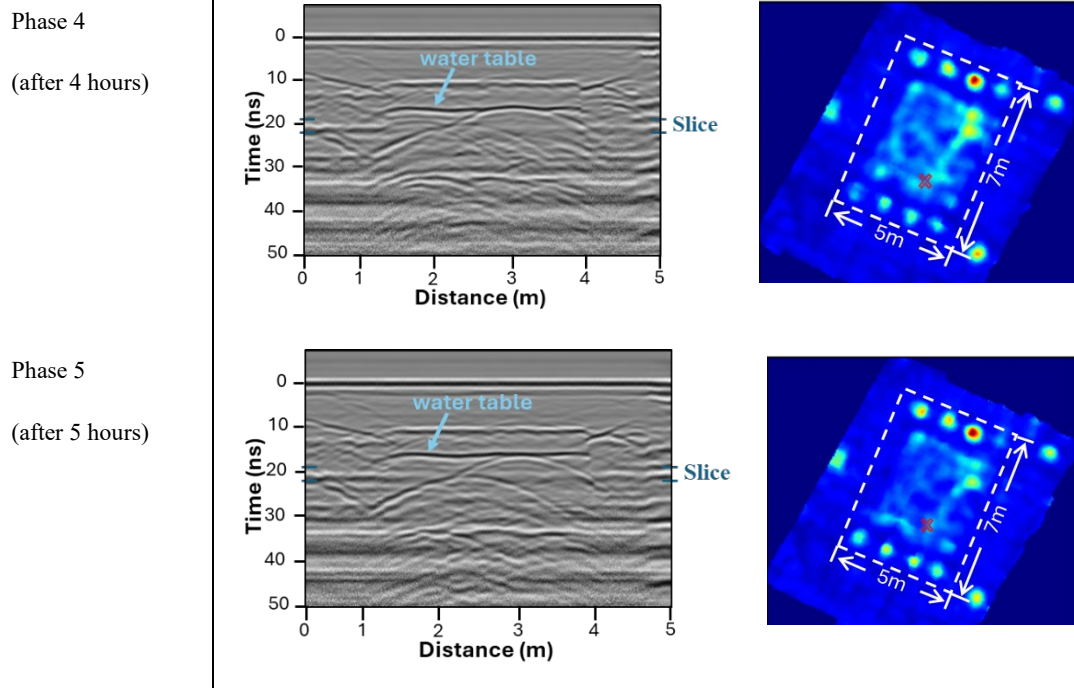


Table 4-6 provides a representative sample from Dataset Three, showcasing five GPR measurements conducted at one-hour intervals following a simulated leakage event on Cell 2. These measurements were carried out using the Raptor45 MCGPR system with 8 channels. The high efficiency of the system allowed for full cell coverage within approximately 10 minutes, ensuring minimal leakage volume differences within a single measurement. The table includes two key representations: B-scans, displaying profiles that cross buried pipes and leakage points, and amplitude slices, highlighting reflections within the depth range of 18.5 ns to 21.5 ns. The baseline slice (without leakage) serves as a reference, with the leakage point marked by a cross.

As leakage progresses, the water table layer becomes increasingly shallow, indicating the growing volume of leakage water in the tank. From Phase 2 onward, the bottom layer of the tank becomes more distinct and begins to shift downward within the range of 28 ns–32 ns, a phenomenon attributed to the reduced wave velocity in water-saturated regions. Following Phase 2, hyperbolas associated with buried pipes become more pronounced, while patterns below the water table appear smoothed. In later phases (e.g., Phases 4 and 5), reflections between pipe hyperbolas and the bottom tank layer also exhibit smoother characteristics,

suggesting water's influence on signal propagation.

Amplitude slices reveal notable changes across phases. Early phases (2 and 3) show increased amplitudes, reflecting enhanced material interface contrasts, such as at the water table. Conversely, later phases (4 and 5) exhibit reduced amplitudes, likely due to wave attenuation in water-saturated areas. The consistency between baseline and Phase 1 slices underscores the reliability and quality of the amplitude data.

The analysis highlights two primary effects of water leakage on reflected GPR signals:

- (1). Water amplifies the contrast of different materials. Leakage enhances reflections at material interfaces, creating distinct features such as the water table.
- (2). Water reduces the heterogeneities within the same materials.

In regions with high structural heterogeneity, water smooths reflections, diminishing variations, and produces clearer signals. The temporal patterns observed in Dataset Three underscore the potential for leakage detection through time-series GPR measurements. Amplitude slices effectively capture both amplification and attenuation effects. With its comprehensive collection of realistic simulated leakage data, this dataset serves as an excellent resource for developing advanced change detection strategies, paving the way for more precise and reliable monitoring of leakage events.

4.2 Characterization of Leakage Signatures using Instantaneous Frequency

As introduced in Section 2.4, detecting localized water leakages using GPR presents complex patterns in the time domain. However, water content also exhibits distinct characteristics in the frequency domain, offering additional insights for analysis. Water trapped and bound within dielectric materials—such as soil, asphalt, and concrete—produces effects

similar to free water but at significantly lower relaxation frequencies. For example, water has general relaxation frequencies of approximately 500 MHz in soil (De Loor, 1983; Jol, 2009a), 1000 MHz in asphalt (Shang & Umana, 1999b), and 1200 MHz in concrete (Lai et al., 2011; Xiao et al., 2017).

A material's relaxation frequency depends on the volumetric fractions and forms of free water, which vary based on the water's state—whether fully bound, partially bound, or free—and the volumetric content of pore water across a wide range of porosities in soil, concrete, and asphalt. These properties make the frequency domain a valuable tool for characterizing water infiltration and leakage. Previous studies have demonstrated the potential of joint time-frequency analysis—including wavelet and STFT—to characterize transient wetting and drying processes in soil using GPR data. However, the limitations of these techniques, such as the windowing restrictions in STFT and wavelet transforms, can restrict their ability to accurately capture localized leakage dynamics. This study aims to address these limitations by developing a new method tailored to real-world leakage scenarios.

To achieve this, we utilized a well-designed experimental site that emulates realistic leakage conditions in buried pipes. Our dataset, introduced in the previous section, serves as the foundation for this analysis. GPR measurements were conducted using antennas operating at frequencies ranging from 250 MHz to 700 MHz, covering the relevant relaxation frequencies of the materials in the experimental setup, including soil, asphalt, and concrete. Our analysis is focused on IFS, which provides spatially resolved insights into variations in GPR signal propagation caused by changes in water content. The objectives of this study are twofold:

- (1). To establish a robust workflow for generating IFS that accurately represents spatial instantaneous frequency variations.
- (2). To investigate and characterize the patterns of leakage in buried pipes using IFS, identifying

signatures specific to leakage-induced water content fluctuations.

By addressing these objectives, this study contributes to the advancement of GPR-based water leakage detection methods, presenting a novel approach to overcome the challenges of time-frequency analysis and providing a more detailed understanding of leakage dynamics in subsurface environments.

4.2.1 Workflow of IFS Generation based on GPR Data

In our research, we focused on the characterization of buried pipes and the recognition of patterns from IFS to characterize the leakage patterns of buried pipes. Figure 4-8 provides a comprehensive outline of the procedures involved in generating IFS, also the AS was included.

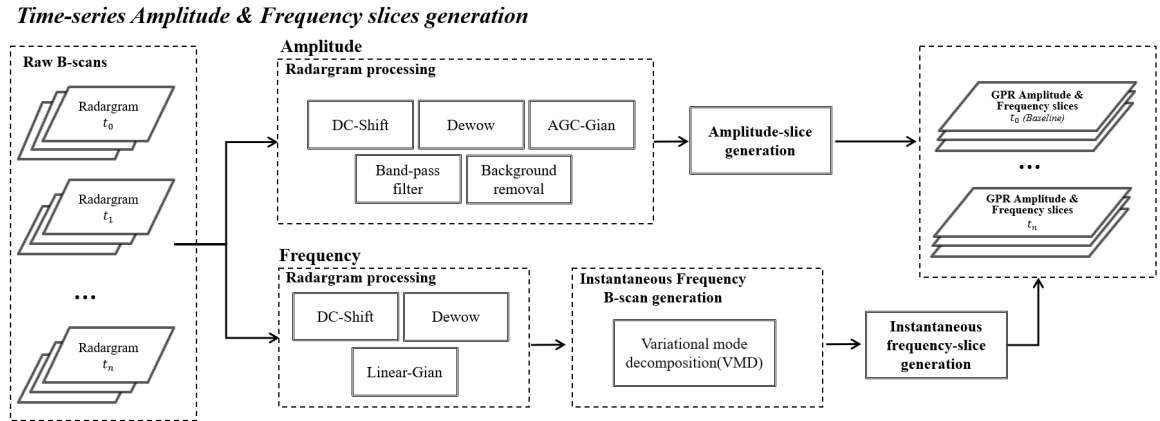


Figure 4-8 The workflow of the AS and IFS generation

In this study, the data originates from Dataset Two, as described in Section 4.1. The details of the leakage experiment and AS generation are summarized in Table 4-7 for reference.

Table 4-7 Configurations of GPR survey and parameters of C-scan generation.

GPR survey		C-scan generation	
Antenna	Leica DS2000/IDS Hi-mode	Grid spacing (X & Y)	0.25 m
Frequency	700 MHz/600 MHz/250 MHz	Profile spacing	0.25 m
Time window	50 ns	Slice thickness	0.6λ
Sampling interval	0.0114 m/0.0089 m/0.0114 m	Interpolation method	Bi-linear

While AS provides insights into underground conditions by extracting energy distribution from reflected signals—deriving from the permittivity of host materials to image underground objects horizontally—IFS aims to reveal the dominant instantaneous frequency at various depths. The signal processing functions can introduce artificial noise in the frequency domain (such as bandpass), and the generation of IFS adopts a different approach. Although AGC gain and background removal functions do not directly impact frequency components, during the conversion from time amplitude to time-frequency, the amplitude is utilized to determine the dominant frequency component at a specific depth. Therefore, it is crucial to maintain a consistent gain at the same depth. To achieve this, the AGC gain was replaced with a two-point linear gain (-5 dB at the electronic zero and +45 dB at the end of the A-scan) (Lai et al., 2014). Initially, the A-scans were processed using a direct current shift and a De-wow filter. Following this basic 2D signal processing, variational mode decomposition (VMD) (Dragomiretskiy & Zosso, 2014) was utilized to generate the IFS.

Joint time-frequency analysis (JTFA) techniques are essential tools for signal analysis, particularly useful in interpreting GPR data. These methods effectively reveal how frequency components change over time. VMD is used to decompose the GPR signal and obtain the instantaneous frequency, leveraging its high time-frequency resolution and adaptive features. These benefits have been demonstrated through comparative studies (He et al., 2020). Benefiting from the characterizations of VMD, there are various related applications in GPR fields. Zhang et al. proposed methods that utilize VMD for GPR radargram denoising (Zhang et al., 2018). Lu et al. demonstrate the performance of using VMD for oil-contaminated soil detection (Lu et al., 2023). Besides, He et al. evaluate the capability and vast potential of VMD-based methods to locate underground utilities (He et al., 2020).

In the VMD algorithm, the decomposition could be expressed by solving the

constrained optimization as follows:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \|\partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \|^2_2 \right\} \\ s. t. \sum_k u_k(t) = f(t) \end{cases}, \quad [4-1]$$

where $\{u_k\}$ represent the mode components, $\{\omega_k\}$ is the centre frequency of each mod. The symbol f denotes the input signal. Dragomiretskiy et al. introduced the Lagrange multipliers and the penalty item into Equation [4-1], represented as:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_k \|\partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \|^2_2 + \|f(t) - \sum_k u_k(t)\|^2_2 + \\ & \langle \lambda, f - \sum_k u_k(t) \rangle. \end{aligned} \quad [4-2]$$

Detailed descriptions of these components can be found in the literature (Dragomiretskiy & Zosso, 2014; Zhang, Nilot, et al., 2018). The number of IMFs k , is related to the number of stationary sub-components within the non-stationary signal. The penalty parameter α controls the bandwidth of IMFs. Normally, the initial values of k and α are manually pre-determined based on prior information about the region of interest. These parameters are then iteratively updated based on the results of the decomposition. In the scenarios lacking prior information, a swarm optimization method can be employed to determine the optimal combinations of k and α . In this paper, k and α were set as 4 and 2048, respectively, based on reasonable ranges referenced from existing literature (Liu et al., 2022). VMD offers a way to convert a time-amplitude domain signal into the time-frequency domain. This transformation is instrumental in extracting quantifiable wave properties, which serve as valuable descriptors for characterizing variations in water content.

The generation of an instantaneous frequency slice commences with the creation of an instantaneous frequency B-scan (IF B-scan). In this process, each A-scan is decomposed into several intrinsic mode functions (IMFs) using VMD. The raw A-scan can be conceptualized as a blend of these IMFs, represented as:

$$f(t) = \sum_{k=1}^K a_k(t) \cos(\theta_k(t)), \quad [4 - 3]$$

where $f(t)$ is the input signal as a function of time, $a_k(t)$ represents the amplitude of the k th intrinsic mode at time t , and $\theta_k(t)$ denotes the phase of the k th intrinsic mode at time t . After the decomposition, each IMF is subjected to a transformation into instantaneous phase, frequency, and amplitude through the function of an amplitude-normalized Hilbert transform algorithm.

The creation of an instantaneous frequency A-scan (IF A-scan) involves assigning the instantaneous frequency of each trace to identify major reflecting objects at distinct depths. The instantaneous frequency is derived from the IMF with the highest instantaneous amplitude, as shown in the workflow in Figure 4-9.

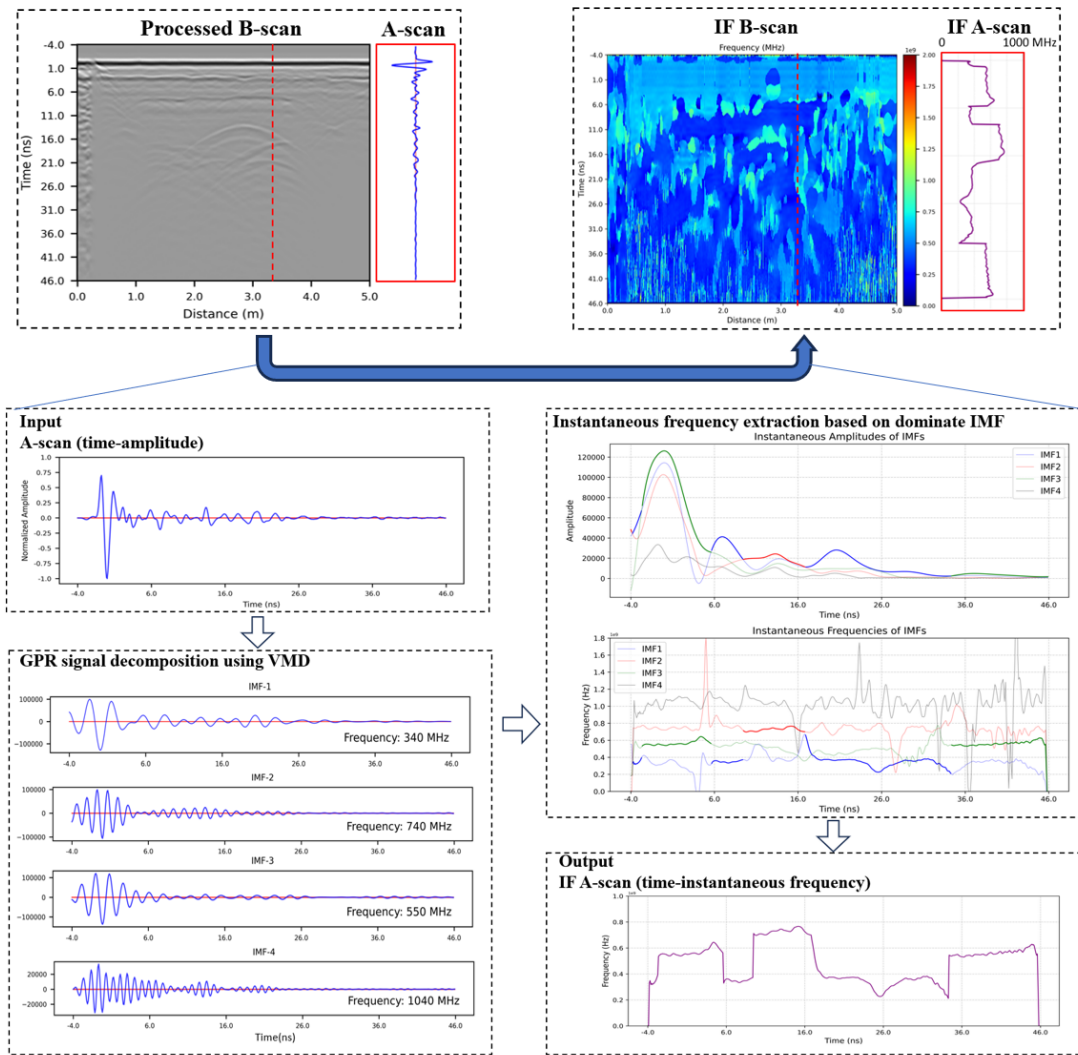


Figure 4-9 Schematic diagram of the process for generating instantaneous frequency B-scan via VMD from processed radargrams

Figure 4-9 illustrates the process of generating instantaneous frequency B-scan (IF B-scan) using VMD. The processed B-scan serves as the starting point, with a selected A-scan (time amplitude) undergoing decomposition into IMFs via VMD. The dominant IMF, based on its instantaneous amplitude, is then used to extract the instantaneous frequency for each point in the A-scan, resulting in the IF A-scan. The IF A-scans are subsequently assembled to create an IF B-scan, providing a two-dimensional representation of instantaneous frequency variations across the survey area. This process is outlined step-by-step in the schematic diagram, offering a comprehensive overview of the transformation from a processed radargram to its

corresponding IF B-scan.

This study aims to characterize pipe leakage by comparing the differences in the same underground layer before and after hydration. The increase in moisture content significantly diminishes the velocity of electromagnetic waves. To ensure that each slice accurately represents the same depth underground, a correction for slice thickness is required. Velocity estimation is a prominent topic in the GPR field. In this paper, we employed the most common ray-based model. To maintain focus, we did not delve into an extensive description of hyperbolic pattern fitting, as detailed discussions are available in the literature (He & Lai, 2024).

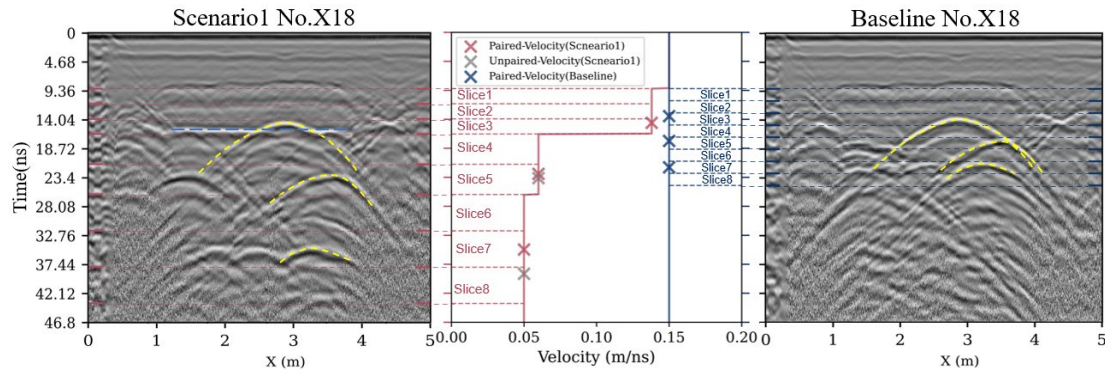


Figure 4-10 Illustration of dynamically adjusting slice thickness based on hyperbolic-based velocity estimation

Initially, baseline data were obtained in the absence of any water leakage. A constant velocity of 0.15 m/ns was estimated using the hyperbolic fitting method (ASTM, 2011). The fitting hyperbolas are indicated by yellow dashed lines in the radargram and blue crosses in the depth-velocity plot (as shown in the middle of Figure 4-10). Blue dashed lines divide the radargram into eight slices. In contrast, in leakage scenarios, the slice thickness varies at different depths due to the different velocities estimated from the hyperbolas. The locations of the hyperbolas used for velocity estimation are marked with red and grey crosses. We identified three paired hyperbolas in both the baseline radargram and the leakage scenario radargram. The relative relationship between the slices and paired hyperbolas demonstrates the effectiveness of thickness correction using estimated velocities.

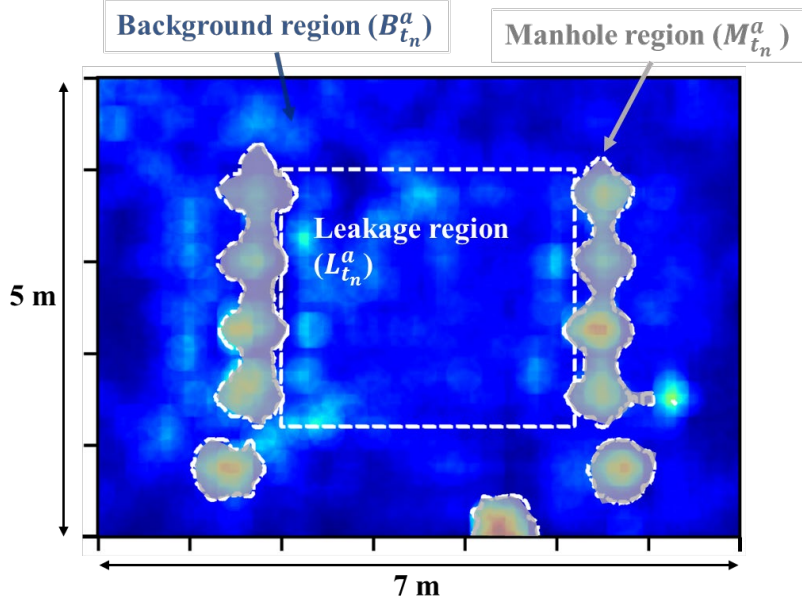


Figure 4-11 Illustration of the leakage region, background region, and manhole region at the experiment site

For targeted analysis, the study site has been partitioned into three distinct regions: the leakage region, the background region, and the manhole region, shown in Figure 4-11. These regions correspond to the pixel sets $L_{t_n}^a$, $B_{t_n}^a$ and $M_{t_n}^a$ in the slice, where t_n refers to the acquisition time in the time series, and a is the index of IFS ($1 \leq a \leq A$). Initially, the manhole region is omitted due to its irregular signal reflection, characterized by significant reverberation—a result of the high-energy surface reflection of manholes. In Figure 4-11, this region is represented by a grey mask derived using a manhole filter. The leakage region is designed as a subterranean concrete enclosure to prevent the spread of leakage water into the background region. Consequently, variations in the background region, unaffected by leakage water, represent natural environmental changes across different scenarios.

The primary aim of this study is to understand the characteristics of buried leakage pipes and to identify patterns from slices. Therefore, we employ three metrics to quantify the changes induced by leakage using paired slices. The centre drift, which may be the centre amplitude drift or IF drift, is defined as the difference in averages between times t_n and t_0 .

Given pixel set $R_{t_n} = \{r_{k_n} | k = 1, 2, \dots, K\}$ represent the slice at t_n the center drift can be represented mathematically as:

$$Center\ drift(R_{t_0}, R_{t_n}) = \sum_k^K r_{k_n} - r_{k_0}/K, \quad [4 - 4]$$

where r_{k_n} is the value of pixel k in the slice.

The Intersection Over Unit (IOU) measures the similarity between two histograms, focusing on the shape of the histogram curves, illustrated in Figure 4-12. It is computed as

$$IOU(R_{t_0}, R_{t_n}) = \frac{Intersection(R_{t_0}, R_{t_n})}{Union(R_{t_0}, R_{t_n})}. \quad [4 - 5]$$

IOU, commonly used in detection tasks as an accuracy assessment metric, in this context, is used to describe the similarity between distributions from a pair of slices. Furthermore, this study also employs the Pearson correlation coefficient, a statistical measure that calculates the correlation between two numerical variables. It ranges from -1 to 1, with -1 indicating a perfect negative correlation, 1 indicating a perfect positive correlation, and 0 indicating no correlation. The Pearson correlation coefficient (PCC), denoted as $\rho(R_{t_0}, R_{t_n})$, between two datasets obtained at t_0 and t_n is defined as:

$$\rho(R_{t_0}, R_{t_n}) = Cov(R_{t_0}, R_{t_n}) / (\sigma R_{t_0} * \sigma R_{t_n}), \quad [4 - 6]$$

where $Cov(R_{t_0}, R_{t_n})$ is the covariance between R_{t_0} and R_{t_n} , and σR_{t_0} and σR_{t_n} are the standard deviations of R_{t_0} and R_{t_n} , respectively. In this study, the Pearson correlation is used to measure the correlation of the slices before and after a water leakage event. Higher correlation coefficients (closer to 1) suggest that the water leakage has a minimal effect on the IF distribution, indicating that the signal remains relatively stable. In contrast, lower correlation coefficients (closer to 0 or negative values) indicate more substantial changes in the IF distribution due to the leakage. This provides a valuable measure to assess the impact of water leakage on the IF distribution in our study.

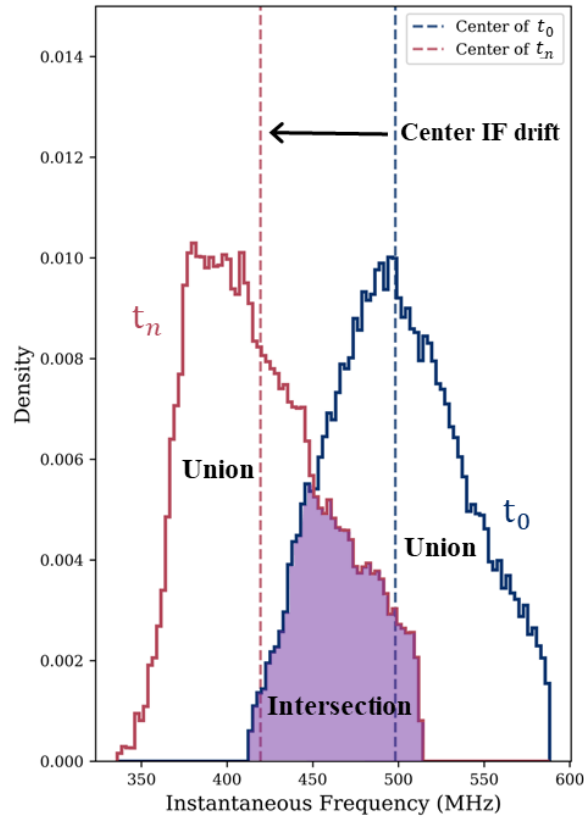


Figure 4-12 Illustration of center IF drift and IOU

4.2.2 Experiment Results and Discussion

The leakage dataset originates from Dataset Two of the leakage signature database described in Section 4.1.3. The buried pipe leakage scenarios were emulated in a custom-designed leakage cell, resembling a closed concrete tank with dimensions 500 cm $\times$ 300 cm $\times$ 150 cm (Figure 4-13). This concrete cell incorporated eight types of pipes commonly used in real-world settings to ensure a comprehensive representation of potential leakage scenarios, as shown in Figure 4-14.

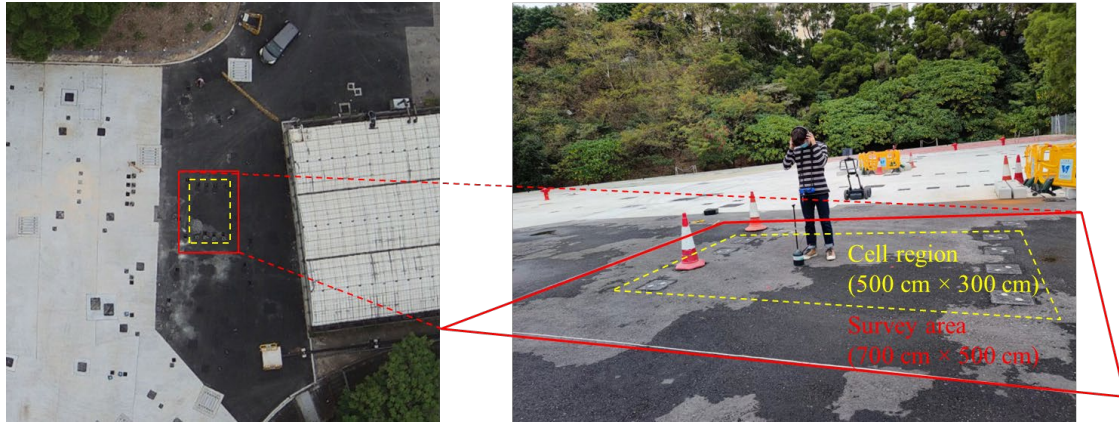


Figure 4-13 Site photos of the designed leakage cell

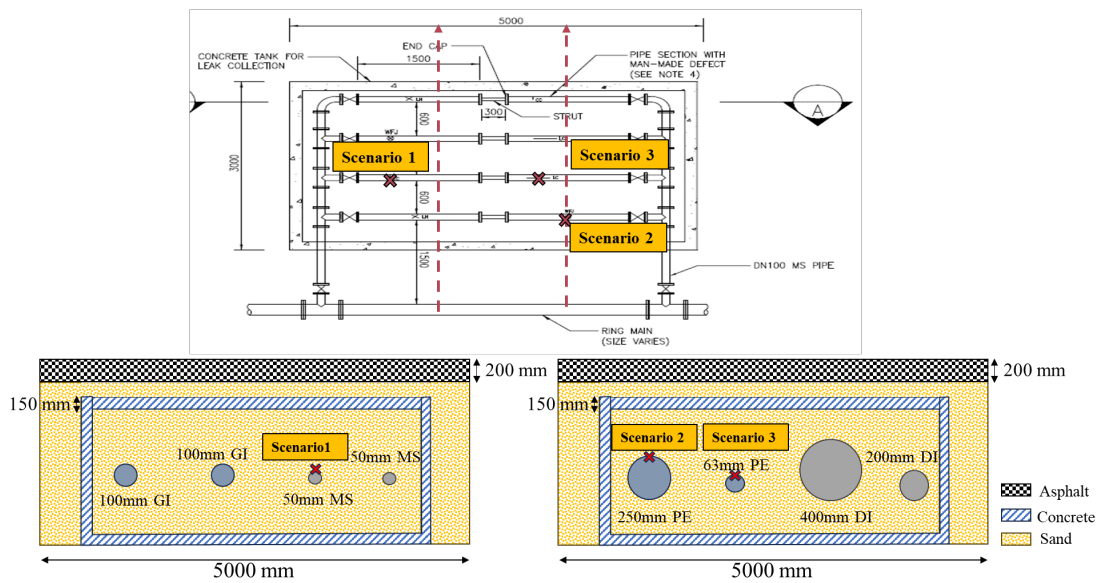


Figure 4-14 Sketch of the leakage pipes with marked leak points and vertical cross-section views of the experimental site

Data collection from the leakage scenarios was accomplished by performing an orthogonal grid survey (in Figure 4-15) with an IDS Hi-mode (operating at 600 MHz) antenna and a Leica DS 2000 (operating at 250 MHz and 700 MHz), both using a 50 ns time window to achieve a penetrating depth of approximately 2 meters. For each leakage scenario, two main stages were observed: the baseline case and the leakage case. Figure 4-15 outlines the survey grid overlaid on the leakage cell schematic, along with the timeline for leakage duration.

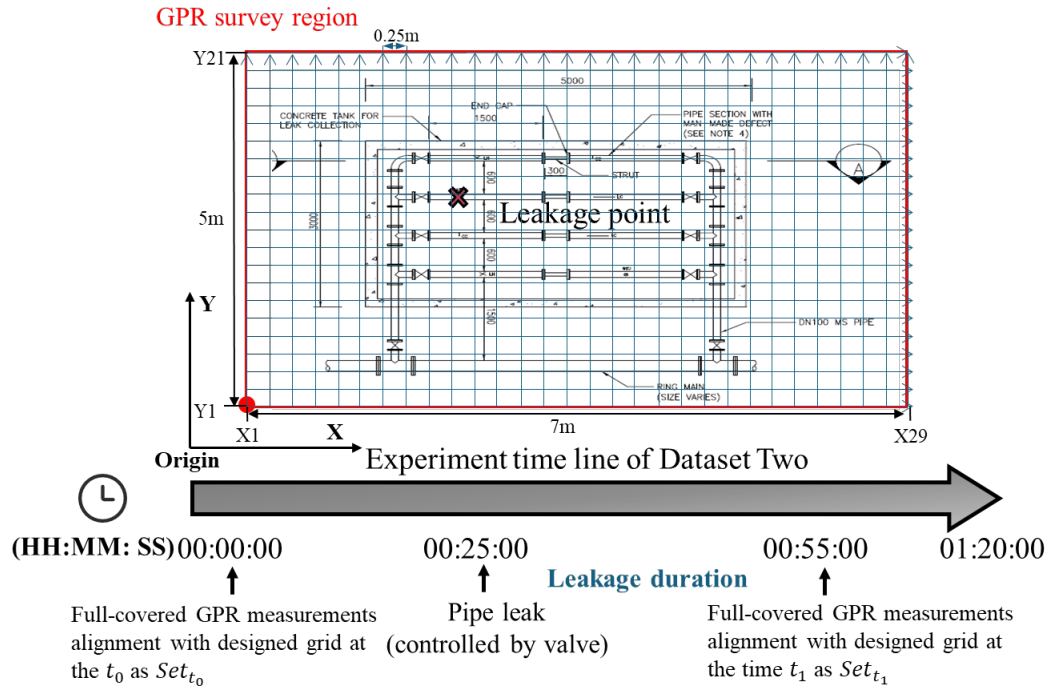


Figure 4-15 Experimental timeline for leakage scenarios and survey grid overlaid with the cell schematic (same as **Figure 4-6**)

Three distinct leakage scenarios from the database were utilized in this study, each varying based on the flow of leakage to simulate different degrees of severity, as detailed in Table 4-8. Each pipe, buried at different depths, was fitted with a gate valve for flow control. To emulate real-world conditions more accurately, asphalt was applied as a pavement material atop the cell. A drainage pipe was also installed at the bottom of the cell to manage water flow effectively.

Table 4-8 Detailed description of simulated leakage scenarios

Leakage scenario	Material hosting the leak pipe	Leak pipe size (mm)	Buried depth (m)	Flow of leakage (L/min)	Data acquisition time
Scenario 1	MS	100	1.110	1430	5 minutes after leakage
Scenario 2	PE	250	1.125	33	5 minutes after leakage
Scenario 3	PE	100	1.160	1281	20 minutes after leakage

As shown in Table 4-8, the leakage rates for Scenario 1 and Scenario 3 are 1430 L/min and 1281 L/min, respectively, both exceeding the drainage capacity. The leaking water

gradually filled the tank. Scenario 1 is characterized by a pipe burst, resulting in rapid and substantial water outflow that remains unobservable from the surface. Scenario 2 represents a minor leakage scenario, with a relatively low leakage rate. Scenario 3 involves the spread of leaked water over a larger area. Approximately 15 minutes after the leak point was opened, water was observed exiting through the manholes. Following this, the leak point was closed, water was allowed to drain partially, and data collection for Scenario 3 was conducted.

4.2.2.1 IF Outcomes and Leakage Signatures

Figure 4-16 illustrates how water leakage alters the GPR signal in a single IFS. The pair of plots on the left (Figure 4-16 (a) and (b)) depicts the IFS of the baseline Scenario (absence of leakage) and Scenario 1 (presence of leakage), respectively, within a specific depth range of 14-16 ns. The corresponding histograms, represented in red and blue, indicate the pixel values within the target and background regions (illustrated in Figure 4-11). Dashed lines denote baseline data, while solid lines represent Scenario 1 data, and vertical lines in each plot mark the centre of the histogram.

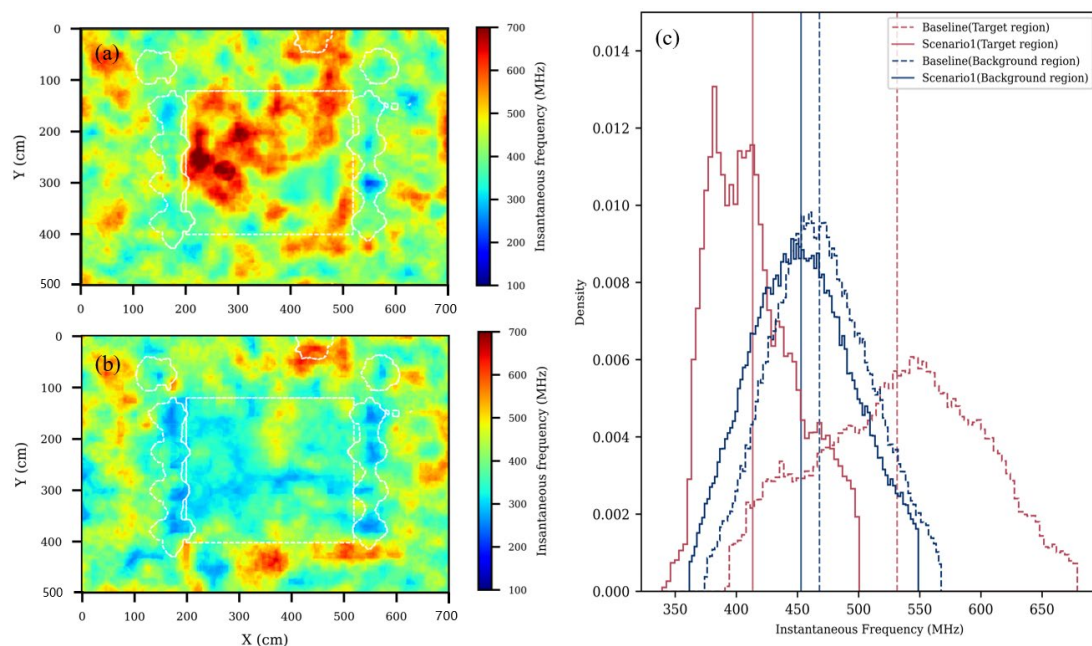


Figure 4-16 (a) IFS of baseline ; (b) Scenario 1 at depth 0.9 – 1.05 m; (c) Histograms of background and target regions

Figure 4-16 highlights the impact of water leakage on the IFS by comparing the baseline and leakage conditions (Scenario 1). A clear reduction in the high-frequency components of the signal is observed after leakage, particularly in the target region. The centre IF in this region shifts significantly from 535 MHz to 417 MHz, representing a decrease of 118 MHz. In contrast, the background region shows a minimal IF shift of only 10 MHz, demonstrating the stability of the signal outside the leakage-affected area.

Quantitative metrics further emphasize this distinction:

- (1). The IOU values between the baseline and Scenario 1 are 0.171 in the target region and 0.818 in the background region.
- (2). The Pearson correlation coefficients are -0.465 in the target region, indicating a negative correlation, and 0.935 in the background region, indicating a strong positive correlation.

These results collectively validate the effectiveness of the proposed method in detecting water leakage. In the target region, the large IF centre drift, low IOU, and negative Pearson correlation highlight the method's sensitivity to leakage-induced changes. In the background region, the minimal IF drift, high IOU, and strong positive correlation illustrate its robustness against noise and its ability to maintain stability in unaffected areas.

IFS provides a horizontal IF distribution within a certain depth range, condensing the information from a single IFS, which can be partial when characterizing the influence of water leakage. It is essential to consider how changes in water content will affect IFS vertically. To this end, a similar analysis was conducted from Slice1 to Slice8 (depth range from 0.675 m to 1.875 m with 0.15 m intervals). The results of this analysis are summarized in the following figures (Figure 4-17 to Figure 4-19). These figures are constructed using three distinct metrics derived from various IFS. In these figures, the values pertaining to the target region and the background region are represented by red and blue circles, respectively. The x-axis of the plot

displays the centre frequency drift, while the y-axis showcases slices at different depths. Differences in the centre drift between the target and background regions are illustrated using grey horizontal lines. The dimension of the circles in the plot corresponds to IOU values, ranging between 0 and 1. The colour gradient of these circles reveals the Pearson correlation coefficients. When the p-value of the Pearson correlation falls beneath the critical threshold of 0.05, the circles are accentuated in red. If the Pearson correlation turns out to be negative, the associated circle is outlined with a dashed line.

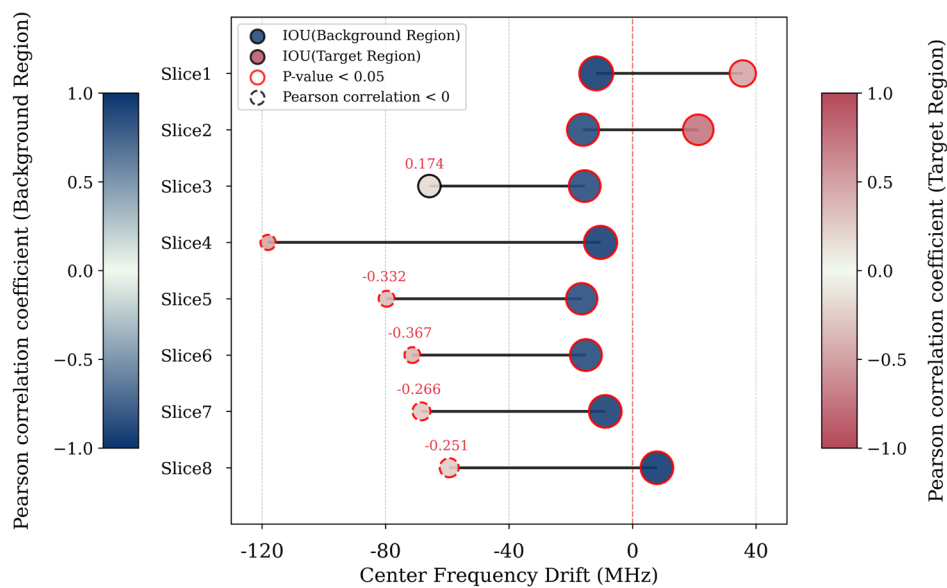


Figure 4-17 Depiction of vertical water leakage patterns in Scenario 1, as illustrated by variation in centre IF drift, IOU, and Pearson correlation coefficient across different IFS

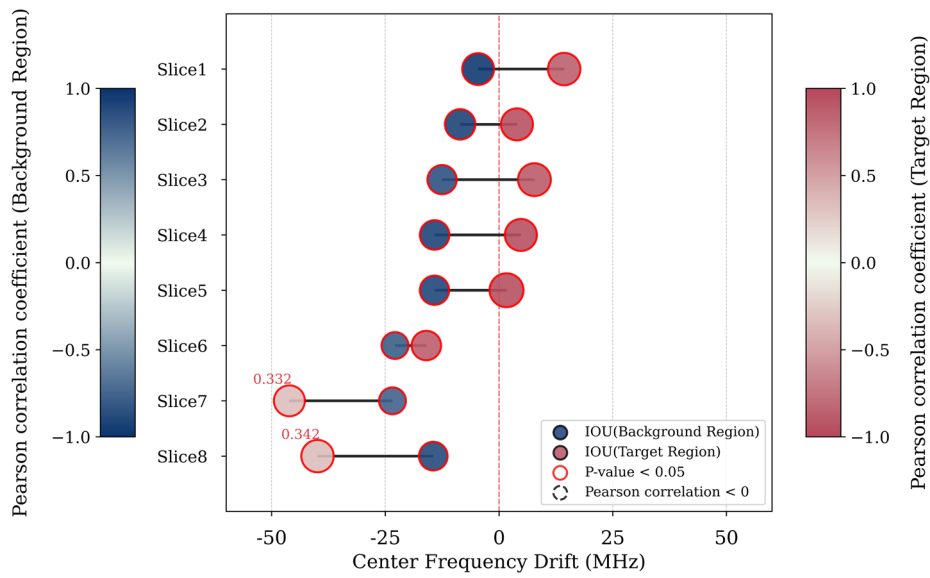


Figure 4-18 Depiction of vertical water leakage patterns in Scenario 2, as illustrated by variation in centre IF drift, IOU, and Pearson correlation coefficient across different IFS

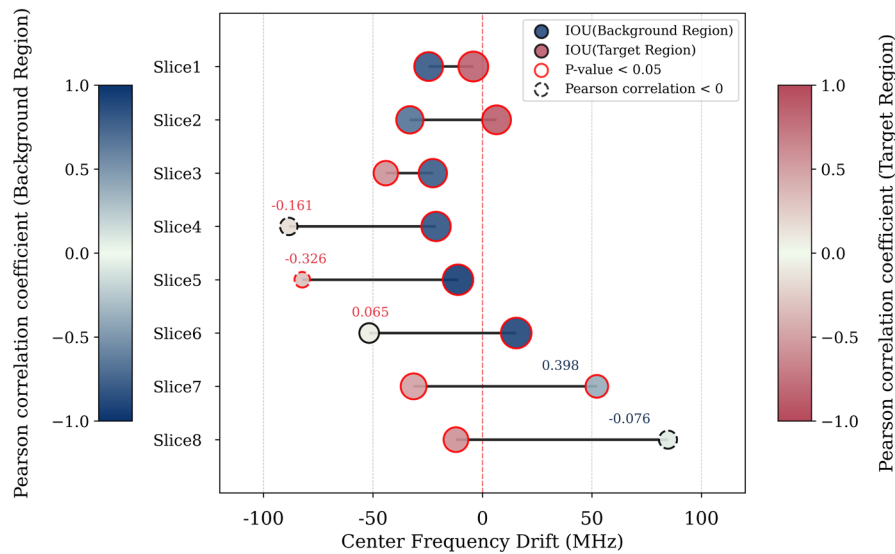


Figure 4-19 Depiction of vertical water leakage patterns in Scenario 3, as illustrated by variation in centre IF drift, IOU, and Pearson correlation coefficient across different IFS

The blue circles serve as anchors, representing environmental water content changes before and after leakage events. The vertical leakage patterns can be highlighted by comparing these three metrics in the target and background regions. In Scenarios 1 and 2, the blue circles are located around the 0 MHz line within the -25 MHz to 25 MHz range, suggesting less variation of centre IF drift in the background region of slices compared to the target regions.

The variations of red circles from shallow to deep slices clearly illustrate a pattern: at the interface from unsaturated to saturated zones, represented by the water table in some cases, the centre frequency drift transitions from positive to negative, along with relatively larger variations in IOU and PCC. In Scenario 1, the centre drift is positive for the first two slices, shifting to negative from slice 3 through slice 8, with the recognized water table within slice 3 (see Figure 4-10). Additionally, IOU and PCC indicate a similar trend, showing smaller circles and decreasing PCC below the water table. Conversely, in Scenario 2, the centre drift maintains a positive value from slice 1 to slice 5, then turns negative from slice 6 to slice 8. The leakage pipe is buried at a depth range of 1.125 m to 1.375 m, which corresponds to slices 4 and 5. Given that the leakage rate is lower than the drainage rate, the wetting zone should spread from the bottom of the leakage pipe. The results from the IFS accurately reflect these known conditions, highlighting the impact of leakage and drainage on the centre frequency drift and associated metrics. These observations suggest two mechanisms influencing the IF: reflection and absorption. At relatively shallow depths, unsaturated soil amplifies the signal's reflection. Due to the proximity of the energy source, the stacking signal experiences minimal attenuation. Increases in water content cause the IF to drift towards the antennas' fundamental frequency. As depth or free water content increases, signal absorption predominates, leading to the attenuation of the high-frequency component. This explains the negative centre IF drift observed in the red circles when the depth is below the interface of saturated and unsaturated zones. For Scenario 3, the centre IF drift lacks the stability seen in the previous two scenarios. This is attributed to the data being collected during drainage procedures, where water backflows and floods in the valve chambers and is dispersed across the site. Despite water seeping into the background region, a similar pattern of target region centre drift—moving from positive to negative with increasing depth—can be discerned. The IOU illustrates the similarity between pre- and post-event histograms at different regions. Comparatively, blue circles tend to be larger

than red circles at the same depth when the underground is significantly affected by water, as in Scenario 1 and Scenario 3. With increasing depth, the size decreases in Scenario 1. However, the IOU does not effectively reflect the underground water content in cases of minor leakage, such as Scenario 2. The Pearson correlation coefficient compensates for this, with a smaller PCC in slices 7 and 8 of the target regions compared to the shallower slices, at a high confidence level (P-value lower than the critical value). Additionally, the correlation coefficients display varying positive or negative correlations across different slices, serving as a potent indicator of water content changes.

In summary, the IF exhibits two primary patterns in response to underground water content changes. First, the IF tends to drift towards the antennas' fundamental frequency when soil moisture content increases but remains unsaturated, indicating reflection dominance. Second, the high-frequency components of IF experience significant attenuation when the soil is saturated and free water content increases, signifying absorption dominance.

4.2.2.2 Compare AS and IFS

AS are traditionally used to represent the subsurface by mapping out energy distribution. One of the major advantages of AS compared to IFS is its higher sensitivity to local interference caused by anomalies. These interferences manifest as oscillating signals, which are the local high-frequency components in the frequency domain and are often filtered out during IFS generation. However, several issues arise when dealing with multiple observations of the same region using AS.

The reflected amplitude operates within a dynamic range determined by the nature of the antennas. Most manufacturers calibrate their antennas' amplitude based on subsurface conditions to establish a suitable dynamic range. Consequently, the interpretation of amplitude signals is largely comparative because the dynamic range determines amplitude in the C-scan,

and this can obscure the interpretation of unregistered AS from different times. As seen in Figures 4-20 to 4-22, the axes display varying amplitude ranges, making it challenging to quantify the severity of leakage with fluctuating values. This characteristic complicates comparisons between observations taken at different times and those obtained from different antennas. One of the major challenges when dealing with time-series AS data is determining whether amplitude changes are derived from leakage or other anomalies. In contrast, IFS offers a promising method to characterize leakage patterns by anchoring water content variations in the time-frequency domain. IFS remains unaffected by the dynamic amplitude range, making it more reliable for time-series observations and water content changes. However, IFS is not sensitive to minor local leakages, which limits its ability to precisely locate leak points.

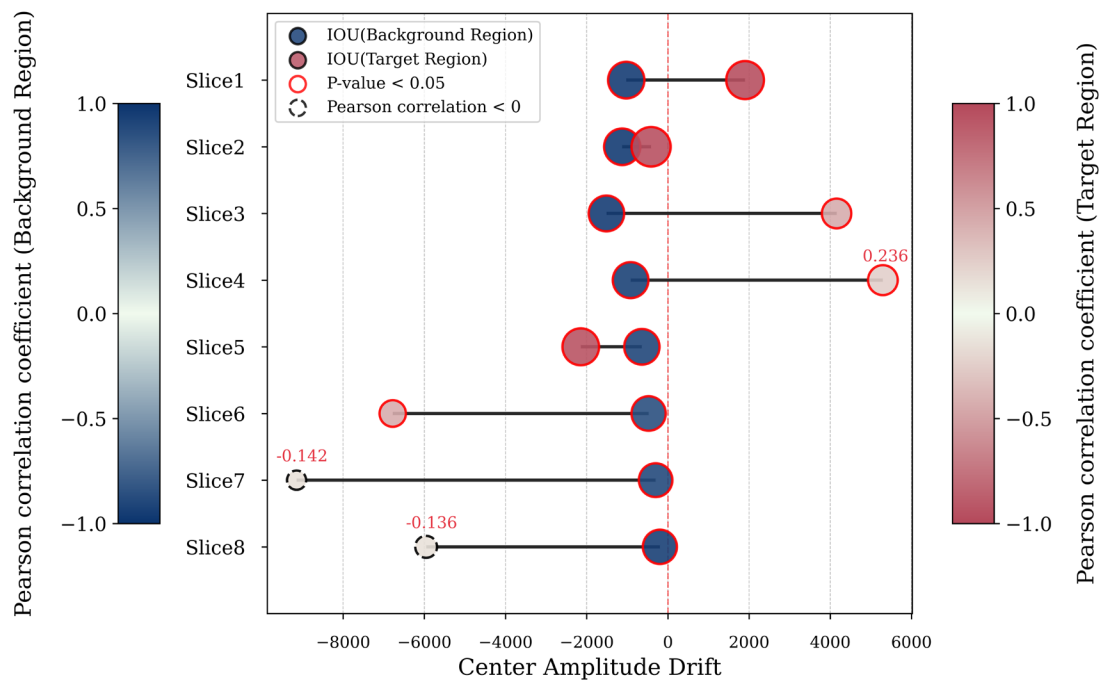


Figure 4-20 Depiction of vertical water leakage patterns in Scenario 1, as illustrated by variation in centre amplitude drift, IOU, and Pearson correlation coefficient across different AS

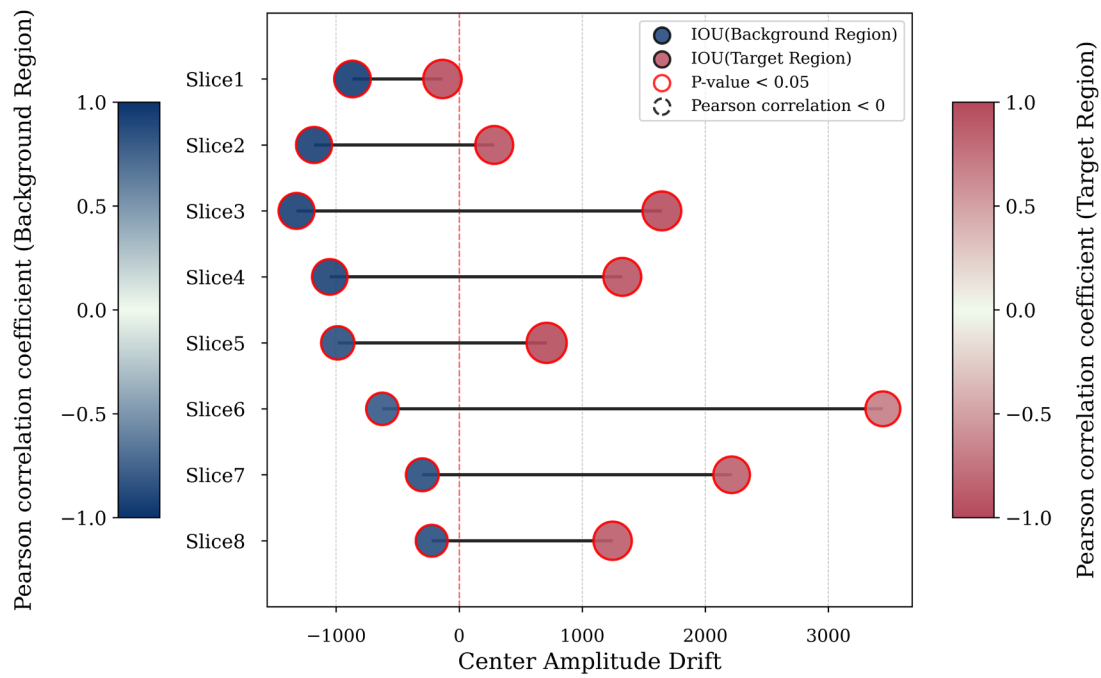


Figure 4-21 Depiction of vertical water leakage patterns in Scenario 2, as illustrated by variation in centre amplitude drift, IOU, and Pearson correlation coefficient across different AS

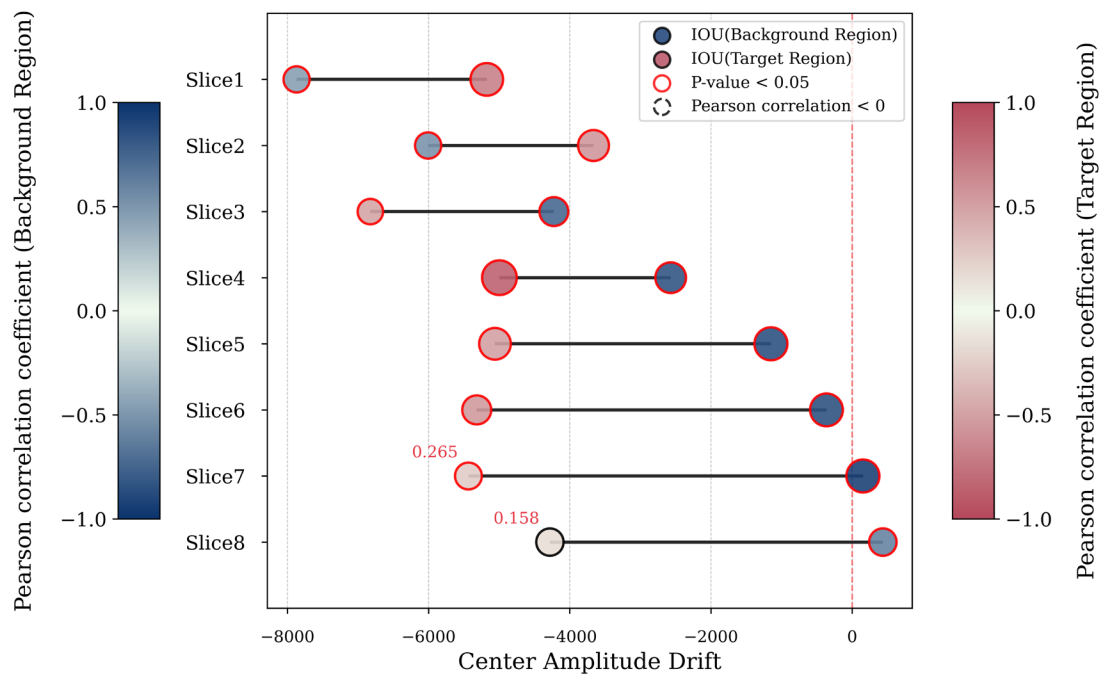


Figure 4-22 Depiction of vertical water leakage patterns in Scenario 3, as illustrated by variation in centre amplitude drift, IOU, and Pearson correlation coefficient across different AS

Another challenge with AS is the need for slice thickness adjustment. As discussed in the data analysis section, the slice thickness must be adjusted based on varying velocities at

different depths to ensure slices represent the same depth range vertically when there are significant differences in water content. This procedure corrects misalignment resulting from different velocities, allowing focus on changes derived from water content changes. Non-linear gain during AS generation can introduce artefacts to subsequent analysis.

In summary, while AS offers strong sensitivity to localized anomalies, its comparative and dynamic nature poses challenges for consistent analysis across time-series data. Proper adjustments for slice thickness and careful consideration of non-linear gain are essential to mitigate these issues. However, for comprehensive and consistent detection of water leakage, particularly in time-series observations, IFS remains the more reliable tool, despite its limitations in detecting minor leakages.

4.2.2.3 Fundamental Working Frequency Selection for IFS

To investigate whether various band-wise fundamental working frequency antennas can capture the water leakage patterns of buried pipes. Three fundamental frequencies' antennas (250 MHz, 600 MHz, and 700 MHz), which are the most commonly used frequencies in underground utility surveying, were used to conduct the leakage scenarios. Scenario 1 was chosen to verify how water content will affect the GPR signal. In order to narrow down our focus, we extract a region with the highest IF drift, as illustrated by the yellow dashed box in Figure 4-23.

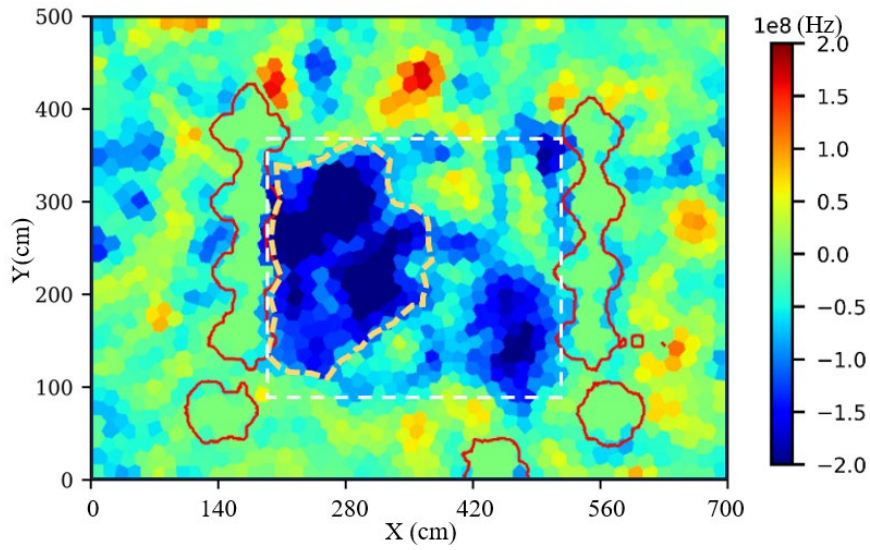


Figure 4-23 Difference image of IFS at a depth range from 1.215m to 1.365m in Scenario 1, highlighting areas significantly affected by water leakage (marked by yellow dashed line)

Figure 4-24 to Figure 4-26, present the average IF within the yellow region, accompanied by a ± 1 standard deviation buffer of the average IF curve, using 700MHz, 600MHz, and 250MHz antennas in Scenario 1. Corresponding amplitude curves are also illustrated with dashed lines. The accompanying three plots exhibit similar patterns. The highest IF drifts are observed within the depth range of Slices 3 or 4 in the results for both the 700 MHz and 600 MHz antennas. Additionally, the shallow slices maintain positive IF drift until the point of proximity to the water table depth.

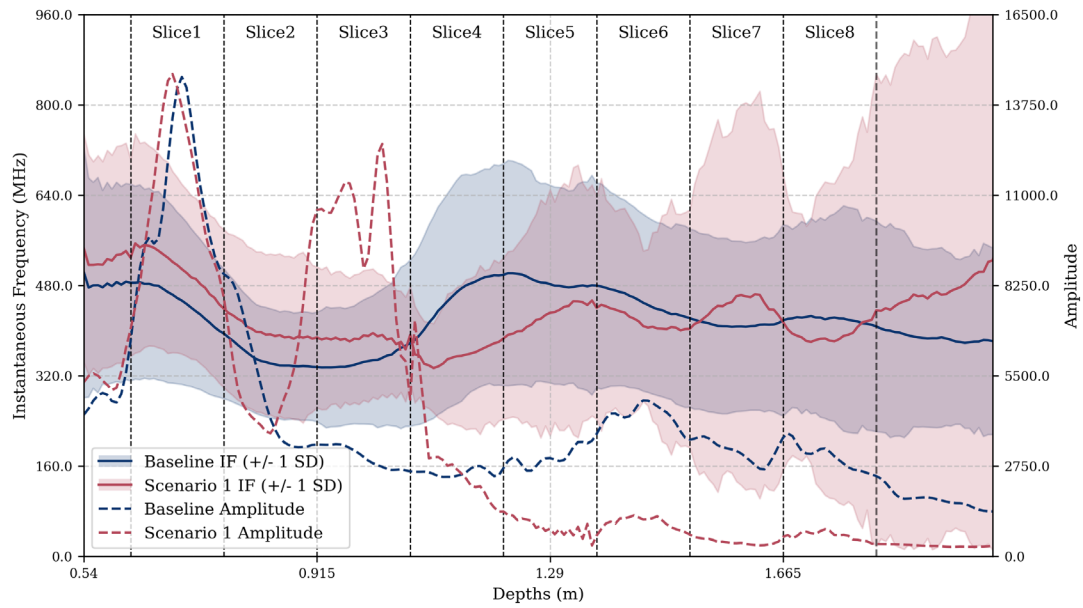


Figure 4-24 Average IF and amplitude curve of 700MHz antenna for areas significantly affected by water leakage in Scenario 1

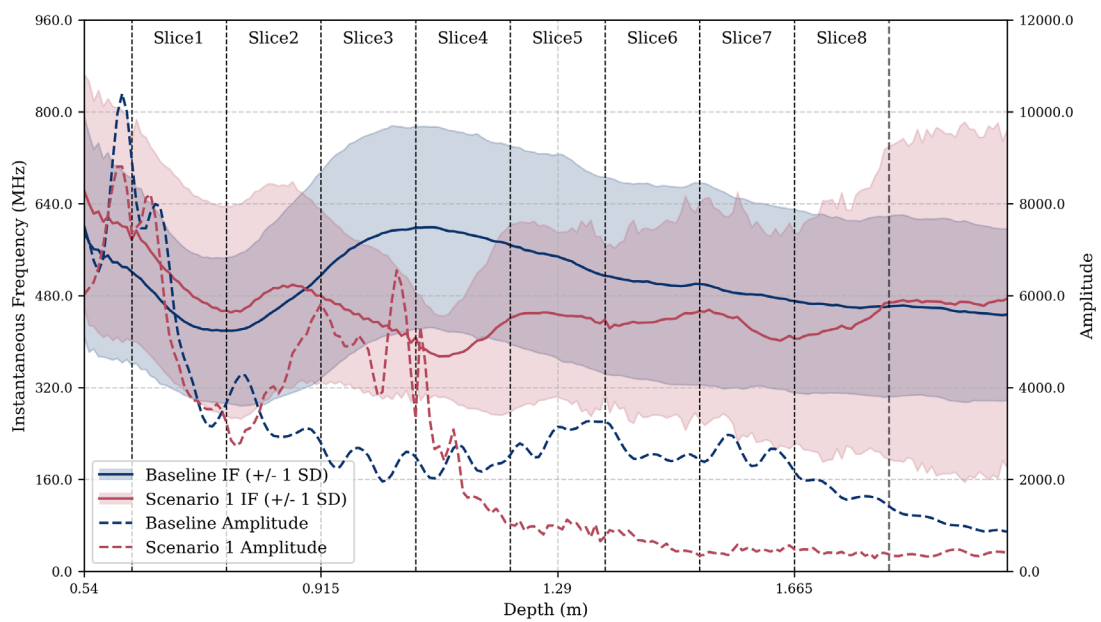


Figure 4-25 Average IF and amplitude curve of 600MHz antenna for areas significantly affected by water leakage in Scenario 1

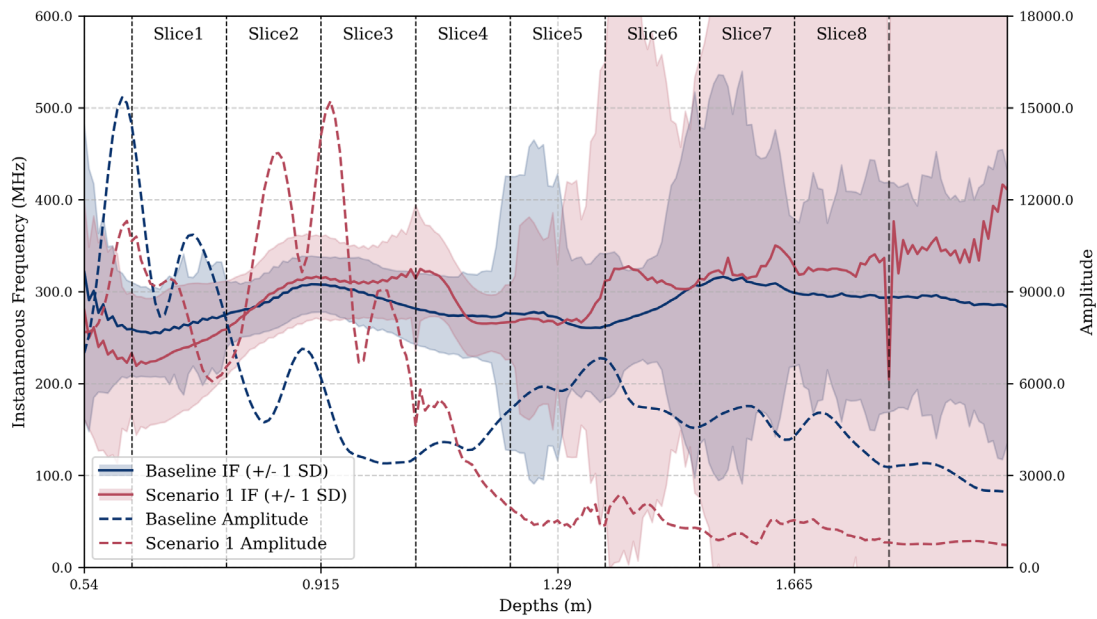


Figure 4-26 Average IF and amplitude curve of 250MHz antenna for areas significantly affected by water leakage in Scenario 1

Beyond the observations summarized in the previous plots, additional findings highlight key insights into the relationship between antenna performance and frequency-dependent water content characteristics. The most significant negative IF drifts are observed in the ranges of 500 MHz to 410 MHz for the 700 MHz antenna and 600 MHz to 390 MHz for the 600 MHz antenna, indicating substantial frequency shifts caused by water leakage. Variations across different frequencies also reveal important patterns: as water content increases, the signal becomes more chaotic and deviates beyond the buffer of ± 1 standard deviation of the IF curves.

The 600 MHz antenna demonstrates the most stable reflected signal, making it particularly effective for detecting leakage patterns. In contrast, the 250 MHz antenna struggles to capture the most responsive frequency range relevant to water content, which typically centres around 500 MHz in soil according to relaxation frequency studies (De Loor, 1983; Jol, 2009a). This mismatch results in the 250 MHz signal quickly devolving into erratic variations, rendering it suboptimal for detecting localized water leakages.

These mechanisms include clear reflections before approaching the water table and a significant absorption of energy when beneath it. Dynamic range variations were also observed across different antennas and observation times. Notably, the 250 MHz and 700 MHz antennas, which operate as dual channels of the same equipment, exhibit a more consistent dynamic range compared to the 600 MHz antenna, which belongs to a separate system. As introduced in Section 2.2, water content in dielectric materials such as soil, asphalt, and concrete exhibits distinct relaxation frequencies—approximately 500 MHz in soil, 1000 MHz in asphalt, and 1200 MHz in concrete. These relaxation frequencies enhance the interaction between electromagnetic waves and water, offering critical insights into leakage detection. Frequencies that align closely with these relaxation frequencies are more effective for identifying water leakage patterns using GPR.

In conclusion, selecting a fundamental frequency near the relaxation frequency—such as 500 MHz for soil—can significantly improve the efficacy of IFS for detecting water leakage. When specific background information is unavailable, a 500 MHz reference frequency serves as a practical starting point for general leakage detection.

In this section, the primary focus is placed on the characterization of leakage patterns through the use of controlled leakage experiments. Two observations are utilized: one conducted before the leakage and the other after. By comparing GPR data from these scenarios, IFS-based leakage patterns are summarized. In practical applications, obtaining pre-leakage data is often not feasible. The detection of leakage from a single observation is recognized as the most common scenario in real-world conditions. In complex and unknown underground environments, significant challenges are presented by the absence of baseline data, as horizontal and vertical leakage patterns derived from IFS may be obscured within the signals. The IF value, particularly when close to the relaxation frequency of water in the soil, is regarded as one of the most important indicators of water content. To improve detection accuracy, the survey

region can be clustered, and a relatively stable region can be identified to serve as the background region, allowing IFS values to be anchored. Although IFS-based leakage patterns demonstrate potential benefits for leakage detection, further validation is required to confirm the sensitivity and reliability of this method across diverse scenarios and environmental conditions. Such validation is essential for advancing the practical application of IFS in identifying water leakage.

4.3 Chapter Summary and Contribution

This chapter addressed Objective 1: Establish a Leakage Characterization Framework, presenting two key advancements in underground leakage detection research:

(1). Development of a Comprehensive Leakage Signature Database

A leakage signature database was established by simulating buried pipe leakages at the Q-Leak Center. This database comprises three distinct datasets, each designed to address specific research objectives and characterize leakage patterns under diverse conditions. By providing a structured repository of GPR signatures for leakage scenarios, this database serves as a foundational resource for understanding the variations in leakage patterns and supports future research in this domain.

(2). Characterization of Leakage Patterns Using IFS

A new representation, IFS, was proposed to improve the analysis of GPR data volumes. A detailed workflow was introduced to guide the transformation of raw GPR data into IFS. Using Dataset Two, the effectiveness of IFS was validated, demonstrating its capability to characterize underground leakage patterns in the frequency domain. This representation extends the feature space for identifying water leakage patterns, particularly in horizontal slices, offering a novel approach to subsurface analysis and enhancing the understanding of leakage behaviour.

These contributions address the objectives of establishing a database and characterizing leakage patterns, advancing both the methodological framework and the practical tools for underground leakage detection. The results provide a strong foundation for further research and potential applications in real-world scenarios.

Chapter 5. Underground Hazards Detection using Temporal GPR Data

As discussed in Chapter 2, the strategy of multiple observations combined with change detection methods shows significant potential for identifying underground hazards from a dynamic perspective.

In Section 5.1, an unsupervised void recognition method based on change detection was introduced and validated through laboratory experiments. The workflow begins with an intensity registration module, which aligns one or multiple pairs of C-scans for consistent comparison. Next, superpixels are segmented from the image differences and used as the basic processing unit throughout the change detection workflow. This segmentation reduces the effects of noise and filters out minor and insignificant changes. At last, the change detection process combines fuzzy c-means clustering (FCM) and the Markov random field (MRF) algorithm, both adapted for C-scan analysis. FCM is employed to classify superpixel-based differences in the images, which are generated using the simple linear iterative clustering (SLIC) method, into initial change maps. Unlike traditional hard classification methods that label areas as strictly "changed" or "unchanged," this approach represents changes using probabilities, creating "soft" relationships. These probabilities offer a more nuanced representation of changes, which are particularly effective when integrated into the MRF model. The developed MRF model incorporates both spectral and spatial components, leveraging intensity information and the spatial relationships among neighbouring superpixels across all three dimensions. This integration enhances the robustness and accuracy of the change detection process.

In Section 5.2, a leakage detection method is proposed by combining amplitude and IF features. This method consists of two stages: Firstly, the 3D FCM algorithm is applied to

identify regions with significant amplitude changes in amplitude slices, generating an initial change map where areas with high amplitude are highlighted. This approach extends the clustering process into three dimensions, ensuring spatial continuity across adjacent slices. Additionally, the probabilistic nature of FCM allows for flexibility in adjusting the sensitivity of the detection by modifying the probability threshold. Secondly, the IFS-based change detection method is then applied to refine the results by exploiting the unique characteristics of water leakage (as discussed in Section 4.2). This step identifies potential leakage regions among the areas initially flagged as high-amplitude changes, leveraging the distinct signal patterns associated with leaks.

5.1 Underground Void Detection based on GPR Time-lapse Slices Integrating Fuzzy C-Means and Markov Random Field Method

In the past two decades, various change-detection methods for images have been proposed in the remote sensing field. These methods can be classified according to their scale, including pixel-level (Chi et al., 2007; Rokni et al., 2015; Stow, 1999), object-level (Desclée et al., 2006; Simanjuntak et al., 2019; Walter, 2004), feature-level (Lv et al., 2018; Seo et al., 2022; Zhang & Shi, 2020), and hybrid-level (Hussain et al., 2013) methods. In the GPR field, it has been suggested that a single measurement is not reliable for determining small changes due to its pseudo-grey scale nature (Luo & Lai, 2020b). Instead, establishing a time series of time-lapse GPR C-scans has been proposed as an alternative for identifying potential subsurface defects. By comparing C-scans with a baseline, regions of change can be extracted, which is particularly useful for detecting underground hazards. Since underground hazards tend to develop over time, change detection offers an effective strategy to identify potential hazards and narrow down regions with high risk. This significantly reduces inspection and maintenance

workloads.

However, despite its potential, research on change detection using GPR slices has been limited due to several inherent challenges (Luo & Lai, 2020b). These challenges are particularly pronounced with C-scans, which provide horizontal representations of the underground and enable the detection of changes over large areas. There are three main challenges:

- (1). Accurate image registration is crucial for aligning multiple surveys in temporal GPR data, as GPR signals are highly sensitive to environmental variations (e.g., moisture conditions, temperature, and antenna energy), which introduce noise and affect the consistency of measurements.
- (2). Pattern recognition methods must be specifically adapted for C-scans because they are generated through interpolation, which compresses vertical information referring to the signal propagation into horizontal intensity distributions, causing a loss of wave patterns and introducing artificial uncertainty due to interpolation parameters.
- (3). Integrating neighbourhood information is essential because C-scans compress continuous underground space into horizontal slices, often obscuring changes occurring in the vertical direction that can only be captured by incorporating three-dimensional spatial relationships from adjacent slices.

Despite these challenges, the proposed methods in this study address these issues. Image registration mitigates noise introduced by environmental variations, while tailored pattern recognition techniques effectively extract meaningful changes from C-scan data. Additionally, integrating neighbourhood information ensures that vertical changes are captured, enabling a more comprehensive and robust change detection process.

5.1.1 Underground Void Detection Methodology

Figure 5-1 shows the flowchart of the proposed superpixel-based change detection

method combining FCM and MRF. The method encompasses two main sub-modules: time-series C-scan generation and C-scan change detection.

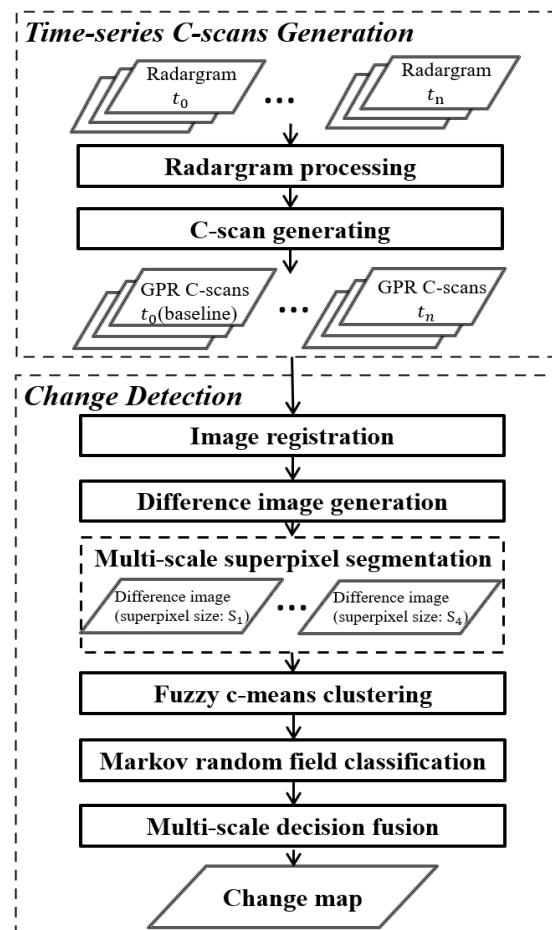


Figure 5-1 Workflow of the proposed superpixel-based change detection method

Radargrams collected at any two different times are processed by basic 2D signal processing functions automatically in-house using Python programs. These functions include de-wow to remove low-frequency noise, direct-current shift to correct the shift in the waveform, static correction to adjust time zero, range gain for consistent amplitude contrast, bandpass, background removal, and migration (Jol, 2009a). Subsequently, C-scans were generated using GPR Slice, a third-party commercial GPR imaging software package (Goodman et al., 2013). The B-scans were stacked into C-scans based on a predesigned survey grid. Thus, each GPR signal was registered using local coordinates, and the profile spacing was less than 3λ (λ

refers to the wavelength of the EM wave) following the 3D GPR imaging criteria in Figure 5-2. To generate the C-scans, we specified three main parameters: profile spacing (PS), slice thickness (ST), and search radius (SR), and used bi-linear interpolation for filling in any missing data points. The profile spacing was set to 10cm less than 3λ , where λ is the wavelength of the GPR signal, so that the measurement effort and data resolution can be leveraged. The ST was set as 0.9λ (15cm) for sufficient resolution in the vertical direction. To interpolate missing data points, we used bi-linear interpolation with a search radius of 0.9λ (15cm), which is recommended to be larger than the profile spacing to involve information from adjacent survey lines.

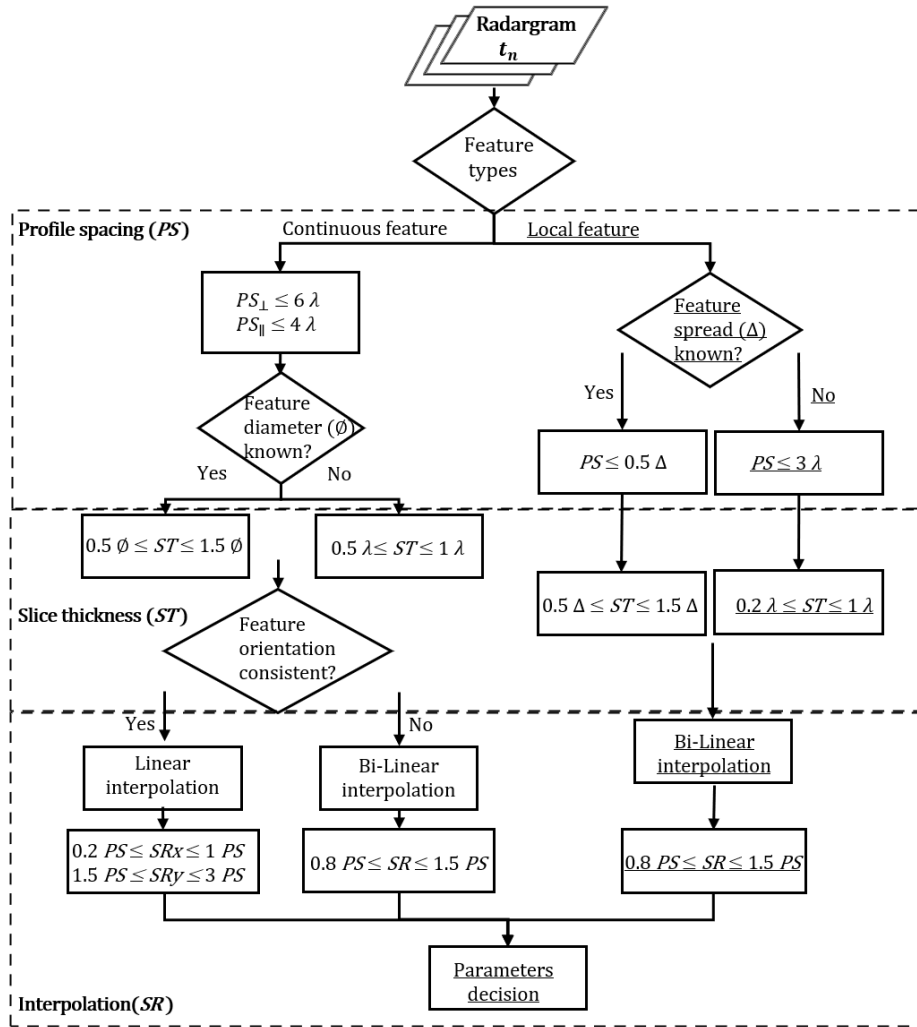


Figure 5-2 3D GPR imaging workflow. *PS*: profile spacing; *ST*: slice thickness; *SR*: search radius; λ : wavelength of the EM wave propagation underground (The underlined marks in the figure denote the parameter choices used in our method) (Luo et al., 2019b)

The reflected signals and locations in GPR surveys are subject to multiple factors that may introduce variations in the GPR signals collected over time. To address this issue, image registration is typically employed as the first step in adjusting a pair of C-scans to a similar intensity range. Intensity-based image registration estimates the transformation that aligns target images to baseline images based on their intensity values. This task can be formulated as an optimization problem with four primary components: transformation type, similarity metric function, interpolator, and optimizer. In the present study, we implemented affine transformation, correlation, linear interpolation, and gradient descent optimization in C-scan intensity registration. The to-be-registered image is transformed to match the reference image in the registration procedure, and the optimizer adjusts the affine transformation matrix iteratively based on the gradient descent direction until the correlation between the two images is maximized. To achieve C-scan image intensity registration, we employed the SimpleITK registration framework (Lowe et al., 2013), which provides stable performance and can be used to effectively design each component using multiple programming language packages.

After applying the C-scan image registration module, the pair of C-scans was aligned and corrected to comparable intensity ranges. In this study, superpixels—homogeneous blocks of pixels that share similar intensity values within a certain distance—were used as the basic processing unit. This approach was adopted for two key reasons: first, it helps to reduce local random noise; second, it scales up the unit of analysis, allowing the focus to remain on changes that exceed a given size threshold. To generate the superpixel-based difference image (DI), two methods were employed: contrast limited adaptive histogram equalization (CLAHE) to normalize the intensity values and the simple linear iterative clustering (SLIC) algorithm to

segment the image into homogeneous superpixels (Achanta et al., 2012; Reza, 2004). The DI was then calculated using the following formula:

$$S_D^a = |CLAHE(S_{t_x}^a) - CLAHE(S_{t_0}^a)|, \quad [5 - 1]$$

where $S_{t_0}^a$ and $S_{t_x}^a$ are a^{th} registered C-scans generated by the data acquired in t_0 and t_x , where a is the index of the C-scans ($1 < a < A$). S_D^a denotes the DI of a pair of a^{th} registered C-scans after CLAHE enhancement.

In this study, the SLIC method was used to segment DI into superpixel-based DI, considering each superpixel as a homogeneous block with similar image features. Specifically, SLIC adapts k -means clustering with limited search regions to generate superpixels iteratively following intensity and spatial proximity. The intensity and spatial proximity with intensity distance d_I and spatial distance d_S between two pixels i, j is written as follows:

$$\begin{cases} D = \sqrt{\left(\frac{d_I}{M_I}\right)^2 + \left(\frac{d_S}{M_S}\right)^2} m \\ M_S = S = \sqrt{N_P/K} \\ M_I = I_{max} - I_{min} \end{cases}, \quad [5 - 2]$$

$$\begin{cases} d_I = \sqrt{(I_i - I_j)^2} \\ d_S = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \end{cases}, \quad [5 - 3]$$

where distance D was calculated by d_I and d_S , and M_I and M_S represent the maximum intensity distance and maximum spatial distance, respectively. m is a relative factor to weigh the intensity and spatial similarity. I_i, x_i , and y_i refer to the intensity value and the position coordinates of the i^{th} pixel, respectively. The maximum spatial distance within a given cluster is related to the sample interval, $M_S = S = \sqrt{N_P/K}$, in which N_P is the number of pixels in a whole C-scan, and K is the number of superpixels. The maximum spectral distance is the difference between the maximum and minimum intensity values of the pixels in the given cluster. The scale has been stressed as a key parameter for discrimination (Inglada & Mercier,

2007) between changed and unchanged areas. A multiscale strategy is implemented to reduce misdetection or detections by generating difference images with varying sizes of superpixels

$$S_n = \left\{ \frac{S}{n} \mid n = 1, 2, 3, 4 \right\}.$$

The superpixel-based DI was generated using the previous equations (Equations [5-2] and [5-3]); hence, superpixels of the DI were used instead of pixels as the basic processing unit for the rest of the modules. Given a set with K superpixels, $R_a^{S_n} = \{r_{k_a}^{S_n} \mid k = 1 \dots K, n = 1, 2, 3, 4\}$ represents a superpixel-based DI generated by a pair of a^{th} C-scans with superpixel size S_n , where $I_{r_{k_a}^{S_n}}$ is the intensity of the superpixel $r_{k_a}^{S_n}$ and the intensity of the superpixel was obtained by the mean value of the pixel-based DI within the cluster region. Figure 5-3 provides an example of our superpixel segmentation approach. The 'target size', a user-specified parameter, sets the size of level-1 superpixels. The 'target size' is meticulously tailored to correspond to the scale of the features targeted for detection and analysis. Consequently, a larger 'target size' may potentially underemphasize minor variations within the image, while a smaller 'target size' may accentuate these nuances. Level-2 superpixels are one-fourth of the target size, while level-3 and level-4 superpixels are one-ninth and one-sixteenth of the target size, respectively (Luo & Lai, 2020).

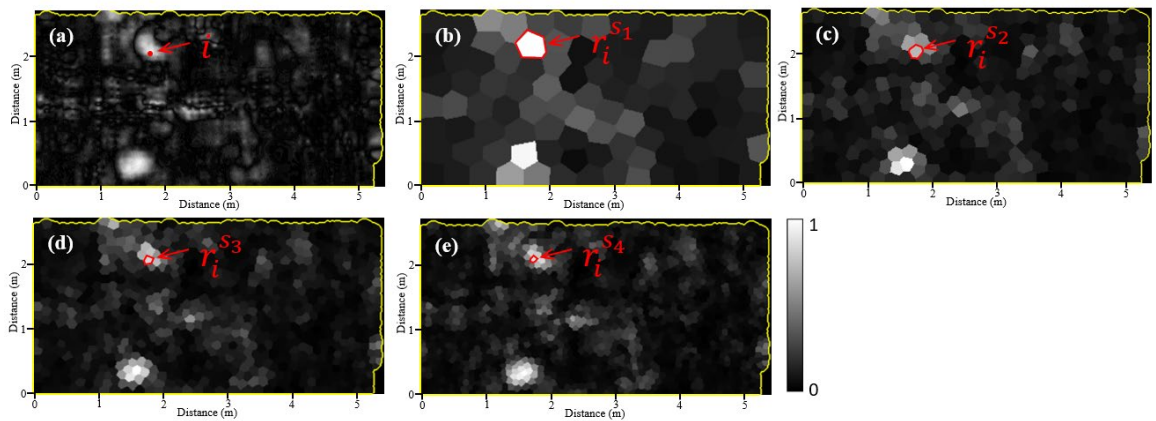


Figure 5-3 Illustrations of pixel-based DI map and multiscale DI maps (a) pixel-based DI map; superpixel-based DI maps: (b) level-1; (c) level-2; (d) level-3; (e) level-4

FCM is a widely used clustering algorithm (Bezdek et al., 1984). The change detection task can be regarded as a binary classification, and the clustering algorithm divides the DI into changed and unchanged regions. FCM, combined with fuzzy set theory, uses the possibility of change occurrence instead of a hard boundary between changed and unchanged states, as depicted in Figure 5-4.

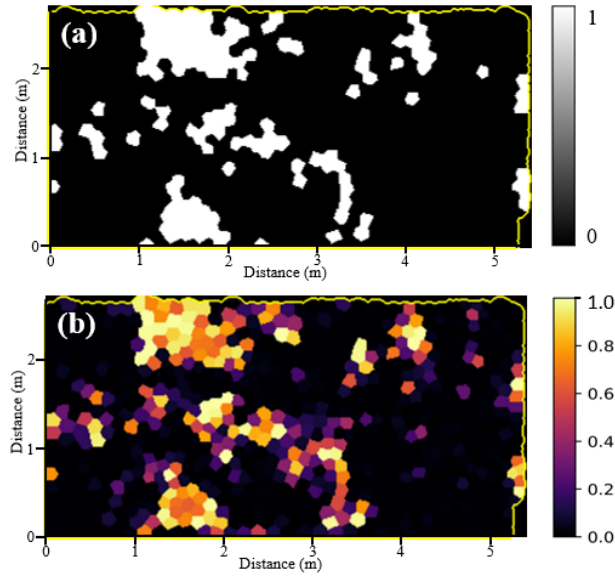


Figure 5-4 Soft label clustering illustration: (a) hard classification: 0 indicates unchanged and 1 indicates changed; (b) soft classification: the values refer to the probability of change

Given C classes, the DI of a^{th} C-scan with level-n superpixel size is represented by $R_a^{S_n}$. For simplicity and clarity, they are represented by $R = \{r_k \mid k = 1 \dots K\}$ in this subsection. FCM aims to minimize the objective function by assigning each r_k to different classes and corresponding memberships iteratively, as follows:

$$J = \sum_{k=1}^K \sum_{c=1}^C u_{kc}^q \|I_{r_k} - v_c\|^2, \quad [5-4]$$

$$\begin{cases} u_{kc} = \frac{\sum_{h=1}^C \left(\frac{\|I_{r_k} - v_c\|}{\|I_{r_k} - v_h\|} \right)^2}{\sum_{h=1}^C \left(\frac{\|I_{r_k} - v_c\|}{\|I_{r_k} - v_h\|} \right)^2} \\ v_c = \frac{\sum_{k=1}^K u_{kc}^q \cdot I_{r_k}}{\sum_{k=1}^K u_{kc}^q} \end{cases}, \quad [5-5]$$

where K is the number of superpixels, C is the number of classes and equals 2, q is the weighting exponent of FCM and equals 2, u_{kc} is the partition of the k^{th} superpixel in the cluster C of the R , and v_c is the c^{th} cluster center. We randomly assigned initial membership and cluster centres. To ensure convergence of the algorithm, we set two termination conditions. The first was based on the convergence of the objective function. We terminated the algorithm when the decrease in the objective function between two consecutive iterations was below a threshold of 0.001. The second termination condition was the maximum iteration number, which we set to 500.

After clustering, the possibility of each superpixel belonging to the changed cluster was normalized to $[0,1]$, according to DI values. In addition, the initial single C-scan change map was generated by assigning the membership of each superpixel.

C-scans are present underground through a series of horizontal images based on depth range. Although changes may exhibit clear patterns, vertical information should also be considered. To obtain effective neighborhood information in three dimensions, the superpixel neighborhoods were defined as an adjacent set $R_{aj}^{S_n} = \{r_{kaj}^{S_n} \mid j = 1 \dots J\}$, as illustrated in Figure 5-5.

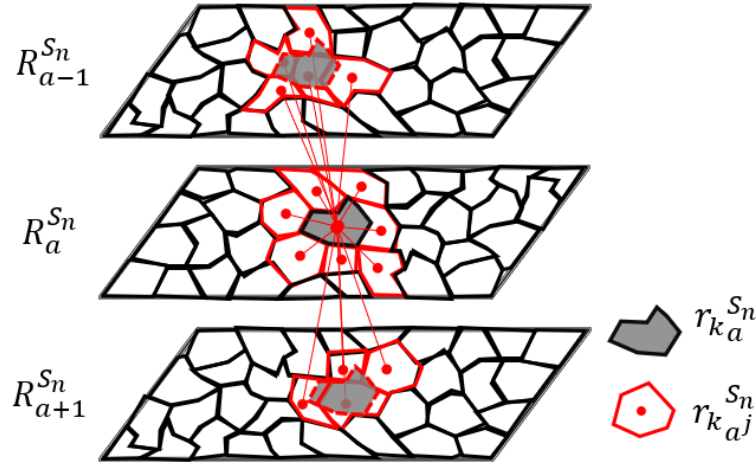


Figure 5-5 Superpixel-adjacent neighbourhoods in three dimensions

Suppose that $L = \{l_{r_{ka}^{S_n}} \mid k = 1 \dots K, n = 1, 2, 3, 4\}$ represents the label sets of the superpixels, in which $l_{r_{ka}^{S_n}}$ is the label of superpixel $r_{ka}^{S_n}$, which is equal to 0 or 1 (changed or unchanged, respectively). Let $X = \{r_{ka}^{S_n} \mid k = 1 \dots K, n = 1, 2, 3, 4\}$ denote the C-scan set composed of K superpixels. The maximum a posteriori estimation is used to obtain the labels of a superpixel, namely

$$L = \arg \max \{P(L)P(X|L)\}, \quad [5 - 6]$$

where $P(X|L)$ is the joint conditional probability distribution of the superpixel set X and $P(L)$ is the prior probability of the class label L . Under the MRF approach, the MAP is equivalent to maximizing the following energy function U_{MRF} with the spectral and spatial components from multi-scaled DI images:

$$U_{MRF} = \sum_{n=1}^N U_{MRF}^{S_n} \quad (n = 1, 2, 3, 4), \quad [5 - 7]$$

$$U_{MRF}^{S_n} = U_{spectral}^{S_n} + \beta * U_{spatial}^{S_n}, \quad [5 - 8]$$

$$\begin{cases} U_{spectral}^{S_n}(r_{ka}^{S_n}) = \frac{1}{2} \ln(2\pi\sigma_l^{S_n}) + \frac{1}{2} (I_{r_{ka}^{S_n}} - \mu_l^{S_n})^2 (\sigma_l^{S_n})^{-1} \\ U_{spatial}^{S_n}(r_{ka}^{S_n}) = \sum_{j=1}^J I(l_{r_{ka}^{S_n}}, l_{r_{ka^j}^{S_n}}) \end{cases}, \quad [5-9]$$

where μ_l and σ_l are the mean and variance intensity of superpixels belonging to the label l , $I_{r_{ka}^{S_n}}$ is the intensity of superpixel $r_{ka}^{S_n}$, and β is the relative weight parameter between spectral energy and spatial energy.

The spatial energy term can be described in detail as follows:

$$I(l_{r_{ka}^{S_n}}, l_{r_{ka^j}^{S_n}}) = \begin{cases} w_{kj}, & l_{r_{ka}^{S_n}} = l_{r_{ka^j}^{S_n}} = 1 \\ 0, & l_{r_{ka}^{S_n}} \neq l_{r_{ka^j}^{S_n}} \end{cases}, \quad [5-10]$$

$$w_{kj} = l(f_k) \times l(f_k^j) \times \frac{1}{R(r_{ka}^{S_n}, r_{ka^j}^{S_n})}, \quad [5-11]$$

$$R(r_{ka}^{S_n}, r_{ka^j}^{S_n}) = \begin{cases} \sqrt{(x_{r_{ka}^{S_n}} - x_{r_{ka^j}^{S_n}})^2 + (y_{r_{ka}^{S_n}} - y_{r_{ka^j}^{S_n}})^2}, & A(r_{ka}^{S_n}) = A(r_{ka^j}^{S_n}) \\ \sqrt{(x_{r_{ka}^{S_n}} - x_{r_{ka^j}^{S_n}})^2 + (y_{r_{ka}^{S_n}} - y_{r_{ka^j}^{S_n}})^2 + ST^2}, & A(r_{ka}^{S_n}) \neq A(r_{ka^j}^{S_n}) \end{cases} \quad [5-12]$$

where w_{kj} denotes the spatial attraction between $r_{ka}^{S_n}$ and its adjacent neighbourhoods $r_{ka^j}^{S_n}$.

$l(f_k)$ are the members belonging to class l of the superpixel $r_{ka}^{S_n}$ from $u_{ka}^{S_n}$ of FCM

results, $R(r_{ka}^{S_n}, r_{ka^j}^{S_n})$ denotes the Euclidean distance between the centroids of superpixel $r_{ka}^{S_n}$

and $r_{ka^j}^{S_n}$. Equation [5-12] is to calculate the Euclidean distance, taking into account whether

they belong to the same depth or not. $A(r_{ka}^{S_n})$ denotes the index of the C-scan, which $r_{ka}^{S_n}$ is

belonging to, and ST refers to the slicing thickness from the C-scan generation function, dealing

with $r_{ka}^{S_n}$ and $r_{ka^j}^{S_n}$ belonging to different C-scans. The iteration condition model was used

in this study to solve Equation [5-7] (Zhu & Liu, 2013). The final change map is obtained when

the iterative termination condition is met or the maximum iteration number is reached.

To quantify the accuracy of the proposed superpixel-based change detection method in the designed situations. Several commonly used indices, TP (true positive), FP (false positive), TN (true negative), FN (false negative), F1-score, and the kappa coefficient, were selected to estimate the results. TP means the correctly detected changed pixels that contain voids, FP means incorrectly detected changed pixels that don't contain voids, TN means correctly detected unchanged pixels that don't contain voids, and FN means incorrectly detected unchanged pixels that contain voids. The F1-score and kappa coefficient were calculated as follows:

$$F1 - score = \frac{2 * (precision * recall)}{(precision + recall)}, \quad [5 - 13]$$

$$\begin{cases} precision = \frac{TP}{TP + FP} \\ recall = \frac{TP}{TP + FN} \end{cases}, \quad [5 - 14]$$

$$Kappa\ coefficient = \frac{2 \times (TP \times TN - FN \times FP)}{(TP + FP) \times (FP + TN) + (TP + FN) \times (FN + TN)}. \quad [5 - 15]$$

5.1.2 Experiment Results and Discussion

Subsurface hazards can emerge suddenly and unexpectedly, leading to serious damage and safety risks, such as road collapses. Despite the potential benefits of GPR-image-based disaster monitoring, the lack of validation data poses a significant challenge for method development. To overcome this limitation, it is essential to have an experimental site that can simulate various underground conditions. The utility laboratory at The Hong Kong Polytechnic University provides a manipulable test bench, which allows for the observation of time-series data under controlled underground conditions. As depicted in Figure 5-6 (a) and (b), the entire experimental site (3 m × 5 m) has a removable cover. It also includes pipes of different materials and manholes in various orientations, making up a complex underground scenario. Utilities of

different sizes were buried under soil with covers on top, allowing the creation of diverse underground situations. Figure 5-6 (b) also shows IDS Hi-Mod used in this experiment, following the grid shown in Figure 5-6 (c) with orthogonal grid cells at 10 cm intervals. Different kinds of targets, including spherical fitness balls and peanut-shaped fitness balls, were put into the soil to simulate different shapes of underground cavities; see Figure 5-6 (d), (e), and (f).

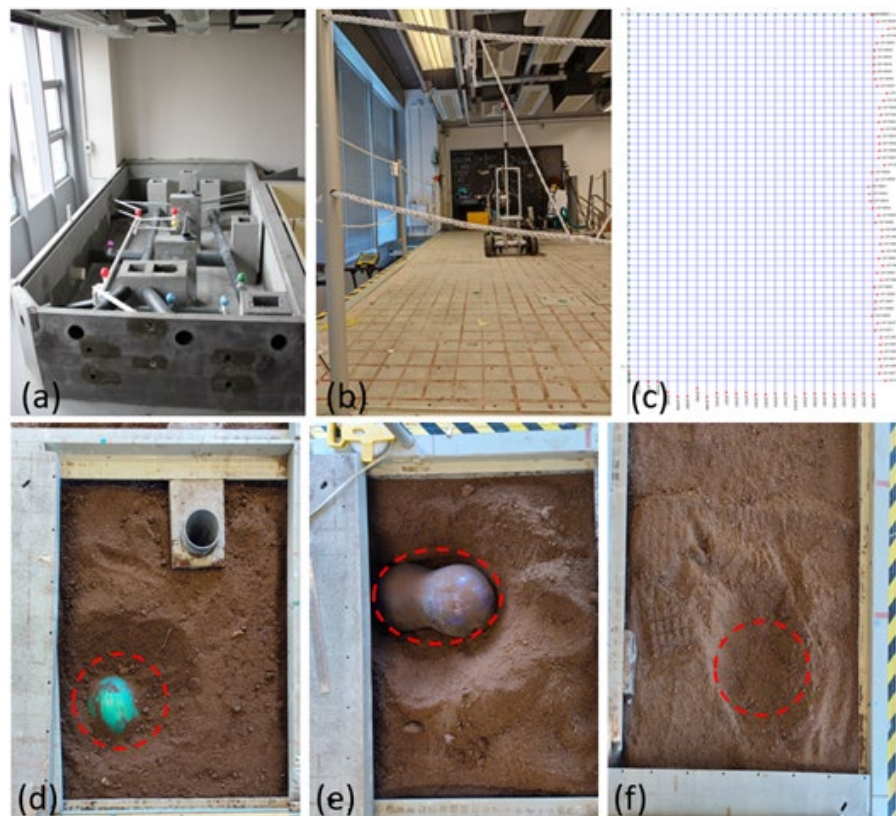
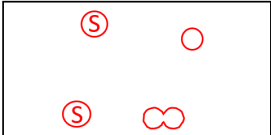
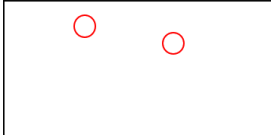
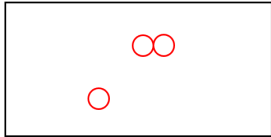
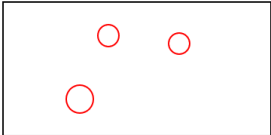


Figure 5-6 Photos of the experiment site and targets. (a) the pipe and manhole distribution at the test site; (b) the test site and the IDS hi-mod antenna; (c) surveying grid of the test site; (d) a buried spherical fitness ball; (e) a buried peanut-shaped ball; (f) a void between pavement and soil

In this study, we designed four scenarios to evaluate the performance of the GPR system in detecting underground voids, as summarized in Table 5-1. We employed an auto-tracked position system to track survey lines on an orthogonal grid, collecting data using a 600 MHz antenna at 1 cm sampling intervals. The experimental design aimed to simulate the formation of underground voids in practical situations. Specifically, we placed the buried balls

near the pipes, mimicking the scenario where water leaks from pipes and causes soil erosion to form voids. We also designed an isolated void to replicate the situation where the heavy pressure of the road results in soil loss and cavity formation. To verify the boundary of effective detection, balls of various sizes were utilized. Additionally, we used peanut-shaped balls to create voids with multiple centres. To imitate the time-series environment of void formation and remediation, each of the four scenarios was measured twice: once with voids present (void case) and once with the voids excavated and filled with soil (no void case).

Table 5-1 *Distribution and sizes of buried targets of the four scenarios*

	Distribution of the buried balls	Size of buried balls
Scenario A	Depth:30cm 	Shallow local voids (diameter: 36cm,36cm) marked as S; One ball (diameter: 28cm) One peanut shape ball (diameter: 28cm on both sides)
Scenario B	Depth:45cm 	Two balls (diameter: 28cm,28cm)
Scenario C	Depth:45cm 	One ball (diameter: 28cm); One peanut shape ball (diameter: 28cm on both sides)
Scenario D	Depth:60cm 	Three balls (diameter: 36cm,28cm,28cm)

designed methods. The pre-processing CLAHE function enhances the image details, while the superpixel segmentation and MRF modules contribute to noise resistance. Moreover, a multi-scale strategy is implemented to balance detail preservation and noise reduction.

A common phenomenon in change detection is that the changed areas often constitute a small fraction of the entire survey site, resulting in a large concentration of pixels with DI values concentrated around zero, as evident from the green histograms in Figure 5-7 (c) and (f). In such instances, image enhancement functions are necessary to improve the quality of the raw images for better analysis. The pre-processing CLAHE function plays a crucial role in enhancing image details. As illustrated in Figure 5-7, the DIs generated with CLAHE processing (b) and (e) and without (a) and (d) exhibit significant differences in the visibility of changes, as indicated by the red rectangles. Without CLAHE processing, some correct changes are compressed, making them difficult to discern. In contrast, the DI generated with CLAHE processing reveals hidden patterns in the raw images, as indicated by the contrast between the blue and green histograms in Figure 5-7 (c) and (f). This is because CLAHE extends the raw pixel values to a larger range, making concealed patterns more discernible.

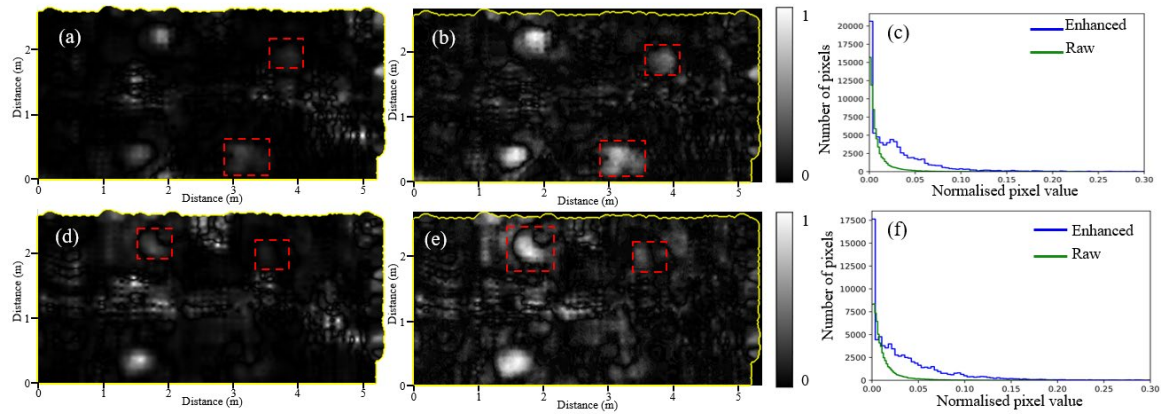


Figure 5-7 Example of the CLAHE image enhancement function for small intensity differences in DI. (a) and (d) are raw DI, (b) and (e) are enhanced DI, and (c) and (f) are histograms of C-scans. Red rectangles indicate the relatively small intensity change in raw and enhanced DI

In terms of noise resistance, two major types of noise are present in our task: local noise

and noise clusters. Local noise, characterized by high pixel values and random distribution, is effectively diminished after the superpixel segmentation module. In this module, each superpixel is assigned the mean value of the pixels within it. Averaged by neighbourhoods, superpixels containing local noise are less likely to be classified as changed areas in the subsequent FCM clustering and MRF classification. For instance, pixel-based DI and superpixel-based DI generated by methods (a) and (b) in Figure 5-8, respectively, demonstrate that some local noise (within the red rectangles) remaining in the pixel-based DI is eliminated through superpixel segmentation (shown in red rectangles in the panel). The other type of noise is cluster noise derived from the multiple representations of unchanged objects. In DI generation, the same targets in different C-scans sometimes remain incompletely eliminated, leaving some blocks at their edges, as shown by the white rectangles in (c), the superpixel-based DI created by (a) and (b) in Figure 5-9. The MRF module also plays an essential role in creating resistance to this kind of noise. The spatial component of MRF considers adjacent neighbourhoods of superpixels. For cluster noise in the superpixel-based DI, if there are no similar clusters in an adjacent field, the likelihood of classification as a changing area will decrease in the MRF model. It is unlikely that noise will cluster and repeat at the same location shown in C-scans at different depths. For example, Figure 5-9 (c) shows clustering noise with white rectangles, and there are no similar patterns in the adjacent C-scan (d). According to the energy function of MRF, these superpixels will usually be weighted as unchanged.

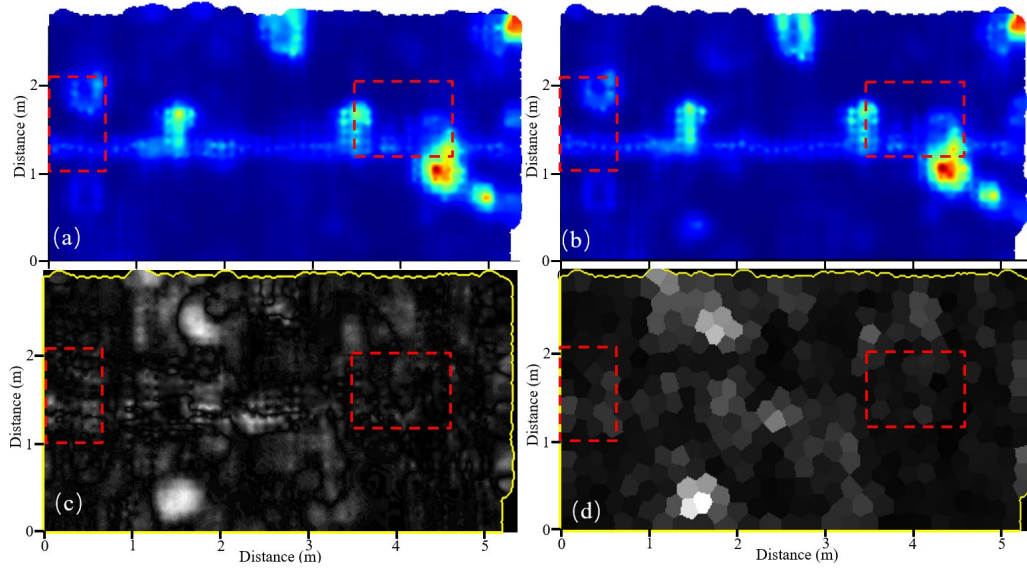


Figure 5-8 Illustration of superpixel segmentation module for local noise resistance. (a) and (b) C-scans with and without void, (c) pixel-based DI, (d) superpixel-based DI

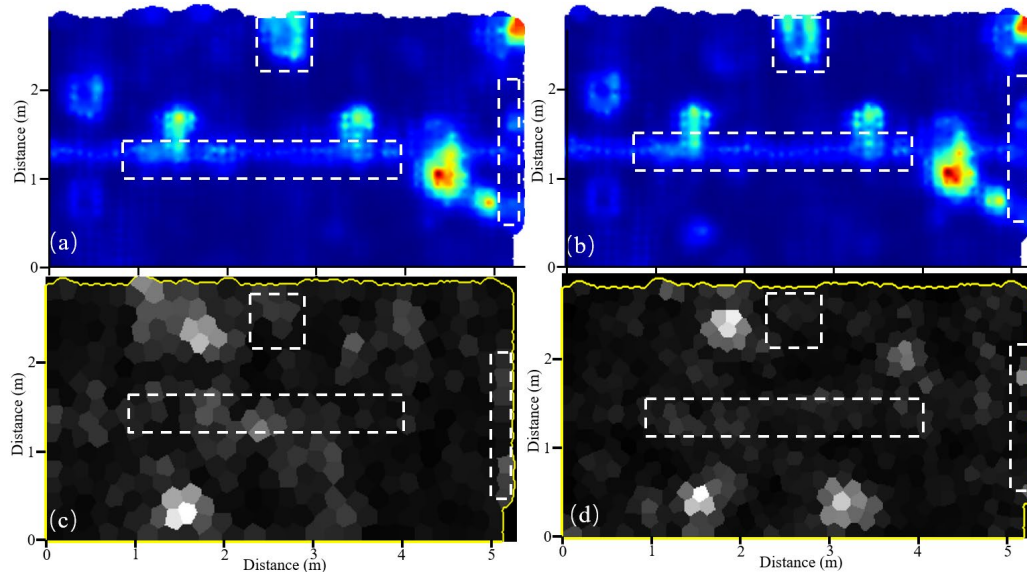


Figure 5-9 Illustration of MRF module for noise cluster resistance. (a) and (b) C-scans with and without void, (c) superpixel-based DI, (d) adjacent superpixel-based DI of (c)

Balancing image details and noise is integral in our proposed method. Equally important is the balance between the spatial and spectral components. The MRF model integrates these two types of information for change detection. In this experiment, we kept all other parameters fixed and varied only the value of β , the relative weight parameter between spectral energy and spatial energy in Equation [5-8], to record the F1-score achieved by different values of β . Figure 5-10 shows the F1 scores for the four scenarios. When the

parameter β is set to 0, the spatial component does not influence the energy function of the MRF, leading to subpar performance across all four scenarios, particularly the last three. As β increases from 0 to 4, all scenarios exhibit an increase in the F1-score, underscoring the beneficial impact of the spatial component. However, once β surpasses approximately 4, the F1-score begins to decline across all scenarios, as in Figure 5-10. Interestingly, Scenario A consistently outperforms the other scenarios in terms of the F1-score, irrespective of the β value. This suggests that the MRF's contribution is less significant in Scenario A and that the spectral component alone can yield satisfactory results. This is inherent to the nature of Scenario A, where all targets are centred within a single C-scan, thus enabling a single C-scan to capture all changes. Consequently, the spatial component's contribution is minimized in this scenario. Nonetheless, even with this minimized contribution, the F1-score still improves by 10% due to the spatial component. Based on our results, we selected β equals 4 as the optimal value, achieving the best balance between the spatial and spectral components in our method.

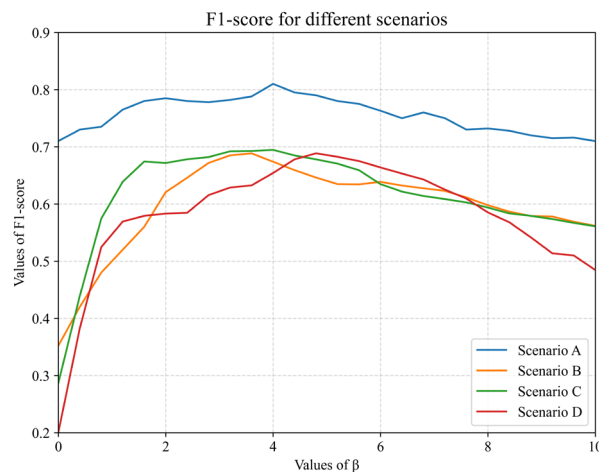


Figure 5-10 Influence of parameter β associated with four scenarios

Figure 5-11 demonstrates the effectiveness of the proposed method in detecting changes in the subsurface environment. The first and second columns of Figure 5-11. present the C-scans for the no-void cases and void cases in the four scenarios, respectively. Buried objects, such as continuous pipes and manholes, can be identified in these normalized amplitude

images. The third column displays the change maps generated by the proposed method using the pair of C-scans in the first two columns. The red circles in the change maps indicate the actual positions of the buried targets, and the white shapes represent the areas of change detected by the proposed method, with black areas indicating no change. As shown in Figure 5-11, the proposed method successfully extracted the main body of the buried targets in all four scenarios, although the detected changed regions did not fully align with the ground truth. The proposed method achieved a 100% detection rate at the decision level, and the assessment metrics (TP, TN, FP, FN, F1-score, and kappa coefficient) calculated at the pixel level are summarized in Table 5-2. The F1-score (0.8, 0.69, 0.69, and 0.65) and kappa coefficient (0.79, 0.69, 0.68, and 0.64) demonstrate the robust performance of this method. Analysis of the detection results revealed that the F1-score and kappa coefficient for scenario A were relatively higher than those for the other three scenarios at the pixel level. This can be attributed to the relatively shallower depth of the buried balls with two shallow voids, which resulted in a simpler and less complex background. However, as the depth increases, the expansion of the Fresnel zone causes a broader distribution of reflected energy. This increased energy influence affects the signal by incorporating contributions from surrounding areas, leading to greater interference and noise in the C-scans. These factors likely contributed to the observed differences in performance across scenarios. In the following section, we discuss in more detail how the sub-modules contribute to the superior performance of the proposed method.

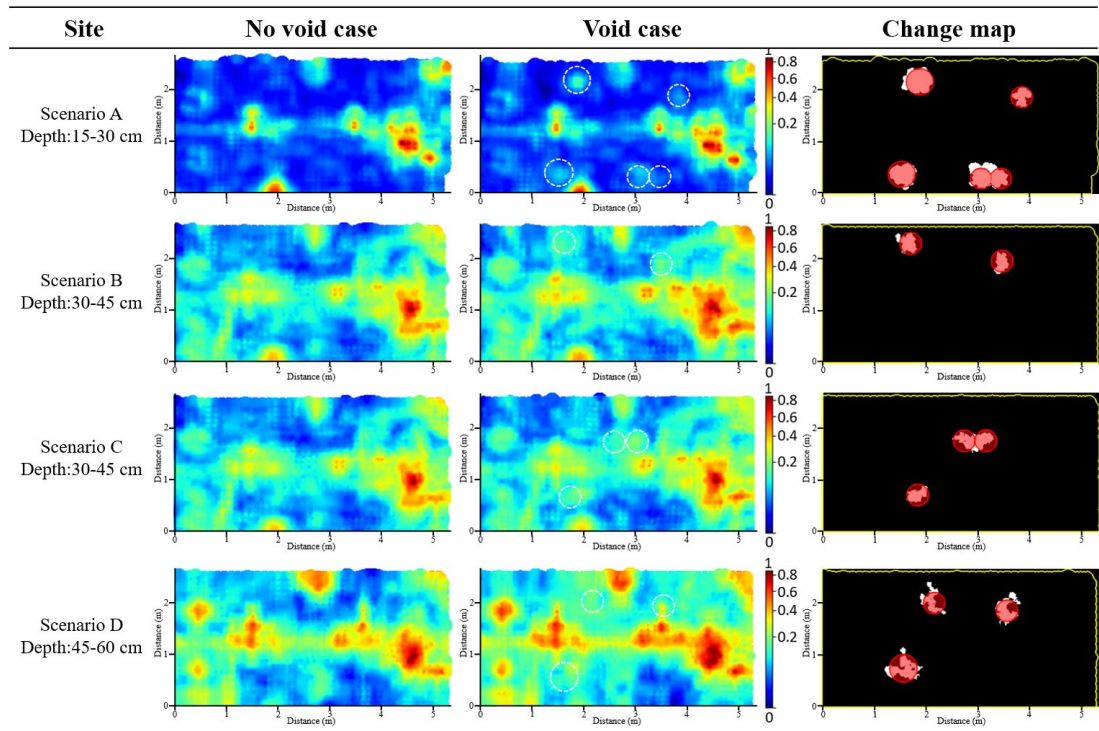


Figure 5-11 Pairs of C-scans (no void case and void case) and change maps created using the proposed method with corresponding references in the four scenarios (White regions are detected as changed areas and red masks refer to ground truth in change maps)

Table 5-2 TP, TN, FP, FN, F1-score, and Kappa Coefficient in the four scenarios

Dataset	TP	TN	FN	FP	F1-score	Kappa coefficient
Scenario A	3145	62,827	699	640	0.82	0.81
Scenario B	702	61,730	524	99	0.69	0.69
Scenario C	1012	61,136	826	94	0.69	0.68
Scenario D	1347	60,305	888	574	0.65	0.64

To assess the impact of the proposed method on the results of C-scan change detection, we performed a comparison among different methods. Due to the absence of prior studies using the same data resources, we designed the comparison using different combinations of sub-modules of our method. The findings are presented in Figure 5-12 to Figure 5-15. and Table 5-3, which depict the outcomes obtained by four different techniques for all scenarios. Figure 5-12 to Figure 5-15 (a) shows the change map generated by the FCM+MRF (pixel-based) method.

When compared with other subplots, the pixel-based method consistently produces scattered local false alarms through all four scenarios. In contrast, the superpixel-based method produces clusters that resist the effect of local noise, thereby emphasizing the contribution of superpixel segmentation in enhancing local noise resistance. Figure 5-12 to Figure 5-15 (b) demonstrates the effectiveness of the MRF filter in eliminating noise clusters by integrating information along adjacent C-scans, particularly in Scenario B, Scenario C, and Scenario D. The change maps generated by the FCM+MRF-1 (superpixel-based) and FCM+MRF-2 (superpixel-based) methods, shown in Figure 5-12 to Figure 5-15 (c) and (d), respectively, illustrate the significance of selecting the appropriate scale for the multi-scale strategy. FCM+MRF-1 uses the largest superpixels, while FCM+MRF-2 uses the smallest superpixels. Figure 5-12 to Figure 5-15 (c) and (d) illustrate the contribution of the multi-scale decision fusion, where coarse superpixels can miss detection in Scenario A and Scenario B, and fine superpixels can result in cluster noise in Scenario C. However, we observed less improvement in some scenarios when comparing the FCM and K-means methods. The primary distinction between these methods is that FCM employs a probability-based approach, as opposed to hard labelling, to characterize each pixel, as discussed in the methodology section. This variance confers an advantage to the overall method when the interpretation of the signal involves uncertainty, particularly in detecting comparatively deep changes. Consequently, while there is a relatively lower improvement observed in shallow Scenario A, an increased improvement is observed in other scenarios. Our proposed method achieved the highest F-score and kappa coefficient compared to all other methods, indicating its superior performance. The comparison results also highlight the contribution of each sub-module in our proposed method.

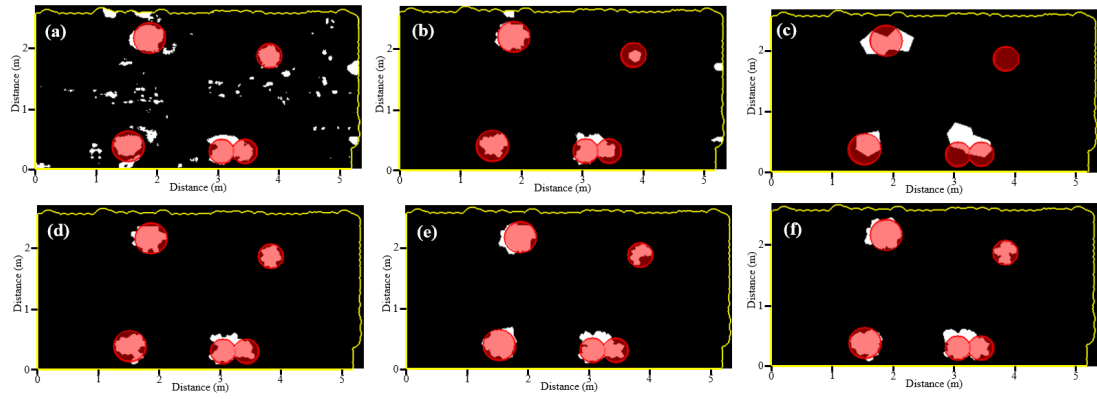


Figure 5-12 Change maps generated by different methods in Scenario A; (a) FCM + MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM + MRF-1 (superpixel-based). (d) FCM + MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based)

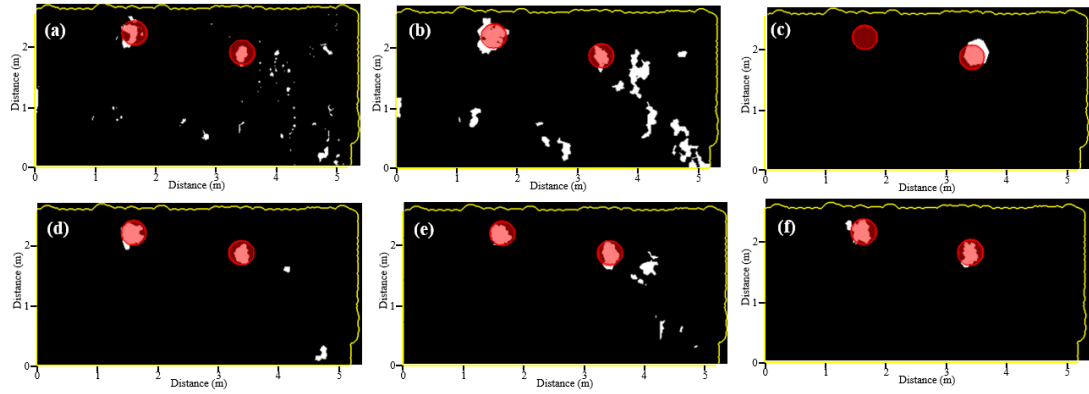


Figure 5-13 Change maps generated by different methods in Scenario B; (a) FCM + MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM + MRF-1 (superpixel-based); (d) FCM + MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based)

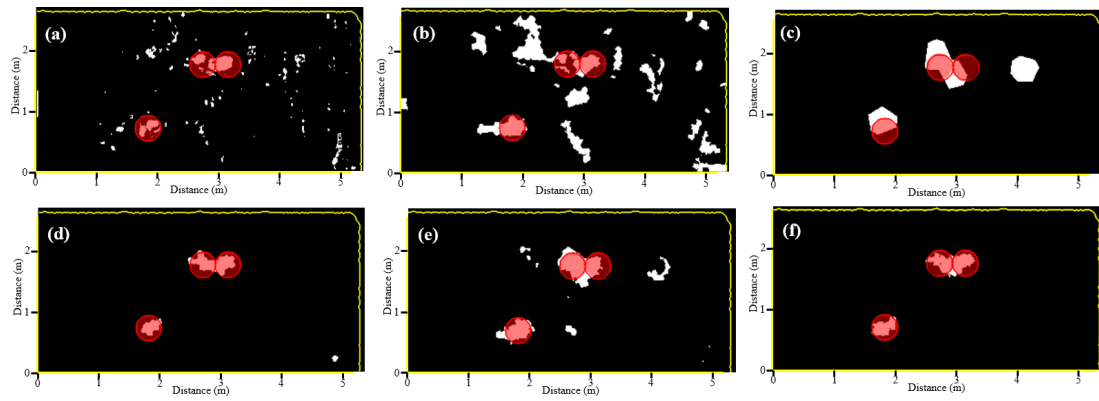


Figure 5-14 Change maps generated by different methods in Scenario C; (a) FCM + MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM + MRF-1 (superpixel-based); (d) FCM + MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based)

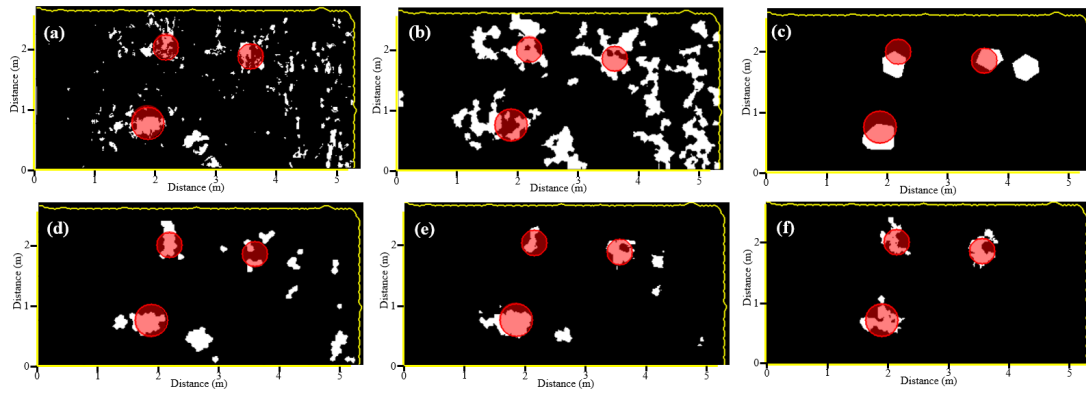


Figure 5-15 Change maps generated by different methods in Scenario D; (a) FCM + MRF (pixel-based); (b) FCM + Multiscale decision fusion (superpixel-based); (c) FCM + MRF-1 (superpixel-based); (d) FCM + MRF-2 (superpixel-based); (e) K-means + MRF + Multiscale decision fusion (superpixel-based); (f) FCM + MRF + Multiscale decision fusion (superpixel-based)

Table 5-3 Comparison of the change detection results obtained by different methods

Methods	Scenario A		Scenario B		Scenario C		Scenario D	
	F1-score	Kappa Coe	F1-score	Kappa Coe	F1-score	Kappa Coe	F1-score	Kappa Coe
FCM + MRF (pixel-based)	0.63	0.61	0.38	0.37	0.39	0.38	0.29	0.26
FCM+ Multiscale decision fusion (superpixel-based)	0.71	0.70	0.35	0.33	0.28	0.25	0.20	0.15

FCM+MRF-1 (superpixel-based)	0.48	0.51	0.45	0.45	0.44	0.42	0.46	0.44
FCM+MRF-2 (superpixel-based)	0.79	0.78	0.68	0.67	0.63	0.62	0.39	0.37
K-means + MRF + Multiscale	0.80	0.80	0.63	0.62	0.65	0.64	0.62	0.61
FCM + MRF + Multiscale decision fusion (superpixel- based)	0.82	0.81	0.69	0.69	0.69	0.68	0.65	0.64

This section presents a novel unsupervised change detection method for subsurface imaging using GPR time-lapse images. The method utilizes a combination of standard processing functions, image registration, superpixel-based DI generation, and image classification via FCM and MRF methods. The proposed method offers three key innovations: first, a clustering algorithm is applied to generate multiscale superpixels, which enhances local noise resistance and reduces the chances of mis- or over-detection due to inappropriate scale selection. Second, a fuzzy C-means clustering algorithm is used to classify the superpixel-based DI, which uses fuzzy membership to represent the relationships between different superpixels instead of hard labels. Finally, the MRF method is employed to integrate vertical information by considering the spatial component, which compensates for the fact that single-slice change detection typically focuses only on horizontal patterns and ignores vertical connectivity. The proposed method was tested in four scenarios, and the results indicate high performance and reliability as evidenced by F1-scores (ranging from 0.65, 0.69, 0.69, and 0.82) and kappa coefficients (ranging from 0.64, 0.68, 0.69, and 0.81). However, there are acknowledged limitations to this study. Specifically, the lack of regular field data to further test the robustness of the method, and the targets and test site were upscaled from a practical case. In conclusion, the proposed unsupervised superpixel-based change detection method offers substantial

potential for detecting subsurface changes in GPR temporal measurements. Further testing and refinement will be necessary to address the limitations of the current study and fully evaluate the effectiveness of this method for real-world applications.

5.2 Enhancing Water Leakage Detection in Urban Infrastructure by Integrating AS and IFS

As discussed in Section 4.2, IFS offers several advantages for characterizing water leakage. One of its key strengths lies in its ability to highlight variations in water content. Compared to the relaxation frequency, the IF value serves as a critical indicator of moisture presence, making IFS particularly effective for detecting water-related anomalies in unknown or complex underground environments. This contrasts with the void detection approach discussed in Section 5.1. Another advantage of IFS is its independence from the dynamic amplitude ranges that often complicate the interpretation of AS. While AS represents pseudo-reflected energy and is highly sensitive to signal propagation and subsurface environmental factors, IFS provides a more consistent and reliable framework for time-series observations and comparative analysis. By focusing on the time-frequency domain, IFS avoids issues such as misalignment and artefacts that can arise from amplitude variations or adjustments to interpolation parameters. This robustness makes IFS particularly suitable for identifying water content changes over time.

Additionally, the unique features of IFS demonstrate its strong potential for isolating abnormal regions in complex environments, providing a clearer understanding of water-related anomalies. To further improve detection performance, this section introduces a hierarchical leakage detection method that integrates both AS and IFS from GPR data volumes.

5.2.1 Leakage Detection Method

As Figure 5-16 shows, the whole leakage detection workflow encompasses three main modules:

- (1) Time-series amplitude and instantaneous frequency slices generation.
- (2) Change detection in amplitude slices.
- (3) Change detection in instantaneous frequency slices. In the following sections, the details of these modules will be introduced.

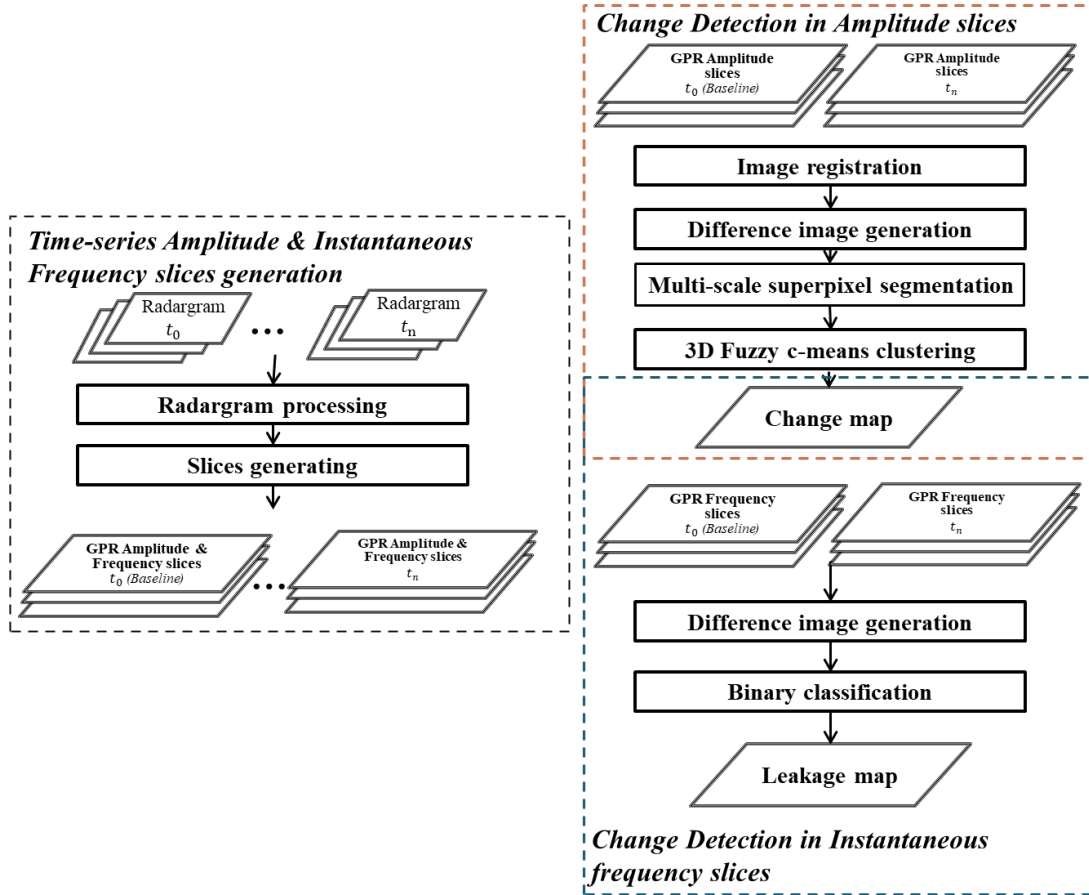


Figure 5-16 The whole workflow of the leakage detection method integrating amplitude and instantaneous frequency slices

The first module is time-series amplitude and instantaneous frequency slices generation through a batch of radargrams collected at different times. The details of the slice generation were illustrated in Figure 4-8 and Figure 4-9, in Chapter 4. Basically, two-dimensional signal

processing was employed to enhance GPR data quality through steps such as De-wow filtering, direct current shift correction, time zero correction, and AGC to amplify signals. A bandpass filter removed irrelevant frequencies, while global background removal reduced horizontal noise. Processed B-scans were assembled into C-scans based on a predefined survey grid, with local coordinates registered for 3D GPR imaging. Key parameters included a profile spacing of 25 cm (less than 3λ for optimal resolution), a slice thickness of 0.6λ , and bi-linear interpolation with a 0.2 m search radius to preserve local details. For generating IFS, a distinct approach was adopted to minimize artificial noise in the frequency domain. Unlike AS, which maps energy distribution horizontally, IFS reveals dominant frequencies at various depths. To ensure consistent gain, AGC was replaced with a two-point linear gain (-5 dB at electronic zero and +45 dB at the end of the A-scan). After basic signal processing, including De-wow and direct current shift correction, VMD was applied to generate the IFS, offering insights into frequency-based subsurface characterization. In this way, time-series AS and IFS were generated to further change detection in AS.

As described in Section 5.1.1, the change detection module consists of three main steps: data registration, difference map generation, and classification. Like the void detection workflow shown in Figure 5-1, image registration involves two components: geolocation registration and amplitude intensity registration. Geolocation registration was achieved using coordinates-based affine transformation, while intensity registration was conducted using affine transformation. The to-be-registered image was iteratively transformed to align with the reference image, with the SimpleITK registration framework ensuring stable and efficient performance. After C-scan image registration, superpixels were used as the basic processing unit to reduce noise and focus on significant changes. Superpixels, generated using the SLIC algorithm, cluster pixels based on intensity and spatial proximity. The DI was calculated as the absolute difference between CLAHE-enhanced registered C-scans at different times. To

improve accuracy, a multiscale strategy was applied, creating DI maps with varying superpixel sizes to balance global and detailed feature detection. This approach reduces noise, highlights meaningful changes, and provides a robust foundation for subsequent analysis.

FCM is one of the most popular unsupervised clustering algorithms. Classifying DI into change and unchanged categories by a clustering algorithm is a common procedure for unsupervised change detection. Assuming that changes have an effect among areas, resulting in high-intensity superpixels clustering together not only in horizontal slices but also in vertical space. In this study, we extend the FCM algorithm for C-scans change detection by introducing a 3D spatial factor in the objective function. Given C classes, superpixel-based DI is represented as $R = \{r_k | k = 1 \dots K\}$. The 3D FCM objective function is defined as:

$$J = \sum_{k=1}^K \sum_{c=1}^C u_{kc}^m (\|x_k - v_c\|^2 + G_{kc}), \quad [5-16]$$

$$u_{kc} = \frac{1}{\sum_{c=1}^C \left(\frac{\|x_k - v_c\|^2}{\|x_k - v_c\|^2} \right)^{\frac{2}{m-1}}}, \quad [5-17]$$

$$v_i = \frac{\sum_{k=1}^K u_{kc}^m x_k}{\sum_{k=1}^K u_{kc}^m}, \quad [5-18]$$

where u_k is the membership of superpixel k to class c , v_c is the centre of the class c , m is the fuzziness parameter, G_{kc} is the spatial factor, and x_k is the superpixel's grey value. The spatial factor G_{kc} is defined as:

$$G_{kc} = \sum_{j \in N_k} \frac{1}{D(k, j)} (1 - u_{kc})^m \|x_j - v_c\|^2, \quad [5-19]$$

$$D(k, j) = \sqrt{d(k, j)^2 + t(k, j)^2}, \quad [5-20]$$

where N_k is the neighbourhood of superpixel k , $D(k, j)$ is the Euclidean distance between superpixel k and j , calculated by projecting distance $d(k, j)$ and the thickness $t(k, j)$.

To further narrow down the leakage regions within the clusters showing high amplitude

changes, a second stage of change detection was performed using IFS. A binary classification was applied to the IFS-based difference map by implementing a thresholding method that considered both the magnitude of IFS changes and the cluster size. This process enabled the extraction of spatial clusters with significant IF changes, which correspond to regions where water content increased in the period. By filtering potential suspect leakage regions based on their IF property changes, this approach facilitates more precise localization of leakage areas across the entire survey region.

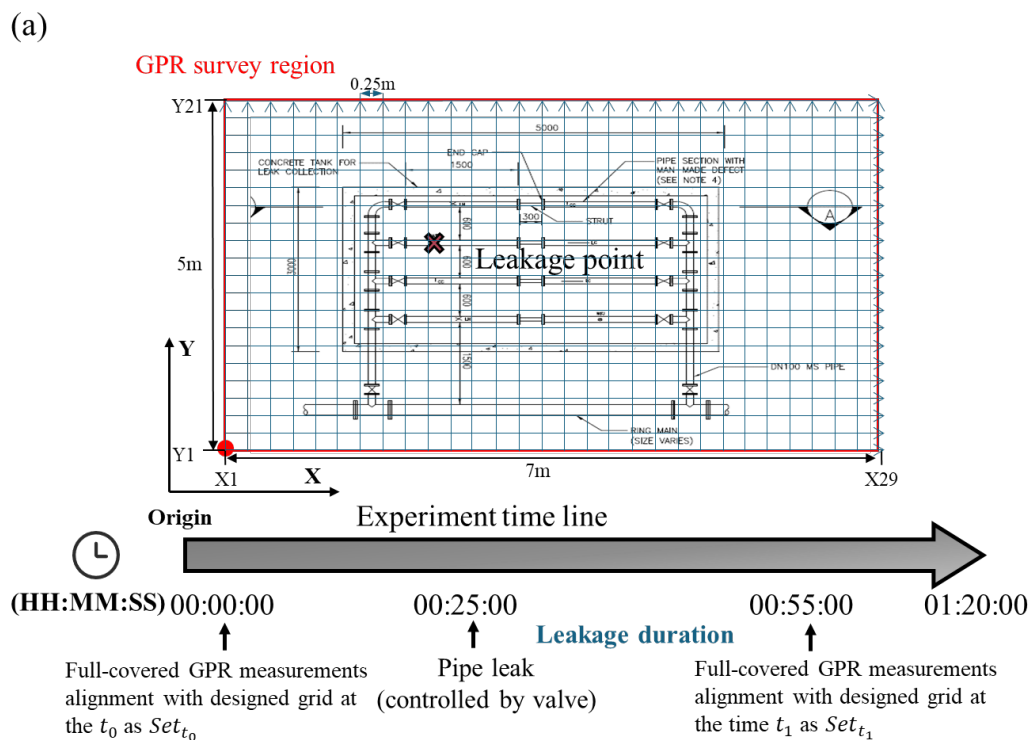
5.2.2 Experiment Results and Discussion

The GPR data used to verify the proposed method is from Dataset Two of the leakage signature database, detailed in Section 4.1.3. GPR measurements were conducted over a predefined survey grid with orthogonal scan directions and a spacing of 0.25 meters. These measurements were first taken before the leakage event. Leakage was then simulated by opening the control valve connected to the buried pipe at the designated leakage point. After a set leakage period, a second set of full-coverage GPR measurements was collected using the same survey grid. To minimize interference from previous leakage scenarios, only one leakage scenario was conducted at a time, with a recovery period of approximately two weeks to allow the leakage cell to return to its natural state.

Figure 5-17 (a) illustrates a leakage scenario where the leakage duration was 5 minutes, with a flow rate of 1430 L/min. The leak point is marked with a red cross. The pipes are enclosed within a concrete tank equipped with a drainage pipe at the bottom to prevent water from escaping the tank. Figure 5-17 (c) shows the cross-sectional view along the red line in Figure 5-17 (b). Based on this setup, the leakage region and background region were segmented based on whether the leaked water would affect the underground conditions in different areas, as illustrated in Figure 5-17 (b). Additionally, the manhole region was identified separately, as

GPR signals cannot penetrate metallic objects, resulting in multiple reverberations in this area. The segmentation of these three regions (leakage region, background region, and manhole region) helps to analyze the GPR data and evaluate the performance of the proposed leakage detection method.

Following the basic assumption that leaked water cannot escape from the tank, changes in the background region between the two GPR measurements are attributed to minor environmental factors. These changes are expected to be minimal compared to those in the leakage region, which represents the emulated water leakage. This segmentation approach ensures a focused evaluation of the proposed method by isolating regions most likely impacted by the leakage event.



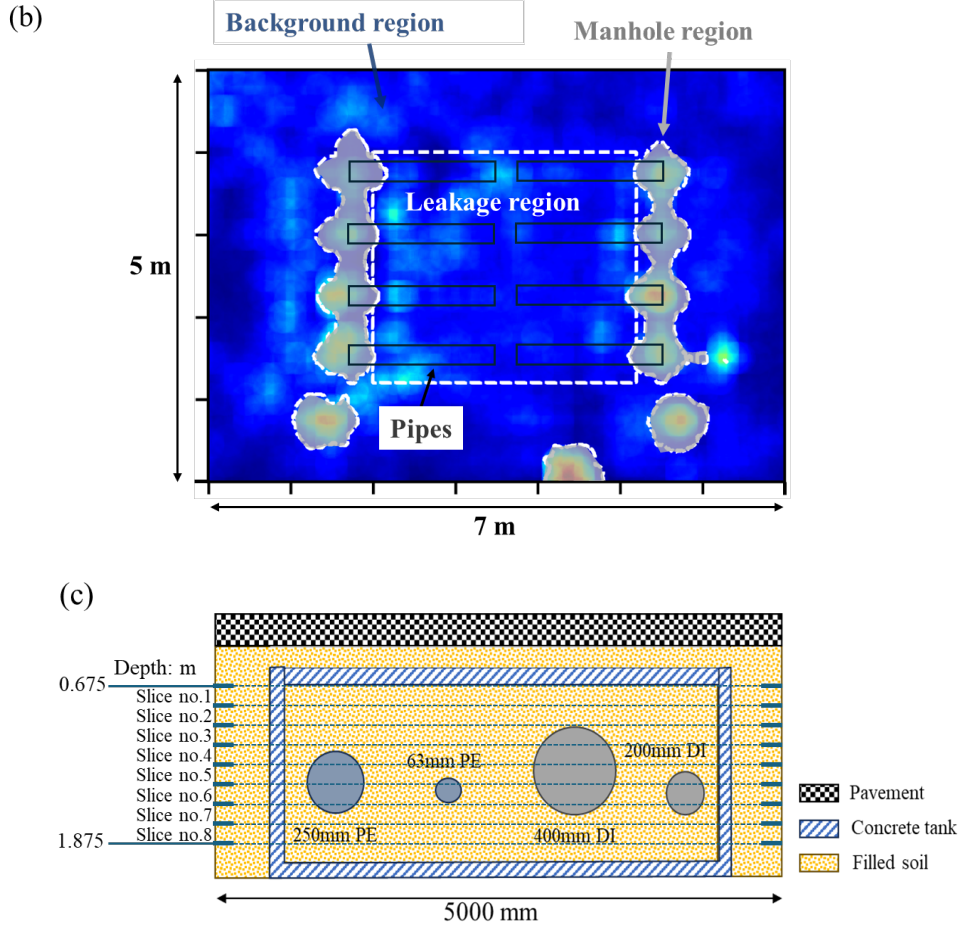


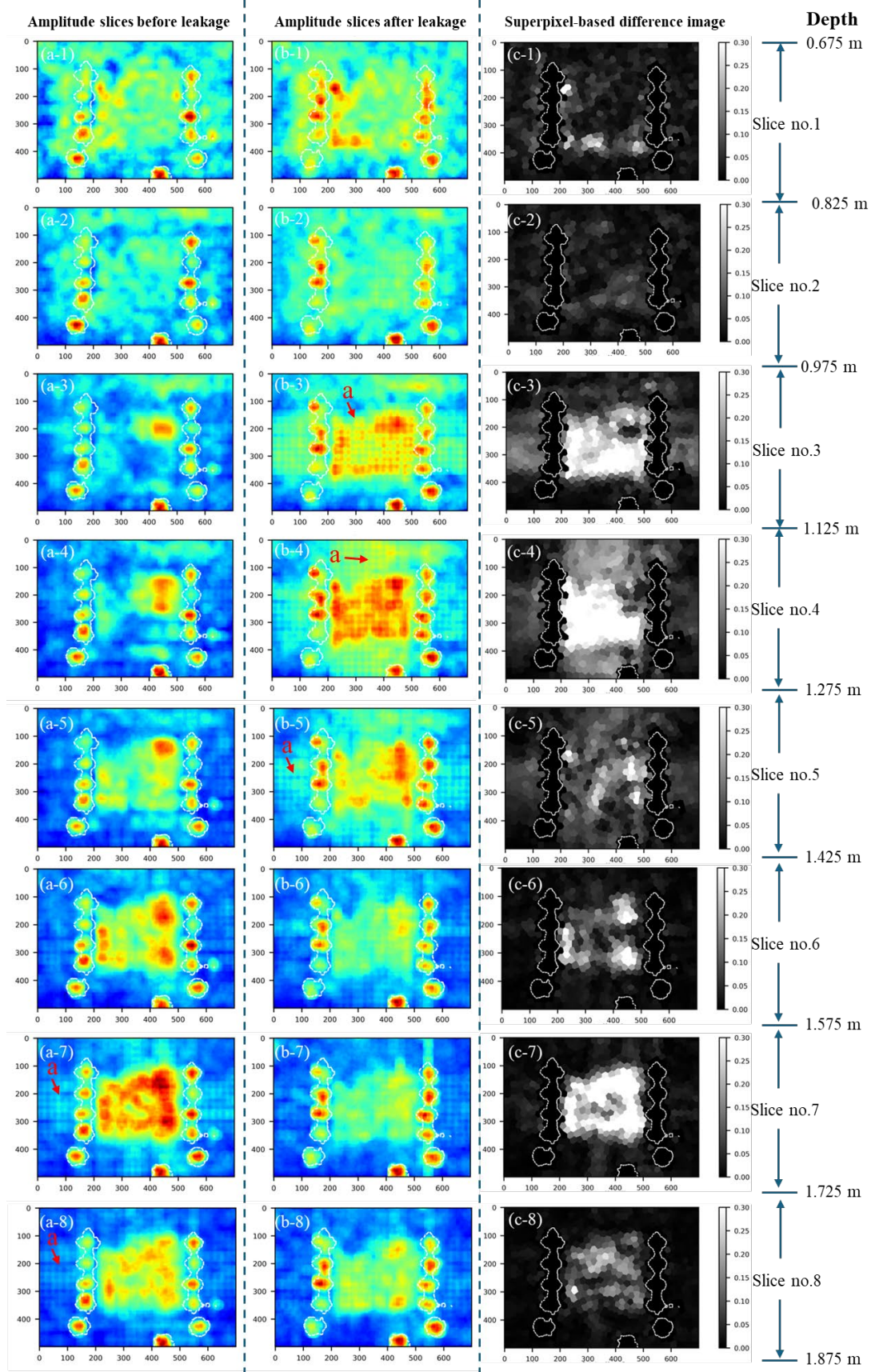
Figure 5-17 (a) Illustration of leakage experiment design; (b) Illustration of background region, leakage region, and manhole region; (c) Vertical cross-section (align with the red line in (b)) views of the experimental site

Figures 5-18 and 5-19 present the results of AS and IFS, respectively, generated at eight depth levels ranging from 0.675 m to 1.875 m (Slice no.1 to Slice no.8) with an interval of 0.15 m, as shown in the slice segmentation of Figure 5-17 (c). The analysis focuses on the leakage, background, and manhole regions, with detailed interpretations of the changes observed in both the amplitude and frequency domains.

In the leakage region, significant changes in amplitude values are observed in the superpixel-based difference images (c-3 to c-7), reflecting the presence of leaked water in the leakage region. In the no-leakage AS (Column 1), random clusters are visible in shallow slices (a-1 and a-2) due to the absence of pipes at those depths. As depth increases, high-amplitude

clusters appear (a-3 to a-7) corresponding to pipes, with the highest amplitudes occurring in Slice no.7. In the leakage AS (Column 2), amplitude values increase significantly from Slice no.1 to Slice no.4, peaking at the water table, and then gradually decrease from Slice no.5 to Slice no.8, indicating the spread of leaked water. In the background region, amplitude values in the no-leakage AS (Column 1) are stable and low across all slices, as no utilities are present. After leakage (Column 2), minor variations in background amplitude values are observed, likely due to environmental factors. These variations, however, are much weaker than the changes in the leakage region.

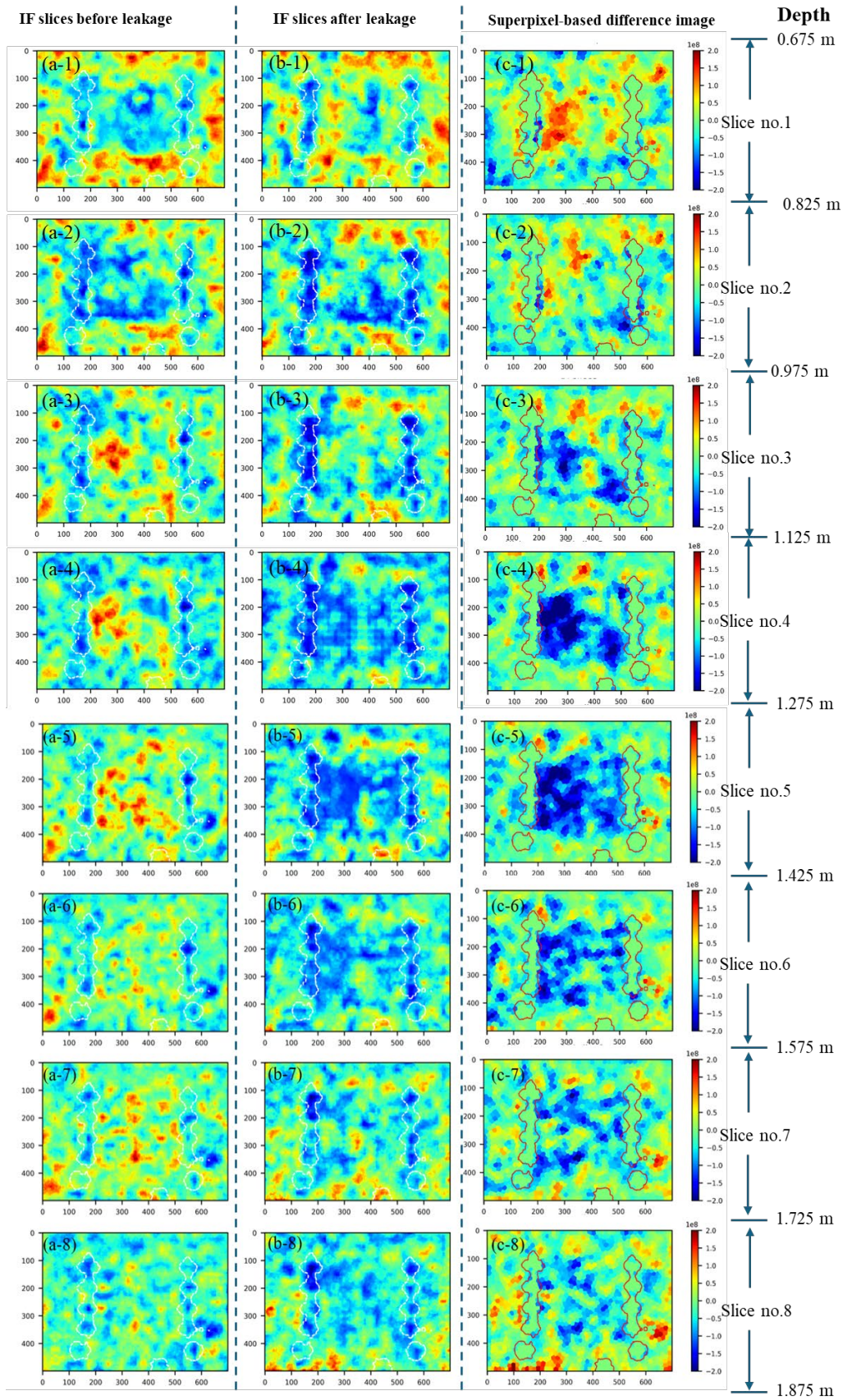
Artefacts such as orthogonal patterns, marked as red 'a' in slices a-6 to a-8 and b-3 to b-5, are visible in both no-leakage and leakage AS. These patterns are likely caused by interpolation artefacts arising from noise components in the signal. The manhole region, on the other hand, consistently shows high amplitude values across all slices due to metallic reflections, unaffected by leakage. The superpixel-based difference images (Column 3) effectively isolate amplitude changes caused by the leakage. High amplitude differences within the leakage region are particularly evident at depths corresponding to the water table, confirming the ability of AS to detect leaked water.



(Figure 5-18 Caption on the next page)

Figure 5-18 *AS without leakage, AS with leakage, and superpixel-based difference images generated from AS with and without leakage. All slices correspond to depth ranges from 0.675 m to 1.875 m at 0.15 m intervals. Remark: Arrows 'a' mark the orthogonal patterns that refer to artefacts during interpolation*

Figure 5-19 follows the same layout as Figure 5-18 but focuses on IFS. The superpixel-based difference images in Column 3 reveal notable changes in the leakage region between the no-leakage and leakage scenarios. These changes, which range from negative to positive values, represent frequency shifts caused by the presence of leaked water. From c-3 to c-7, the leakage region exhibits predominantly negative changes, corresponding to the same depth ranges where significant amplitude changes were observed in the AS difference images. This behaviour aligns with the discussion in Section 4.2, where it was shown that water leakage increases signal dispersion. High-frequency components are attenuated more significantly than lower-frequency components, resulting in a downward shift in instantaneous frequency values. In the background region, IFS values in the no-leakage slices (Column 1) remain stable and low, consistent with the AS results. However, in the leakage slices (Column 2), slight variability is observed, which is likely caused by environmental noise or minor signal disturbances. The manhole region consistently displays high and stable IFS values across all slices due to metallic reflections, which are unaffected by leakage. These reflections dominate the signal in this region, masking any potential frequency shifts caused by the leakage.



(Figure 5-19 Caption on the next page)

Figure 5-19 IFS without leakage, IFS with leakage, and superpixel-based difference images generated from IFS with and without leakage. All slices correspond to depth ranges from 0.675 m to 1.875 m at 0.15 m intervals

Four leakage scenarios from the dataset described in Section 4.1.3 were used to evaluate the performance of the proposed leakage detection method. Table 5-4 provides a detailed description of these emulated scenarios, which vary in terms of pipe material, size, burial depth, leakage flow rate, and data acquisition time. The scenarios include leaks in pipes made of mild steel (MS), polyethylene (PE), and ductile iron (DI), with diameters ranging from 100 mm to 250 mm, buried at depths of approximately 1.1 m. Leakage flow rates span from 28 L/min to 1430 L/min, with data collected at different time intervals after leakage onset, demonstrating a range of conditions for method evaluation.

Table 5-4 Detailed description of emulated leakage scenarios

Leakage scenario	Material hosting the leak pipe	Leak pipe size (mm)	Buried depth (m)	Flow of leakage (L/min)	Data acquisition time
Scenario 1	MS	100	1.110	1430	5 minutes after leakage
Scenario 2	PE	100	1.160	1281	20 minutes after leakage
Scenario 3	PE	250	1.125	33	5 minutes after leakage
Scenario 4	DI	200	1.117	28	5 minutes after leakage

The results are presented in Figure 5-20, which is organized into three columns and four rows. The first column shows the change map derived from AS, while the second column displays the change map derived from IFS. The third column presents the final leakage map, which combines information from both AS and IFS change maps. Each row corresponds to one of the four different leakage scenarios analyzed in the study. In the final leakage maps, the yellow regions represent significant changes identified in the AS change map, and the blue regions highlight significant changes from the IFS change map. The red cross marks the actual location of the leakage point in each scenario. This visualization demonstrates how the

integration of AS and IFS data enhances the accuracy of leakage detection, with the overlap of yellow and blue regions near the red cross indicating the proposed method for identifying leakage zones.

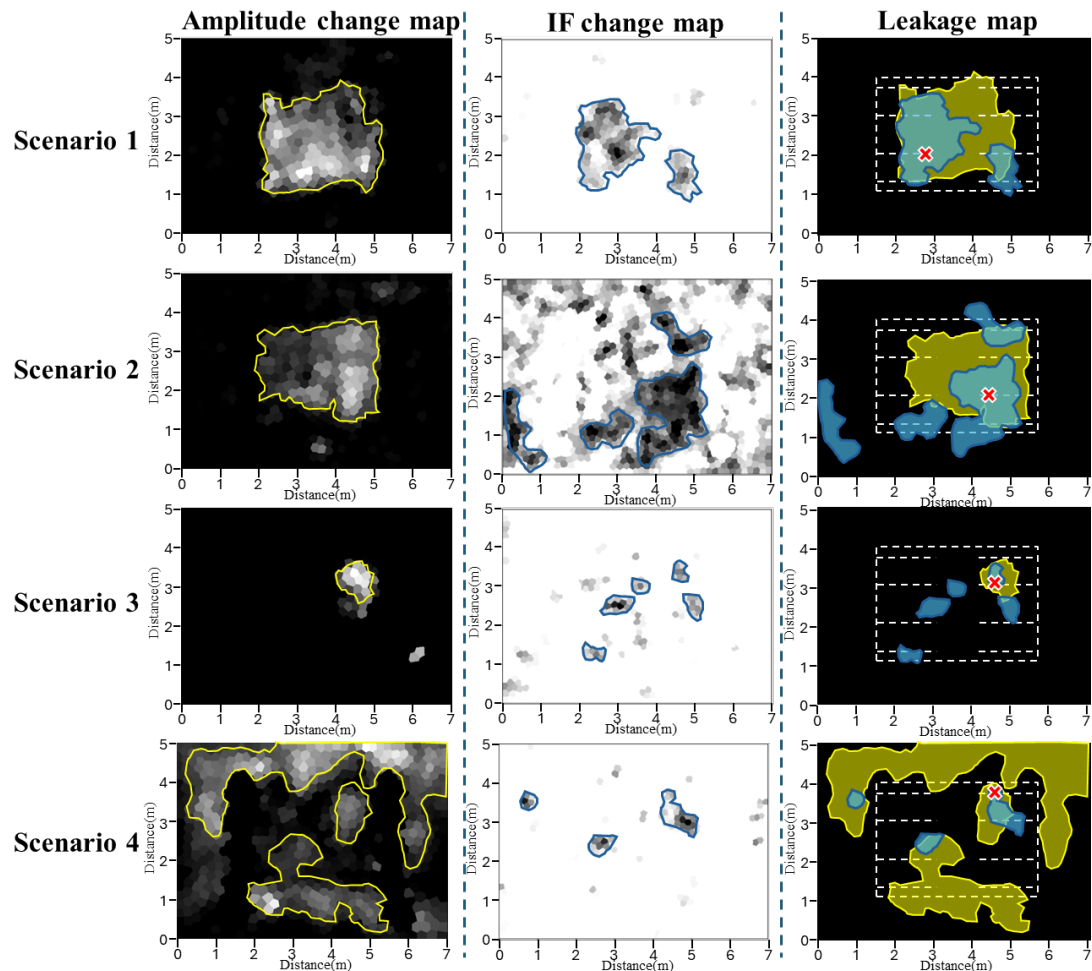


Figure 5-20 Results of four leakage scenarios

The first and second scenarios simulated severe leaks, with water filling the tank. In these cases, both the AS- and IFS-based methods consistently identified the same regions as leakage areas, demonstrating strong agreement and reliable performance in detecting severe leakage. Although the exact spread of underground water from the leaks remains uncertain due to data limitations, the detected leakage regions are closely aligned with the actual leakage points, marked by red crosses. This alignment supports the hypothesis that underground water tends to spread near leakage sources. In contrast, the third and fourth scenarios simulated minor

leaks. For these cases, the AS-based results included additional clusters, introducing some uncertainty due to false positives. The IFS-based results provided further validation by leveraging frequency-domain information, but they also exhibited false alarms at different locations compared to the AS results. This discrepancy suggests that the method has limited capability in detecting very localized and minor leakage.

Despite these limitations, the strong alignment between the leakage map and the actual leak points in severe leakage scenarios demonstrates the method's effectiveness in detecting large-scale leakage. Given that real-world leakage often involves large-scale regions, this approach shows promise for practical applications in severe leakage detection.

5.3 Chapter Summary and Contribution

This chapter presents two innovative unsupervised methods for detecting underground hazards using temporal GPR data, validated through both laboratory experiments and the leakage signature database in Section 4.1. In Section 5.1, a superpixel-based void detection method is introduced. This approach uses FCM clustering to generate initial change maps based on soft probability relationships, while an MRF model integrates spectral and spatial components to achieve robust 3D void detection. Section 5.2 proposes an enhanced water leakage detection method that combines AS and IFS. The AS-based method employs 3D FCM clustering to identify high-amplitude change clusters, while the IFS-based method refines the leakage regions by leveraging frequency-domain characteristics unique to water leakage, significantly improving detection accuracy.

The distinct nature of the input data and change detection processes reflects the differences in how voids and water leakage interact with EM waves. For void detection, changes in the GPR signal result from interactions between EM waves and host materials, with voids causing significant signal disruptions due to reverberation and scattering. These

interferences represent the primary changes in the GPR signal. The MRF model is used to balance spectral and spatial components to accommodate the variable shapes of voids, including flat, layer-like structures commonly referred to as delaminations. These delaminations are characterized by minimal vertical spread but exhibit significant spectral contrast. To ensure effective detection, the void detection method treats spectral contrast and spatial continuity as equally important factors.

In contrast, water leakage detection depends on the gradual changes caused by varying water content, as most underground materials absorb some water. This absorption results in broader regions of similar-scale changes. To distinguish these changes from environmental variations, IFS captures the dispersion of EM waves, reflecting water content alterations from a frequency-domain perspective. Furthermore, as water naturally spreads vertically due to gravitational forces, continuity across adjacent vertical slices becomes a critical condition for clustering in 3D FCM.

The methods proposed in this chapter mark a significant advancement in unsupervised underground hazard detection, offering robust solutions to challenges such as noise, spatial variability, and differing physical characteristics of underground anomalies. By seamlessly integrating spectral, spatial, and temporal information, these frameworks enhance the robustness and accuracy of GPR-based detection. Moreover, the adaptability of these methods to both laboratory and real-world scenarios highlights their potential for practical applications, including infrastructure maintenance and environmental hazard monitoring.

Chapter 6. Time-series MCGPR Registration: From Geolocation to Signal Domain

As discussed in Chapter 2, the transition to large-scale GPR surveys using MCGPR and the shift from single-shot measurements to time-series observations introduce new challenges that must be addressed to ensure accurate and reliable subsurface imaging. The first major challenge is geolocation accuracy. Traditional GPR surveys relied on local coordinate systems based on pre-designed survey grids, whereas modern approaches integrate GNSS-based positioning to establish global coordinate systems. High-quality positioning is essential for MCGPR surveys, as positional distortions and misalignments can lead to significant data misinterpretation. Achieving precise positioning in complex urban environments is particularly challenging due to signal obstructions and multi-path errors. To address these issues, section 6.1 presents a novel post-processing method that enhances positioning accuracy by leveraging cross-channel measurements from MCGPR data. This method involves three key steps: (1) underground features are extracted from quality-assessed MCGPR data using advanced techniques such as the histogram of oriented gradients (HOG) and 2D fast Fourier transform (2D FFT); (2) these features are paired based on their similarity to reconstruct relative positions within the antenna array; and (3) the entire survey trajectory is refined through global optimization using dynamic programming with constraints informed by driving behaviours. A unique advantage of this method is its ability to use MCGPR data as external information to enhance positioning accuracy, ensuring compatibility with various positioning systems and making it particularly suitable for urban environments. The second challenge lies in signal variations within the MCGPR antenna array, where differences between antennas introduce noise that can distort data quality and lead to potential misinterpretation. This issue and its solutions are explored further in section 6.2. By addressing these two challenges—geolocation

accuracy and signal variations—this chapter establishes a framework for improving the reliability and accuracy of time-series MCGPR data, paving the way for more robust subsurface analysis and imaging.

6.1 Enhancing MCGPR Positioning Using Cross-Channel Data

MCGPR systems are increasingly utilized for underground inspections in urban settings due to their efficient data acquisition and the ability to operate without disrupting traffic. Benefit from these features, MCGPR has played a pivotal role in subsurface hazards detection, such as cavities (Kang et al., 2019c; Yamaguchi et al., 2022c), roadbed damage (Xu et al., 2014), in underground utility mapping, including manholes (Yamaguchi & Mizutani, 2021), pavement layer (Liu et al., 2016), and underground pipes (Gabryś & Ortyl, 2020; Yamaguchi et al., 2021), and in underground investigation, for instance, uncharted utilities and no-man land for trial pit. An MCGPR system commonly consists of array-like GPR antennas and a positioning system. High-precision positioning data is indispensable not only for accurately correlating targets detected in GPR images to their real-world geolocations but also for effective interpretation of GPR data. Inaccurate positioning can lead to misalignment and distortion in MCGPR presentations (Zhang et al., 2024), which limits its use. Enhancing the positioning of the MCGPR system has become a critical topic in practical applications.

6.1.1 Position Solutions of MCGPR Survey

Positioning solutions for GPR surveys have always been a crucial topic in the field of GPR applications. In the early stages of a GPR survey, GPR antennas are located by establishing a local acquisition survey grid (Pipan et al., 1999) and recorded by an encoder. Later, the coordinates were obtained by geolocating the start and end points of each survey line using GNSS receivers or total stations. The limitations of this survey grid approach included its

suitability only for enclosed environments, the labour-intensive nature of the process, and the requirement for straight survey lines following pre-designed survey grids. In order to simplify the GPR data acquisition, the positioning system starts to be anchored on the GPR system to ensure continuous and free. Various solutions for GPR surveying have emerged. GNSS has become one of the most popular positioning solutions used in MCGPR (Gabrys & Ortyl, 2020). In open-sky conditions, GNSS can provide continual positioning. With the RTK corrections from the base station, centimetre-level accuracy can be achieved. However, obstruction of tall buildings and roadside tree canopies and multi-path effects in urban environments make it difficult for GNSS-RTK to fix. This degrades GNSS accuracy to several decimeters or even meters (Cornick et al., 2016). In densely populated cities with tall buildings, such as Hong Kong, the fixed rate of GNSS-RTK can drop between 33.4% and 72.4% (Xu & Chen, 2018), see Figure 6-1.

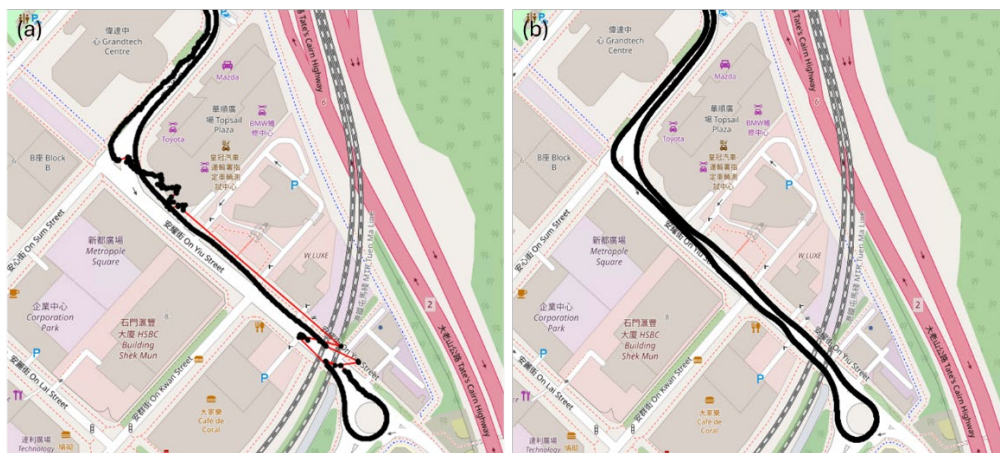


Figure 6-1 illustration of GNSS positioning performance in the urban region; (a) GNSS recorded trajectories; (b) referenced trajectories by mobile mapping system

Various methods have been developed to enhance positioning accuracy in challenging environments. Sensor fusion has proven to be one of the most effective strategies. By integrating a self-tracking system (Boniger & Tronicke, 2010; Ching et al., 2021; Šarlah et al., 2020), centimeter-level positioning accuracy can be achieved. A laser connects the GPR and

the total station through a prism. This approach requires a clear line of sight between the stations and the prism, which limits flexibility and convenience in practical applications. Additionally, some research efforts have focused on fusing vision information with GNSS data (Barzaghi et al., 2016; Sukhanov et al., 2017). These vision-based methods are constrained by the quality of optical features, which can be affected by the camera lens and environmental illumination. Moreover, extensive studies have explored point cloud data, including mobile mapping systems (MMS) and simultaneous localization and mapping (SLAM) systems. These approaches integrate existing commercial mobile measurement systems with GPR and utilize SLAM algorithms to achieve positioning in areas lacking GNSS signals (Ciampoli et al., 2021; Grasmueck & Viggiano, 2007; Hui et al., 2022; Merkle et al., 2021). Reconstructing 3D models from point clouds requires significant computational power and substantial investment in equipment. Taken together, these studies improve positioning accuracy by incorporating diverse environmental data from multiple sensors.

In the context of MCGPR surveys, besides the additional sensors mentioned above, MCGPR systems themselves can serve as supplementary sensors by characterizing underground features, as explored in autonomous driving research (Ort et al., 2020). This study presents a novel post-processing method that, to our knowledge, is the first to leverage cross-channel MCGPR data to enhance positioning accuracy. Our approach utilizes underground features as reference points, akin to feature points employed in LiDAR or optical camera positioning methods on the above surface, to retrieve the relative antenna positions. The key feature of our method is its compatibility. Our approach does not require any additional data and can be applied to various positioning systems regardless of the sensor combinations, potentially enhancing overall positioning accuracy.

6.1.2 MCGPR Positioning Enhancement Method

As illustrated in the workflow in Figure 6-2, our method enhances position accuracy through integrating MCGPR measurements. There are two types of input: raw trajectories obtained from the positioning system and MCGPR data itself. The performance of positioning systems is inherently influenced by the characteristics of the sensors and methodologies employed, leading to variations in trajectory quality across different environments. Such discrepancies can pose challenges for MCGPR applications, where consistent accuracy is crucial. Meanwhile, MCGPR systems demonstrate robust functionality, continuously measuring reflected GPR signals throughout their operational paths. We initiate our approach by categorizing raw trajectories. This classification relies on quality control metrics derived from the positioning systems, such as the quality index found in GNSS records. This process enables us to pinpoint segments in need of enhancement while using good-quality trajectories as reference points for subsequent trajectory reconstruction. As for MCGPR data, we subject the MCGPR measurements to fundamental signal processing techniques designed to mitigate noise and highlight underground features. This basic GPR processing is succeeded by underground feature extraction, which captures underground features in swath-specific segments. The extracted underground features from multiple swaths serve as anchor points for aligning and pairing underground features effectively. To reconstruct the positions of the identified low-quality segments, we employ a global optimization that integrates the results from the underground features pairing. It is important to note that trajectory classification is contingent upon the specific positioning systems used, rendering it a user-defined function that can be tailored to meet situational needs. The subsequent section will focus on MCGPR processing and the analysis of underground features, beginning with an overview of GPR signals. The resolutions of MCGPR data will be introduced, which delineate the boundaries of accuracy for our proposed method, along with the common signal processing techniques

employed to enhance the visibility of underground features within radargrams. Finally, the core components of our proposed methodologies will be explained, including underground feature extraction, pairing, and trajectory reconstruction.

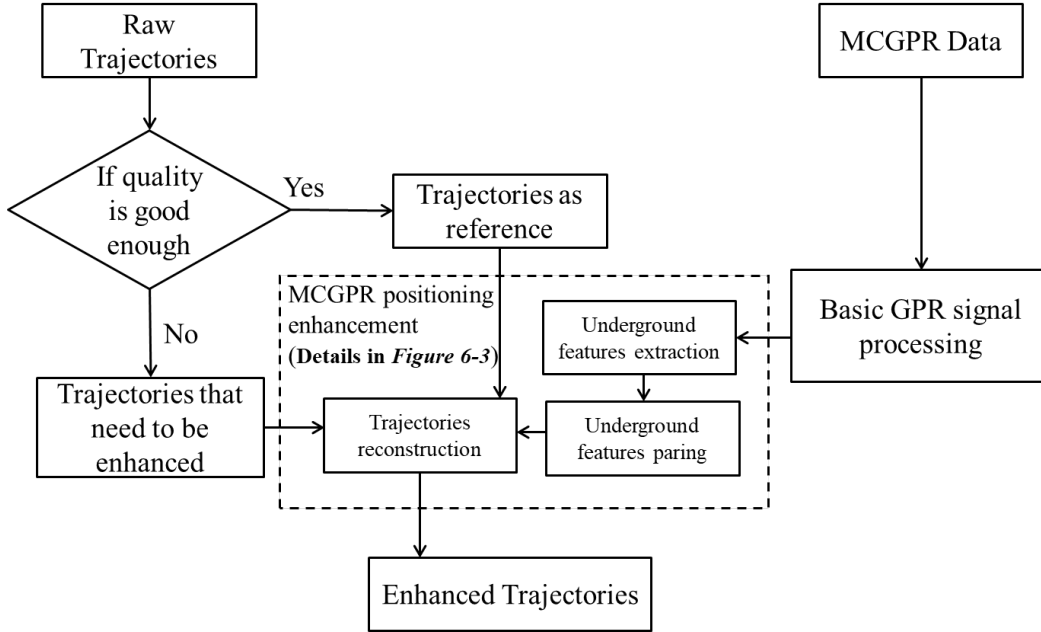


Figure 6-2 The whole workflow of the proposed positioning enhancement method

By extracting these underground features and measuring correlations across multiple MCGPR channels, relative positioning data can be derived to reconstruct accurate trajectories. This approach relies on the hypothesis that underground features are distinctive and stable enough to serve as benchmarks for precise positioning across multiple channels. Two factors derived from the input data specifically influence the effectiveness of this method: the resolution of MCGPR data, which determines the maximum achievable accuracy, and the quality of MCGPR images, which directly affects the clarity and reliability of underground features.

6.1.2.1 MCGPR Data Resolutions

Along-track, cross-track, and range resolutions represent the data resolutions in the three dimensions of MCGPR data volumes. GPR measurements are commonly obtained by

moving a transmitter and receiver set across the surface to detect reflections from below. These components are typically fixed in relation to one another. An odometer attached to the wheel triggers the recording of reflected signals at set intervals, determining the along-track resolution. In the MCGPR system, the cross-track resolution is defined by the spacing between adjacent antennas in the array. The terms "along-track" and "cross-track" pertain to the sampling resolution of the MCGPR system. Additionally, an essential attribute of GPR imagery is the range resolution, denoted as R_{res} . This parameter measures the radar's capability to differentiate between two proximate targets. The discernibility of two such targets is possible when their respective transient returns are temporally well-defined. Typically, the half-width (-6 dB in voltage) or half-power (-3 dB in power) benchmark of the return signal is utilized to signify a distinct separation (Eaves & Reedy, 2012). By defining the range resolution using the half-power point of the normalized *sinc* function, one can derive R_{res} from the maximum theoretical resolution as:

$$R_{res} = \frac{1.39c}{2B\sqrt{\epsilon_r}}, \quad [6 - 1]$$

where c is the speed of light, B is the bandwidth, and ϵ_r is the dielectric constant or relative dielectric permittivity. The factor of 1.39 is associated with the deviation from the theoretical range resolution and is derived empirically.

The accuracy of the proposed method is limited by the along-track and cross-track resolutions in the heading and lateral directions. The along-track and range resolutions are closely related to the quality of underground features extracted from MCGPR images. The following section will detail the processing of GPR images to highlight the underground features needed for positioning enhancement.

6.1.2.2 Basic GPR Signal Processing

Signal processing of GPR images in the proposed method aims to minimize noise and

artefacts while highlighting underground features that can serve as consistent identifiers across different channels. Hyperbolic patterns from localized subsurface objects (e.g., rocks, pipes, utilities) and the layers resulting from material interfaces (e.g., pavement-soil, or soil-clay) are ideal markers for this purpose. The following post-processing workflow is executed to enhance these identifiers: Dewow and DC-shift remove low-frequency noise and direct current bias, respectively, which can obscure the true signal; time-zero correction corrects the different heights of array antennas; AGC adjusts the signal strength to compensate for attenuation with depth. This helps reveal weaker, deeper targets. To enhance local layer visibility and reduce "ringing" effects from the radargrams, background subtraction is applied (Luo et al., 2019b; Zhou et al., 2023; Zhou & Lai, 2024). After basic GPR processing, we move to the main core of the proposed method to leverage underground features for positioning enhancement.

In the event of positioning signal disruptions due to environmental obstructions, the workflow is delineated in Figure 6-3, which shows how the proposed method corrects the poorly positioned data. Captured MCGPR volumes are dimensioned with a length matching the distance of the measurements in the direction of travel, a height that delves from the surface to the maximum depth of penetration, and a width defined by the span of MCGPR coverage. The HOG and 2D FFT are employed for their efficacy in extracting distinctive textural and frequency features. Following underground feature extraction, the underground feature pairing mechanism takes over, operating through a sliding window approach along the trackway, to identify and align analogous features between separate MCGPR swathes. This underground feature pairing process generates a sequence correlation matrix, which is then optimized with dynamic programming to derive the best relative positioning, constrained by reasonable driving behaviors. Using the calculated spatial offsets, the disrupted trajectory segments are reconstructed to restore continuity to the navigation paths. The following sections provide further details on the steps involved: underground feature extraction, underground feature

pairing, and trajectory reconstruction.

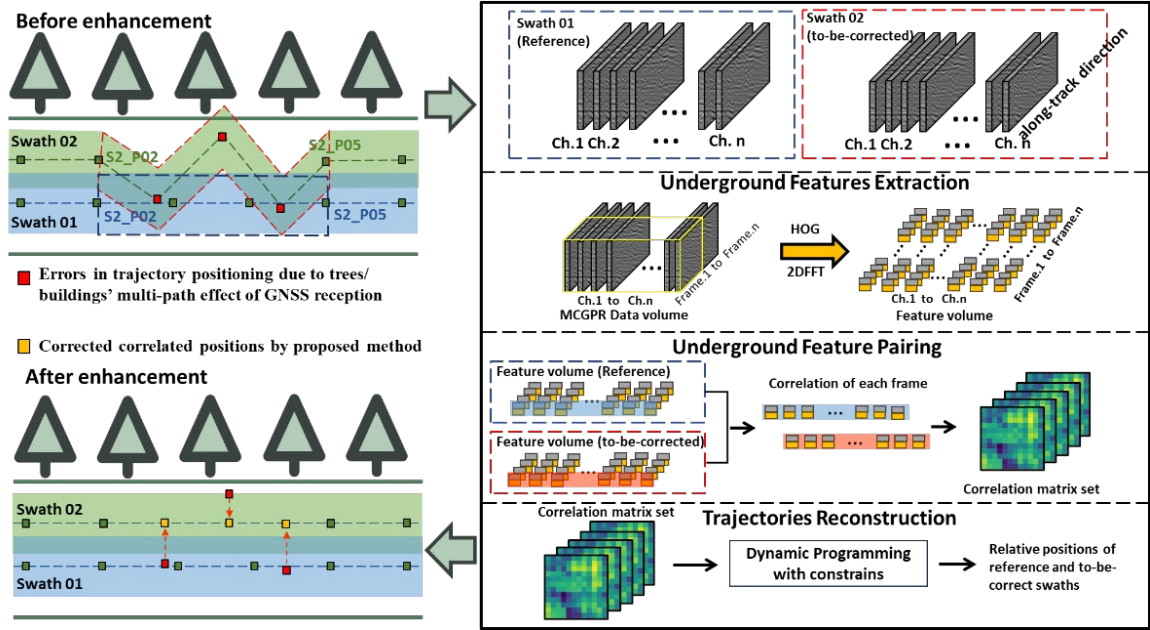


Figure 6-3 Workflow for trajectory reconstruction using cross-channel MCGPR image pairing (extended from **Figure 6-2**)

6.1.2.3 Underground Feature Extraction

The descriptor is engineered for feature extraction by combining time-domain imagery with frequency-domain characteristics, integrating the HOG and the 2D FFT. The HOG is a pivotal feature descriptor for object detection and image analysis, proficient at capturing image shape and appearance through the computation of gradient orientations and magnitudes (Yao et al., 2015). Details can be referred to Figure 6-4. It operates by segmenting the image into smaller regions called cells and accumulating a histogram of gradient directions for the pixels within each cell. This aggregation of histograms forms the descriptor. Due to its robustness to geometric and photometric changes, HOG has found extensive use in various computer vision tasks (Gan & Cheng, 2011; Kim et al., 2015; Wei et al., 2019; Yao et al., 2015). HOG features, which are based on the distributions of gradient angles and magnitudes, are particularly well-suited for GPR data. Two key attributes of HOG that are especially relevant to GPR applications are its insensitivity to small image translations and its resistance to amplitude variations. These

characteristics make HOG a valuable tool in the detection of objects in GPR contexts, such as landmines (Reichman et al., 2017b, 2017a; Stone et al., 2015; Torrione et al., 2014), manholes (Yamaguchi & Mizutani, 2021), and pipes (Bordón et al., 2019). Furthermore, HOG features are advantageous in the proposed method due to their ability to mitigate variations arising from different antennas of MCGPR.

In this section, the MCGPR data volume is divided into smaller units, referred to as "frames," by segmenting the data along the along-track direction using a defined "frame size" parameter. A "frame" represents a specific block of data within the MCGPR volume, as illustrated by the yellow cube in the underground feature extraction function in Figure 6-3. 'Frame size' is a crucial parameter for feature extraction. A larger frame size captures a broader range of GPR signals, resulting in more global features that are better at resisting noise effects. In contrast, a smaller frame size focuses on more localized features, which are more susceptible to noise interference. The key is to select an appropriate frame size that balances feature stability and aligns well with the vehicle's inertia, enhancing overall performance. In the discussion part, the impact of frame size will be further discussed. Given the defined "frame size", the HOG feature extraction process includes two main steps to generate HOG features for every frame, as illustrated in Figure 6-4. The first step of HOG extraction is to compute the magnitude $|\nabla f(x,y)|$ and orientation (angle) $\theta(x,y)$ of the gradient $|\nabla f(x,y)|$. The second step of HOG is to derive the orientation histogram from the orientations and magnitudes. Then, followed by a magnitude normalization, a set of HOG features is able to be obtained.

The 2D FFT is a critical tool for transforming spatial domain data into the frequency domain, revealing the features in the frequency domain, which has been utilized for feature extraction (Angadi & Kagawade, 2017; Liu et al., 2016; Sandirasegaram & English, 2005). In the field of GPR, frequency-domain analysis is essential for signal processing (Tzanis, 2015; Q. Yu et al., 2023) and plays a significant role in various applications, including water content

estimation (Benedetto, 2010b) and object detection (Zhang et al., 2015). With the rise of machine learning and deep learning techniques, 2D FFT has become increasingly valuable for identifying patterns in the frequency domain (Barkataki et al., 2023; Zhang et al., 2015). In our proposed method, frequency features extracted from GPR images using 2D FFT complement the texture features obtained through HOG. These frequency features are particularly advantageous as they exhibit strong resilience to variations as well. By incorporating frequency domain information, we effectively expand the feature space, thereby enhancing the overall performance of our method.

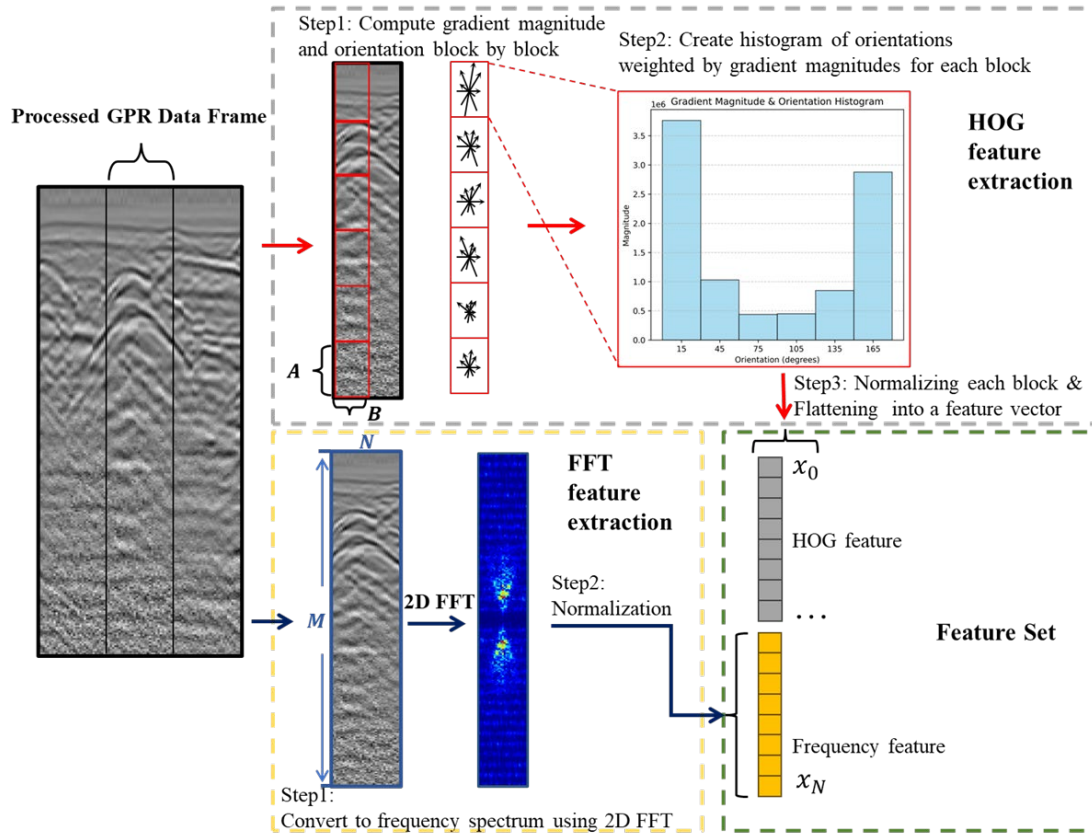


Figure 6-4 Illustration of underground feature extraction by HOG and 2DFFT

2D FFT is employed as a method for converting a frame of MCGPR image from the time domain to the frequency domain, as illustrated in Figure 6-4. FFT features the transformation of a window with size $(M \times N)$ can be denoted as:

$$F(u, v) = \sum_{x=0}^M \sum_{y=0}^N f(x, y) e^{-2\pi i \left(\frac{ux}{M} + \frac{vy}{N} \right)}, \quad [6 - 2]$$

where $f(x, y)$ denotes the intensity value of the image at the spatial coordinates (x, y) . The variables u and v correspond to the spatial frequencies along the x- and y-axes, respectively. To facilitate feature extraction, each frame of GPR data is divided into smaller cells, represented as black squares. Two primary types of features—HOG and 2D FFT—are extracted from these cells and subsequently combined to form a unified feature set for each frame of the GPR image.

For the HOG features, the gradient magnitude and direction are computed at each pixel, which is visually represented by black arrows. These gradient magnitudes are then binned into histograms based on their orientations within individual cells, shown as red squares. The histograms are normalized to ensure consistent representation and are included as part of the feature set. Similarly, for the 2D FFT features, the FFT is applied to each cell to compute the corresponding FFT magnitude spectrum. The resulting spectrum provides frequency-domain characteristics of the cell and is also included in the feature set. By combining the normalized HOG features with the 2D FFT spectrum features, a comprehensive feature set is constructed for each frame of the GPR image, capturing both spatial and frequency-domain information.

6.1.2.4 Underground Features Pairing

To quantify the similarity between the feature sets derived from MCGPR images, correlation functions were employed. The correlation between the two feature sets of two frames (reference and to-be-correct), A and B , is expressed mathematically as:

$$C(A, B) = \frac{\sum_{i=1}^N (A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^N (A_i - \bar{A})^2} \cdot \sqrt{\sum_{i=1}^N (B_i - \bar{B})^2}}, \quad [6 - 3]$$

where N is the total number of features from one frame of GPR data. A_i and B_i are the values of the i^{th} feature in the feature vectors of A and B , respectively. And $\bar{A}$ and $\bar{B}$ are

the mean values. Figure 6-5 illustrates the details of the feature pairing function applied to a single frame. For a given frame t with data from n channels, two feature sets from different swaths are compared to generate an $n \times n$ correlation matrix, C_t . Here, $C_t(i, j)$ represents the normalized correlation coefficient of extracted feature sets from channel i of the swath (to-be-corrected) and channel j of the swath (reference). To determine the optimal pairing of features in this frame, the function identifies the path along the matrix's diagonal with the highest mean value of correlation coefficients. The positioning offsets are indicated by diagonal offsets in correlation matrices. C_t can be simplified as a correlation array C_t^d , represented as:

$$C_t^d = \frac{1}{n - |k|} \sum_{i=1}^{n-|k|} C_t(i, j) \cdot \delta_{j-i, k}, \quad [6 - 4]$$

$$\delta_{x,y} = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}, \quad [6 - 5]$$

where $k \in \{-n, -(n-1), \dots, 0, n-1, n\}$, represents the diagonal offsets in the correlation matrix C_t . The diagonal offsets represent the best alignment of features based on similarity, representing the most likely relative positions between the two swaths at frame t .

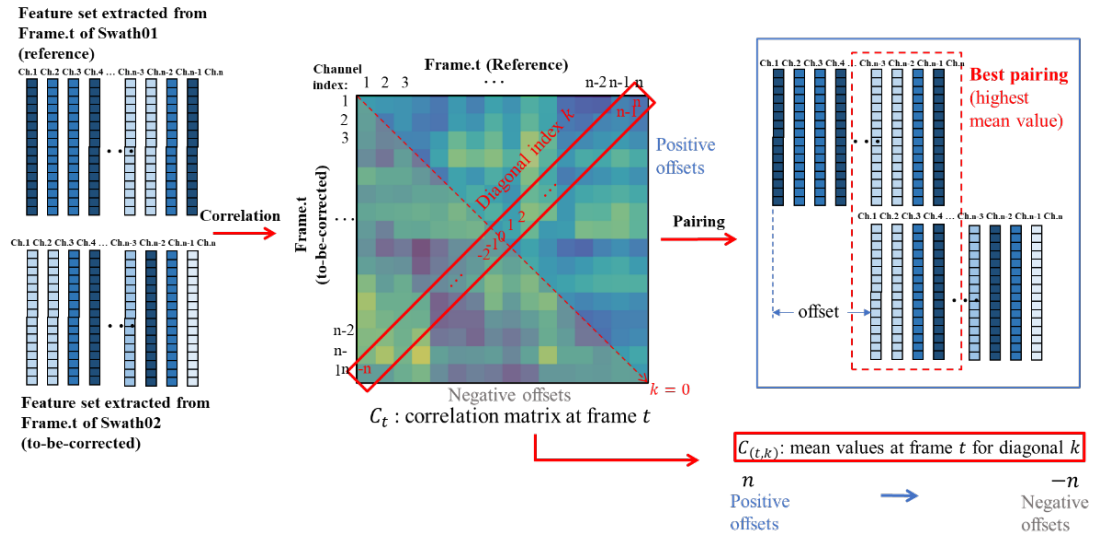


Figure 6-5 Illustration of the underground feature pairing function

Although the relative offsets estimated by correlation at each frame provide local

alignment, they may occasionally produce errors because these positions are estimated locally. For trajectory reconstruction, such local best alignments may result in false matches when viewed globally. To address this issue, a global optimization using a dynamic programming (DP) method is applied. This method integrates sequences of correlation matrices for trajectory reconstruction, as detailed in the next section.

6.1.2.5 Trajectories Reconstruction

The sequence of correlation matrices obtained from underground feature pairings provides a basis for identifying the most likely positions at each frame during trajectory reconstruction. However, when retrieving trajectories based on GPR signals, a significant challenge arises. GPR signals tend to be highly similar when survey swathes are close to each other. While this similarity is useful for feature matching, it can lead to ambiguities, as similar GPR signals are not always captured from the same location. High similarity between GPR images may also occur in distinct areas where the electromagnetic properties of materials along different swathes are highly similar. This ambiguity introduces the risk of false matches, where geometrically inconsistent correlations are mistakenly interpreted as valid positional correspondences.

To mitigate such ambiguities, DP, a global optimization method, was adopted to reconstruct trajectories based on correlation matrices. DP is a powerful optimization technique that solves complex problems by breaking them down into simpler subproblems and solving each subproblem only once, storing the solutions for future reference. This recursive approach makes DP particularly effective for problems with overlapping subproblems and optimal substructure, such as trajectory optimization. Numerous studies have demonstrated that DP excels in trajectory optimization tasks, including those involving flight dynamics (Harada & Miyazawa, 2013), robotic path planning (Lee & Lee, 2022), and road trajectory generation, due

to its ability to consistently find cost-optimal solutions (Abbasov & Gorbunova, 2025) even in challenging environments. In trajectory optimization, DP can be extended to incorporate various constraints, such as those derived from driving behaviours, to ensure the resulting paths are both feasible and realistic. By embedding constraints, DP can generate trajectories that not only optimize a global objective but also adhere to the practical and safety requirements of real-world driving scenarios (Wei et al., 2017).

In the study, trajectory reconstruction was achieved by calculating a series of offsets between the to-be-corrected swath and the reference swath frame by frame. Figure 6-6 is a conceptual illustration designed to demonstrate how DP operates in this context. The yellow dashed line in the figure represents an initial trajectory reconstruction, which connects the best pairing points in each frame based on the correlation matrices. However, due to ambiguities in GPR signal similarity, this naive approach can result in unrealistic trajectories, such as the extreme turns observed in the middle of the path. DP will help optimize the reconstructed trajectories, like the green dashed line generated by DP with a car motion constraint, which is more possible as a car driving trajectory during the MCGPR data acquisition.

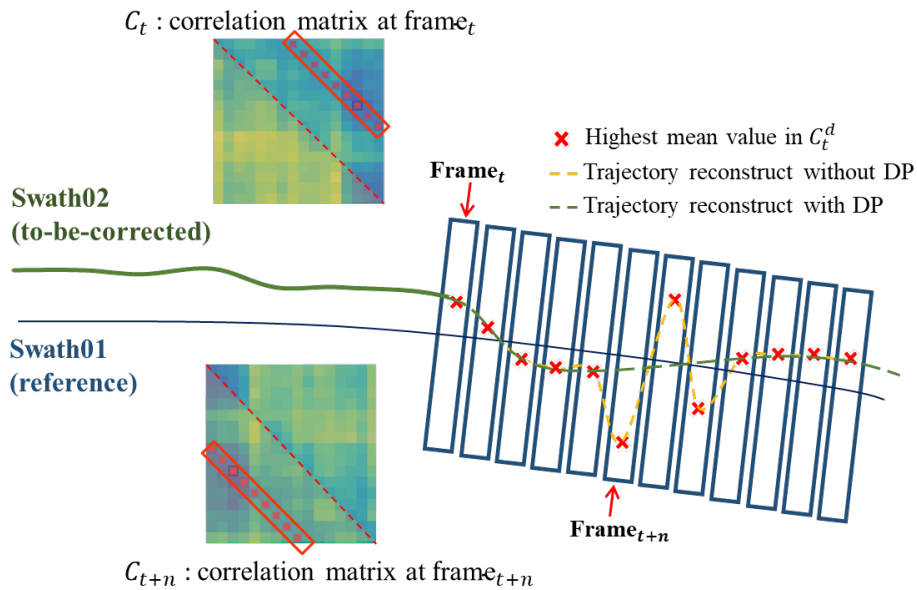


Figure 6-6 Illustration of the DP optimization processing

The constraint of DP in this study is the turning angle constraint derived from the driving scheme used for MCGPR data acquisition, detailed in Section 6.1.3. The turning angle constraint limits the angular change between the headings of consecutive trajectory points. Consider a trajectory consisting of a sequence of M points, $\{p_1, p_2, \dots, p_M\}$, the turning angle constraint for two consecutive points p_i and p_{i+1} can be represented as:

$$|\theta_{i+1} - \theta_i| \leq \theta_{max}, \quad [6 - 6]$$

where θ_i and θ_{i+1} are the headings at the points p_i and p_{i+1} , respectively, and θ_{max} is the maximum allowable turning angle. In this study, the turning angle constraint θ_{max} is defined as the turning angle observed in the components of the recorded reference trajectories in the raw trajectory classification stage.

The objective is to reconstruct the most probable path that the vehicle has traversed based on the similarities of underground features, represented by a sequence of correlation matrices, while satisfying the turning angle constraint. Given a sequence of $C_t^d (t = 1, 2, \dots, T)$, the DP optimization objective is maximizing the total sum of correlation coefficients while satisfying the turning angle constraint. Let value function $V(t, k_t)$ represent the maximum sum of correlation coefficients up to frame t , where the selected index at frame t is k . The state transition is defined as:

$$V(t, k_t) = C_t^d[k_t] + \max_{\substack{k_t \in \{-n, \dots, 0, \dots, n\} \\ |k_t - k_{t-1}| \leq \theta_{max}}} V(t-1, k_{t-1}), \quad [6 - 7]$$

where j is constrained by $|k - j| \leq \theta_{max}$ to represent the turning angle constraint. During the DP optimization, the initialization for the first frame ($t = 1$):

$$V(1, k_1) = C_1^d[k_1], k_1 \in \{-n, \dots, 0, \dots, n\}. \quad [6 - 8]$$

The optimization objective is:

$$\max_{k \in \{-n, \dots, 0, \dots, n\}} V(T, k_T). \quad [6 - 9]$$

Based on the value function and objectives, a series of $k_t (t = 1, 2, \dots, T)$ recursively recovers the trajectory by selecting k_t at each frame, allowing us to optimize the trajectory for the highest cumulative correlation while enforcing turning angle constraints. The DP algorithm works in the following pseudocode to reconstruct trajectories by optimizing the highest correlation with constraints.

Algorithm: Optimal Trajectory Reconstruction with Turing Angle Constraint

1. Input:
2. C : Sequence of correlation arrays $\{C_t^d\}$
3. θ_{max} : Maximum allowable turning angle (constraint)
4. T : Total number of frames
5. n : Number of channels
6. Output:
7. P : Optimal trajectory indices $\{k_t\}$ for frames $t = 1, 2, \dots, T$

8. Procedure:
9. **a. Initialize DP Table and Path Tracker:**
10. Create array $V[1 \dots T][1 \dots 2n + 1]$ to store the maximum cumulative correlation coefficients.
11. Create array $Path[1 \dots T][1 \dots 2n + 1]$ to store the backtracking indices.
12. For k from 1 to $2n + 1$:
13. $V[1][k] \leftarrow C_1^d[k]$ # Initialize the first frame with correlation values
14. $Path[1][k] \leftarrow Null$ # No previous frame for the first frame
15. End For
16. **b. Fill DP Table:**
17. For t from 2 to T # Iterate over frames
18. For k_t from 1 to $2n + 1$: # Iterate over possible indices at frame t
19. $V[t][k_t] \leftarrow -\infty$ # Initialize as negative infinity
20. For k_{t-1} from 1 to $2n + 1$: # Iterate over previous frame indices
21. If $IsFeasible(k_t, k_{t-1}, \theta_{max})$:

Ensure the turning angle constraint is satisfied
22. If $V[t-1][k_{t-1}] + C_t^d[k_t] > V[t][k_t]$:

Update maximum cumulative correlation for the

current frame

```
23.           $V[t][k_t] \leftarrow V[t-1][k_{t-1}] + C_t^d[k_t]$ 
24.           $Path[t][k_t] \leftarrow k_{t-1}$     # Store the backtracking index
25.      End If
26.  End If
27. End For
28. End For
29. End For
30. c. Find the Optimal Solution:
31.   Let  $k_T \leftarrow \operatorname{argmax} (V[T][k])$ 
      # Find the index at frame T with the highest score
32.   Let  $P \leftarrow \text{BacktrackOptimalPath} (Path, k_T, T)$ 
      # Retrieve the trajectory using backtracking
33.   Return  $P$ 

34. Function IsFeasible ( $k_t, k_{t-1}, \theta_{max}$ )
35.   Input:
36.      $k_t$ : Current index
37.      $k_{t-1}$ : Previous index
38.      $\theta_{max}$ : Maximum allowable turning angle
39.   Output:
40.     Boolean: True if the turning angle constraint is satisfied, False otherwise
41.
42.      $\Delta k \leftarrow |k_t - k_{t-1}|$ 
43.     Return ( $\Delta k \leq \theta_{max}$ )
44.
45. Function BacktrackOptimalPath ( $Path, k_T, T$ )
46.   Input:
47.      $Path$ : 2D array storing backtracking indices
48.      $k_T$ : Index of the optimal trajectory at the last frame
49.      $T$ : Total number of frames
50.   Output:
51.     P: Sequence of indices  $\{k_1, k_2, \dots, k_T\}$  representing the optimal trajectory
52.
53.     Create an empty stack  $P$ 
```

```

54.           $i \leftarrow T, k \leftarrow k_T$ 
55.          While  $i \geq 1$ :
56.               $P.\text{push}(k)$ 
57.               $k \leftarrow \text{Path}[i][k]$     # Move to the previous frame's index
58.               $i \leftarrow i - 1$ 
59.          End While
60.          Return  $P.\text{reversed}$ 

```

6.1.3 Experiment Results and Discussion

6.1.3.1 Experiment Sites and Results

In general, extensive coverage of the survey area is critical for ensuring the reliability of findings, for underground utilities mapping, and hazard detection with MCGPR. To achieve comprehensive coverage, the MCGPR system is required to survey multiple times along the road, known as swathes, over each lane—typically two to three swathes with predetermined offsets. The overlapping region between two adjacent swathes often exists, as shown in Figure 6-7. Such overlapping is essential for maintaining the continuity and integrity of the collected data. To geo-locate the MCGPR measurement, the positioning system always connects with the MCGPR system, as illustrated in Figure 6-7. The MCGPR antennas are arranged in an array of multiple transmitter-receiver pairs aligned to span a swath perpendicular to the direction of travel, facilitating data acquisition across a broad area. Simultaneously, the positioning system generates precise location data at regular intervals. The precise placement of data is essential when employing an array of antennas, necessitating positional accuracy that surpasses the minimal channel offset to ensure the discernibility of the antenna.

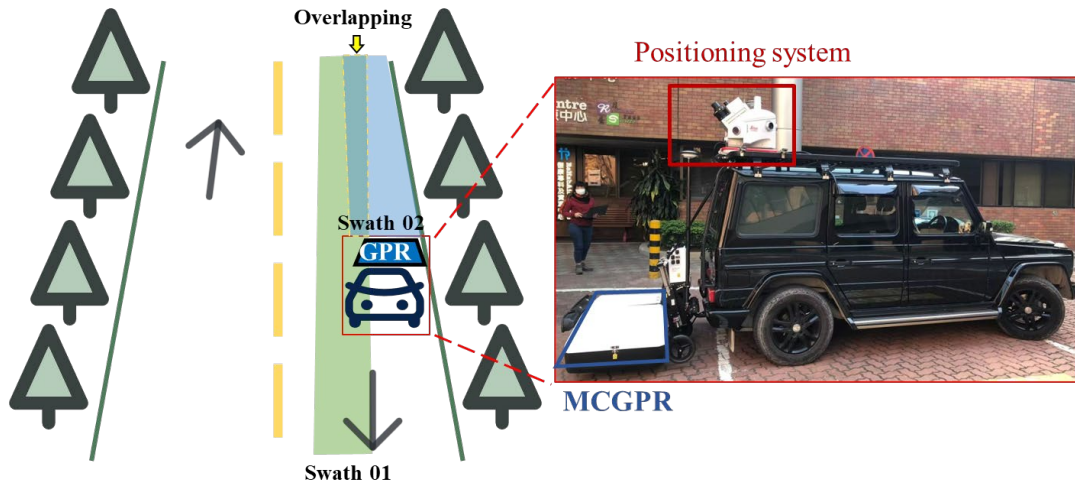


Figure 6-7 Illustration of MCGPR road survey and the components of MCGPR system: MCGPR antennas and position system

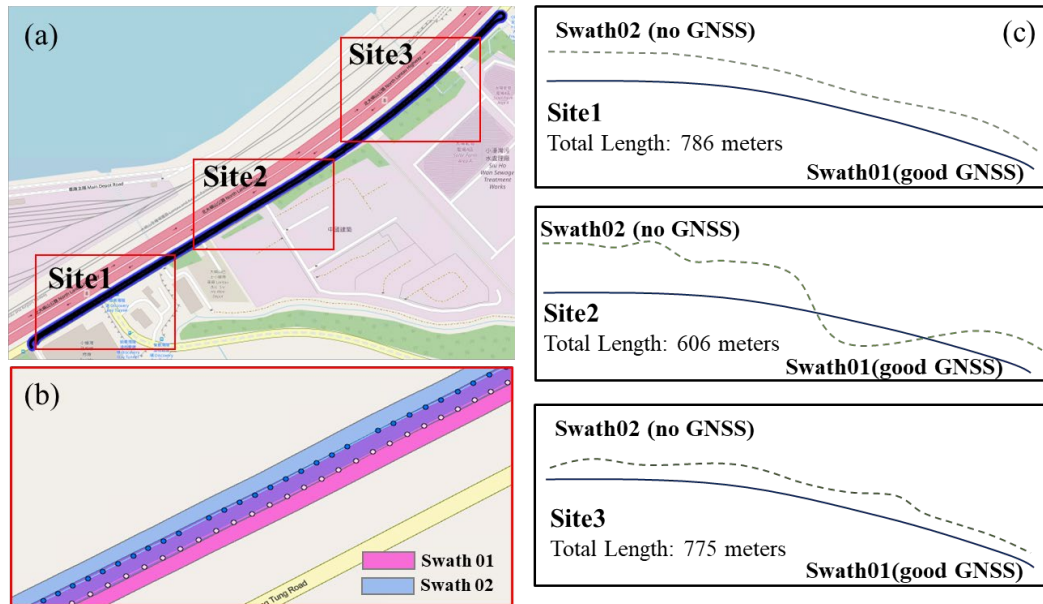


Figure 6-8 (a) Overview of the three validation sites used in the study; (b) Illustration of the overlapping region between Swath 01 and Swath 02, highlighting their spatial coverage; (c) Trajectories of the three sites, showing the reference paths and the masked sections for validation

In this study, three test sites, as shown in Figure 6-8 (a), were selected to validate the accuracy of the proposed positioning enhancement method. The experiment utilized a combination of the Raptor45 ground-penetrating radar system and GNSS-RTK. The Raptor45 system is equipped with 18 channels of 450 MHz impulse antennas, while the GNSS-RTK receiver, mounted alongside the Raptor45, provided centimetre-level accuracy (approximately

4 cm) as a benchmark for ground truth reference. Both MCGPR measurements and trajectory data were effectively captured during the experiment. To assess the performance of the proposed method, portions of the trajectories—specifically the middle sections—were deliberately masked to simulate poor-quality or missing trajectory data. These masked sections were treated as unknowns requiring enhancement using the proposed approach. Figure 6-8 (b) illustrates the swath coverage by Swath 01 and Swath 02, providing a clear representation of the spatial arrangement of the captured data. Additionally, Figure 6-8 (c) presents a comparison of the reference trajectories and the masked to-be-corrected trajectories. The masked sections, where positioning information was removed, serve as the basis for validating the effectiveness of the proposed method in reconstructing and improving trajectory accuracy.

The MCGPR data was recorded based on distance, while the GNSS-RTK positioning data was recorded based on time. To synchronize these two datasets for analysis, resampling was required. In the test sites, the GNSS-RTK system operated at a sampling rate of 10 Hz. With an average driving speed of 30 km/h, the GNSS-RTK system recorded one position approximately every 0.83 m. To ensure that the reconstructed trajectory provided sufficient spatial resolution for comparison with the GNSS-RTK data, the MCGPR data were processed using a frame size of 16 pixels, with the odometer triggering step configured to 0.04 m. Using the proposed algorithm, the spacing between consecutive reconstructed points was maintained at 0.64 m, which is smaller than the 0.83 m interval of the GNSS-RTK system, ensuring a higher spatial resolution in the reconstructed trajectory. To align the reconstructed trajectories with the GNSS-RTK reference data, the GNSS-RTK positions were interpolated using the two nearest recorded points. This interpolation created a synchronized mapping of true GNSS-RTK positions to the reconstructed trajectories at corresponding timestamps. The resulting mapping facilitated the evaluation of relative location accuracy by allowing direct comparisons between the reconstructed trajectories and the reference trajectories. Mean absolute error (MAE) and

root mean square error (RMSE) are employed as the error metrics. Given a trajectory consisting of a sequence of M points, $S = \{p_1, p_2, \dots, p_M\}$, MAE and RMSE are denoted as:

$$MAE_s = \frac{\sum_{i=1}^M |p_i - p_i^{ref}|}{M}, \quad [6 - 10]$$

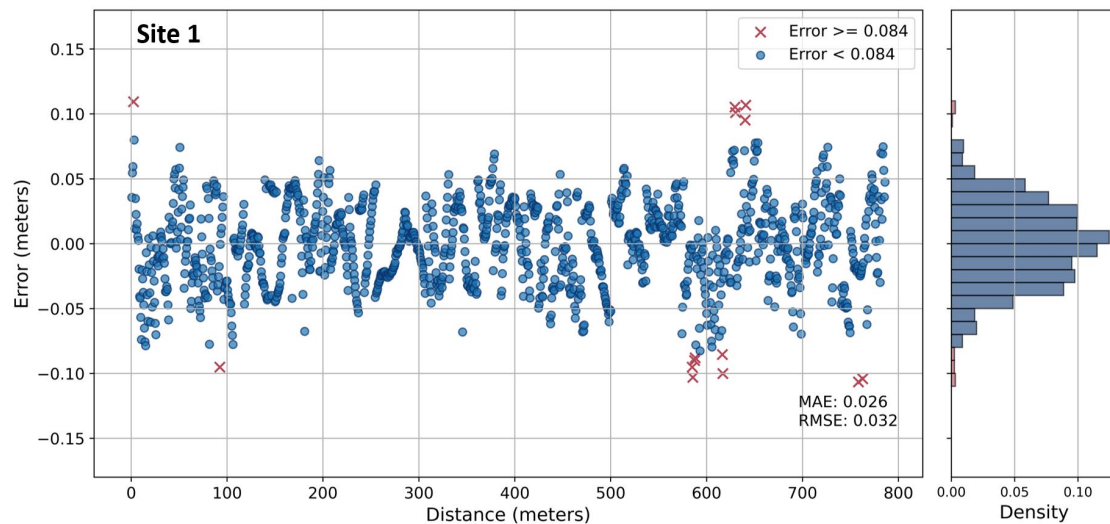
$$RMSE_s = \sqrt{\frac{\sum_{i=1}^M (p_i - p_i^{ref})^2}{M}}. \quad [6 - 11]$$

Table 6-1 presents the accuracy results of the proposed method across the experimental sites, showing an overall MAE of approximately 0.036 meters and an RMSE of around 0.047 meters. This level of accuracy meets the requirements for most MCGPR road survey applications, demonstrating the method's strong positioning performance. Notably, at Site 1 and Site 3, the method achieved an MAE close to 0.03 meters, reflecting high precision and effectiveness. The close alignment of RMSE values with MAE further underscores the method's consistency and reliability, as illustrated in Figure 6-9. This consistency suggests a well-behaved error distribution, with errors following a normal distribution and minimal outliers at Site 1 and Site 3. To assess positioning precision relative to channel separation, a threshold of 0.084 meters, equivalent to the separation distance between antennas, was set. In Figure 6-9, a graphical analysis of error distributions across the three test sites confirms the method's robustness: over 95% of samples maintained absolute errors below the 8.4 cm threshold (visualized as blue circles and bars in the histograms), with red crosses marking instances of relatively poorer performance. However, results for Site 2 show higher MAE and RMSE values compared to the other sites, with more samples exceeding the threshold. As shown in Figure 6-8 (c), Site 2 was designed as a special case where the vehicle crossed the previous measurement path, unlike typical conditions where drivers maintain alignment with prior measurements to ensure overlapping regions. The error distribution pattern across the three sites suggests that errors are not random but rather follow an oscillating trend, particularly pronounced at Site 2.

Red crosses frequently appear at points where the error shifts from increasingly negative to increasingly positive values or vice versa, exceeding the threshold at each extreme. This pattern suggests that trajectory characteristics contribute significantly to error sources. While overall positioning accuracy remains high, recognizing the limitations and uncertainties inherent in the proposed method is essential to explore. A detailed analysis of these error sources and method constraints is provided in the next section.

Table 6-1 MAE and RMSE of three test sites

	Length	MAE	RMSE
Site1	786 meters (1251 samples)	0.026 meter	0.032 meter
Site2	606 meters (965 samples)	0.052 meter	0.071 meter
Site3	775 meters (1233 samples)	0.031 meter	0.039 meter



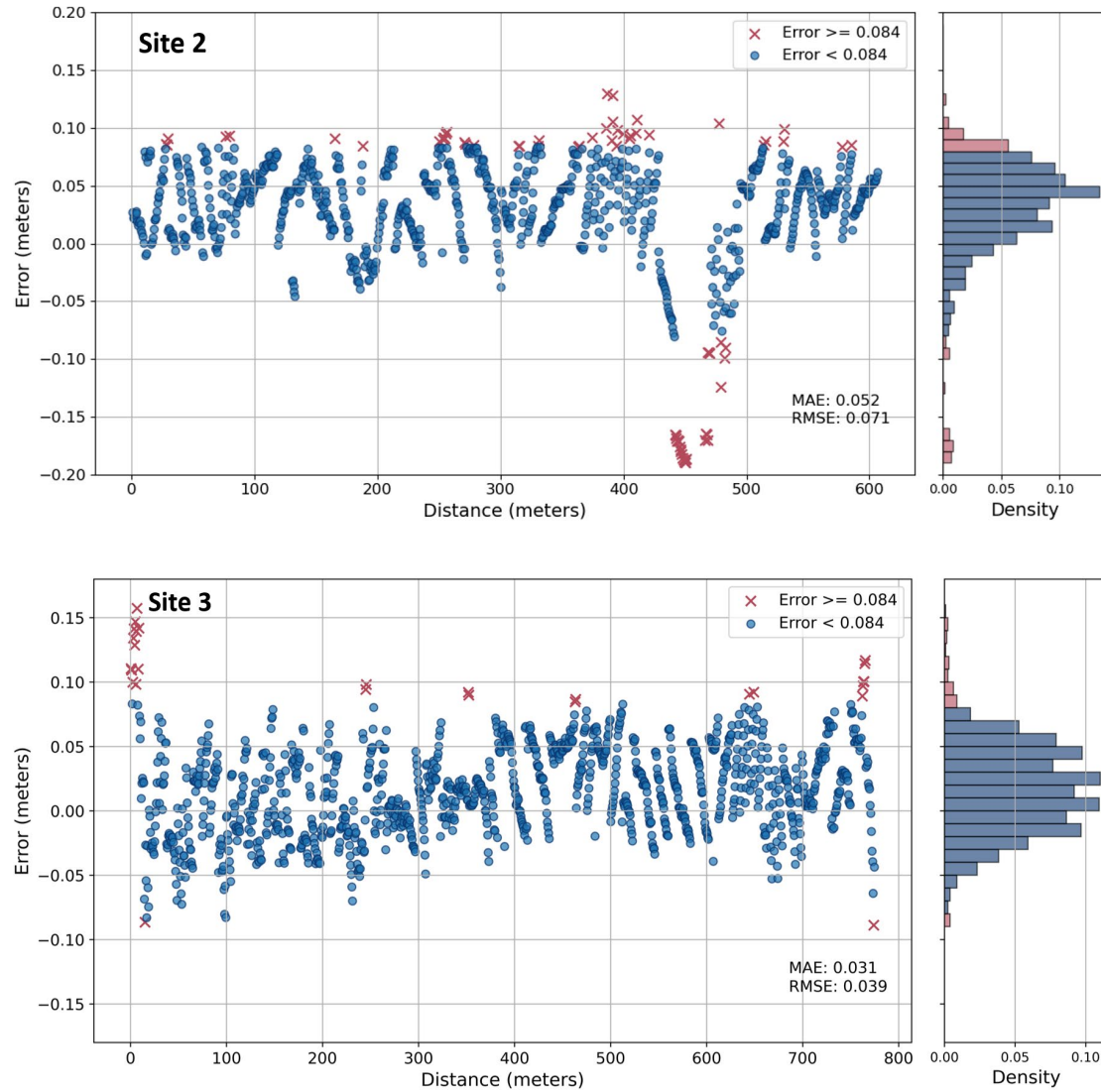


Figure 6-9 The error distributions of three test sites with corresponding MAE and RMSE

6.1.3.2 Discussion

This study investigates positioning enhancement for MCGPR road surveys and introduces, for the first time to our knowledge, a specialized method that leverages cross-channel MCGPR measurements. The proposed approach incorporates tailored methods for underground feature extraction and pairing, enabling accurate positioning, anchoring, and trajectory reconstruction. To further refine results, a global optimization framework constrained by driving behaviour was designed to filter outliers and smooth trajectory paths. The effectiveness of this methodology was empirically validated using data from three real-world

test sites, demonstrating promising results. However, several critical factors that influence performance warrant further exploration.

One key factor is the frame size, which defines the segment length of GPR images used as the basic unit for underground feature extraction. Frame size directly impacts the accuracy of feature pairing and overall positioning performance. A second factor is turning angles, with Site 2 exhibiting significantly more frequent and sharper angle changes compared to other sites, which posed challenges for trajectory reconstruction. Another important consideration is the overlapping rate between adjacent swathes, which is essential for cross-channel feature matching. The influence of overlapping rates on positioning enhancement remains unclear and requires additional analysis. Finally, the availability of reliable position measurements, particularly in challenging environments, plays a crucial role in the optimization process. Incorporating such measurements has been shown to significantly improve trajectory reconstruction accuracy, as observed in specific scenarios.

These factors—frame size, turning angles, overlapping rates, and reliable positioning measurements—affect performance in distinct ways and are analyzed in this section. The discussion begins with a sensitivity analysis of frame size, followed by an evaluation of how turning angles influence trajectory reconstruction, focusing on Site 2's error sources. Next, the impact of overlapping rates is assessed using data from Site 1. Finally, the role of reliable positioning measurements in challenging environments is explored, with Site 2 serving as a case study. This structured evaluation provides deeper insights into the factors affecting MCGPR positioning enhancement and highlights areas for future improvement.

(1). Impact of frame size on positioning performance

The choice of frame size is essential for balancing accuracy and robustness in the proposed method, as it determines the spatial resolution of the features used in positioning.

Frame size refers to the length of GPR signals grouped as a single unit for feature extraction, where one pixel corresponds to 0.04 meters, as along-track resolution set in Raptor45. To assess its impact, we conducted sensitivity testing with frame sizes ranging from 16 to 512 pixels, using RMSE as the evaluation metric across data from the three test sites. As shown in Figure 6-10, RMSE values gradually increase when the frame size exceeds 64 pixels, with a marked performance drop for Site 2 when using frame sizes above 256 pixels. This trend suggests that larger frame sizes may smooth out important localized details, reducing the method's responsiveness to abrupt changes in vehicle trajectory. Interestingly, the method performed consistently well with frame sizes between 16 and 64 pixels, balancing noise resistance and spatial precision. This range likely aligns with the vehicle's inertial stability, as smaller frame sizes enhance accuracy by capturing more localized features—ideal when assuming a stable heading within a single frame. Despite the added computational cost, frame sizes below 64 pixels ($64 \times 0.04 = 2.56$ meters) are recommended to optimize performance. These findings emphasize the importance of selecting a frame size that accommodates both the survey environment and the vehicle's motion characteristics, ultimately improving the precision of positioning enhancement.

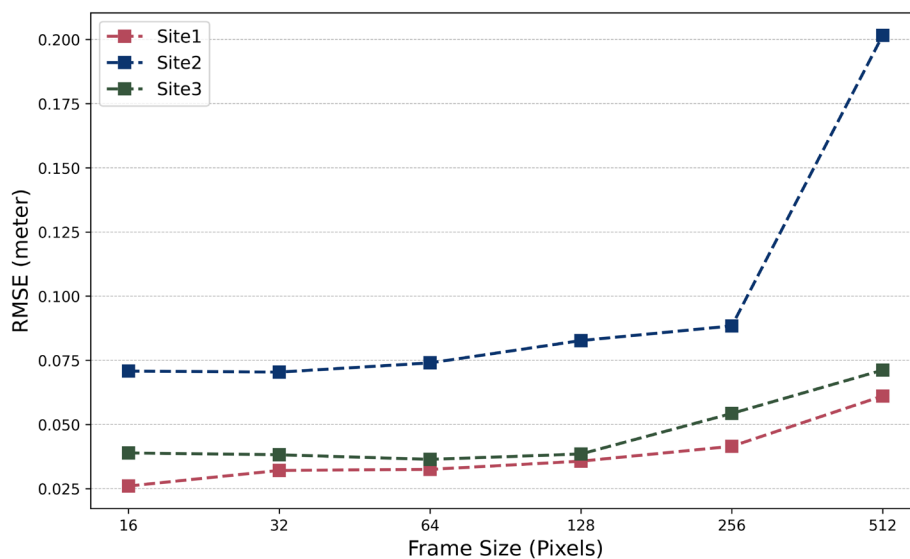


Figure 6-10 RMSE of three testing sites with various frame sizes

(2). Errors associated with turning angles in Site 2

Figure 6-11 presents a simplified trajectory map from Site 2, along with the comparison of error distribution between the proposed method and the true reference. In the map, the blue solid line represents the swath used as the reference, while the green lines denote the swath with poor-quality positioning segments, marked by dashed lines. When comparing the two plots, we observe that high errors are consistently associated with sections where there are significant changes in heading angle. As defined in the methods section, the turning angle represents the angular change between consecutive trajectory points. Unlike typical MCGPR surveys, where swathes are conducted in a parallel scheme to maintain consistent turning angles, Site 2 was intentionally designed with varied angles between swathes. This design difference results in substantial turning angle variations between the reference swath and the swath requiring correction. Minimal differences in turning angle allow trajectories to be approximated as parallel, enabling optimal feature extraction and pairing. However, as the turning angle difference increases, the method's accuracy diminishes. This degradation is primarily due to two factors: the distortion inherent in GPR imagery at high angles and a reduction in the overlap percentage between swathes, which reduces effective feature correlation.

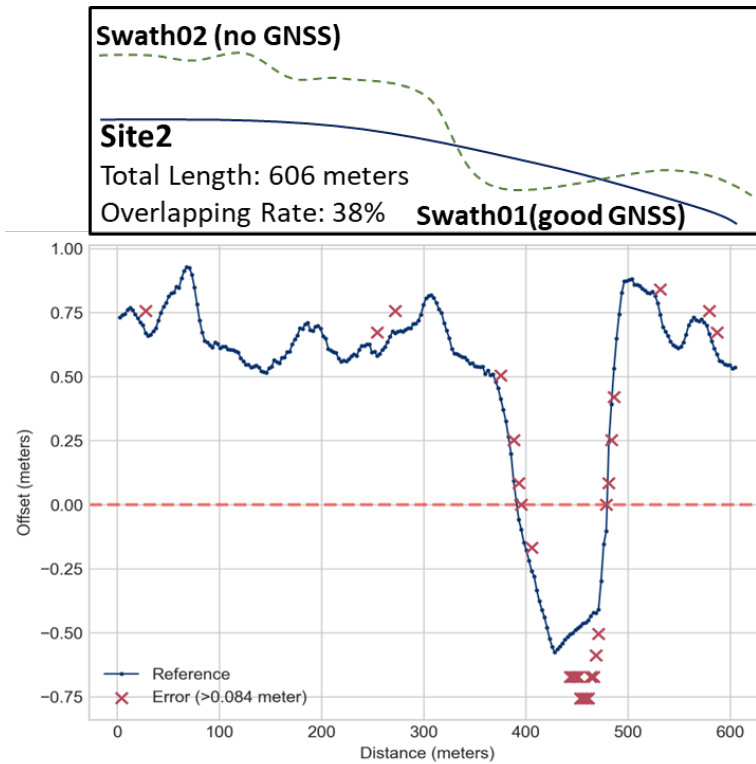


Figure 6-11 Errors distribution of testing Site2 with sketch of trajectories of paired swathes

To illustrate the impact of large turning angle differences between swathes, an example is shown in Figure 6-12. Here, a sample frame is represented by the data points between two array-like yellow crosses. In this example, the reference swath follows a straight path, while the swath needing correction undergoes a turn. This discrepancy introduces challenges for underground feature extraction and pairing from two main factors. First, MCGPR data distortion arises due to different positions along the arc of the turning swath. Since the same odometer records data across all channels, the data length captured is consistent, but it actually represents varying survey distances, as seen in the different lengths from Ch.1 to Ch.5 on Swath02. This effect results in lower resolution in the outer channels compared to the centre, degrading the feature pairing process fundamental to our method. Secondly, the turning angle difference creates a “transition zone”—a central region where feature pairing shifts between adjacent channels. For instance, as Swath02 turns, Ch.4’s pairing transitions from Ch.1 to Ch.2 on the reference Swath01, creating a “transition zone” marked in red.

Within this “transition zone,” systematic misalignment issues make it difficult to accurately pair MCGPR data with the reference swath. These two factors—data distortion from varying arc positions and pairing misalignment in the “transition zone”—emerge as primary causes of performance degradation when turning. In future work, a deeper analysis of these challenges is needed. Potential solutions include addressing data distortion and designing specialized correction methods for the “transition zone.”

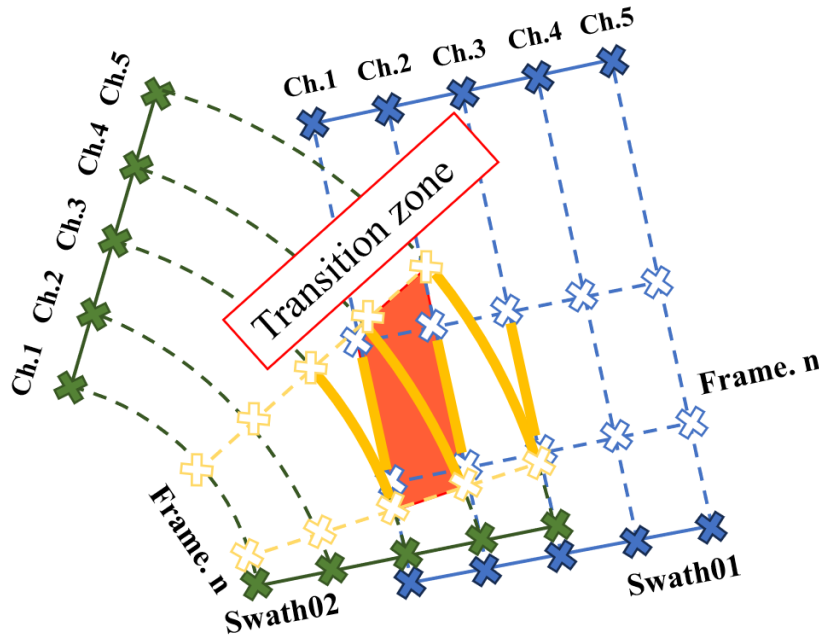


Figure 6-12 Illustration of the influence of turning angles

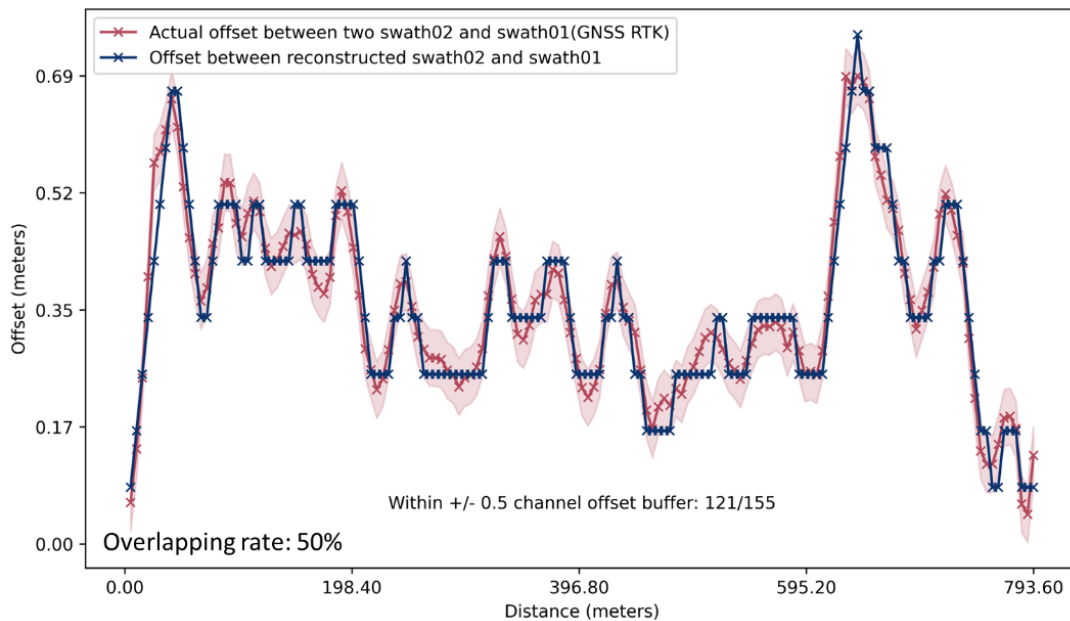
C. Influence of overlapping rates on positioning accuracy in Site1

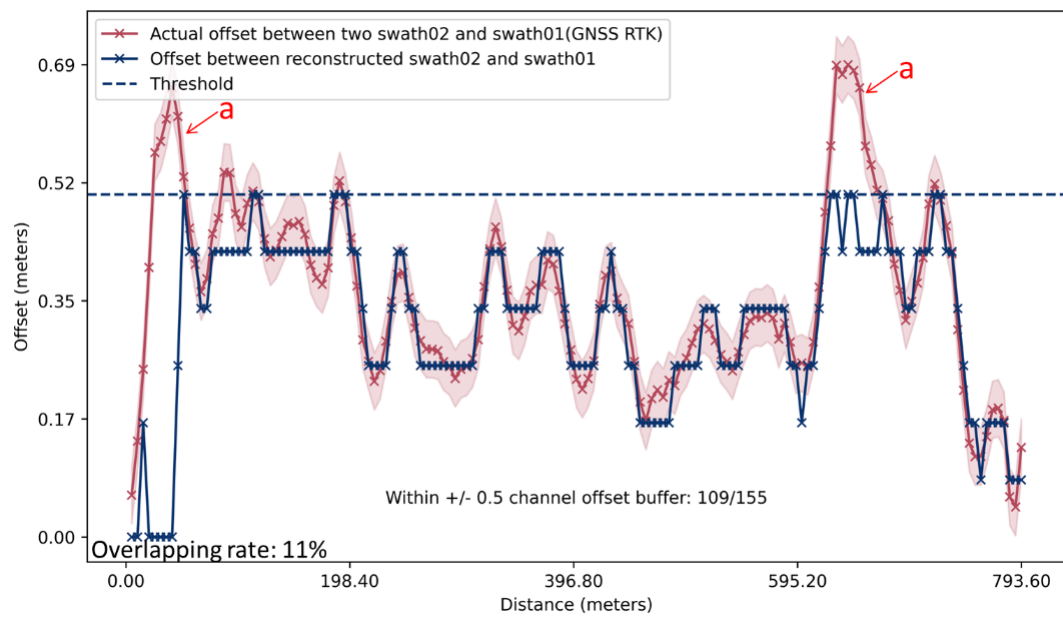
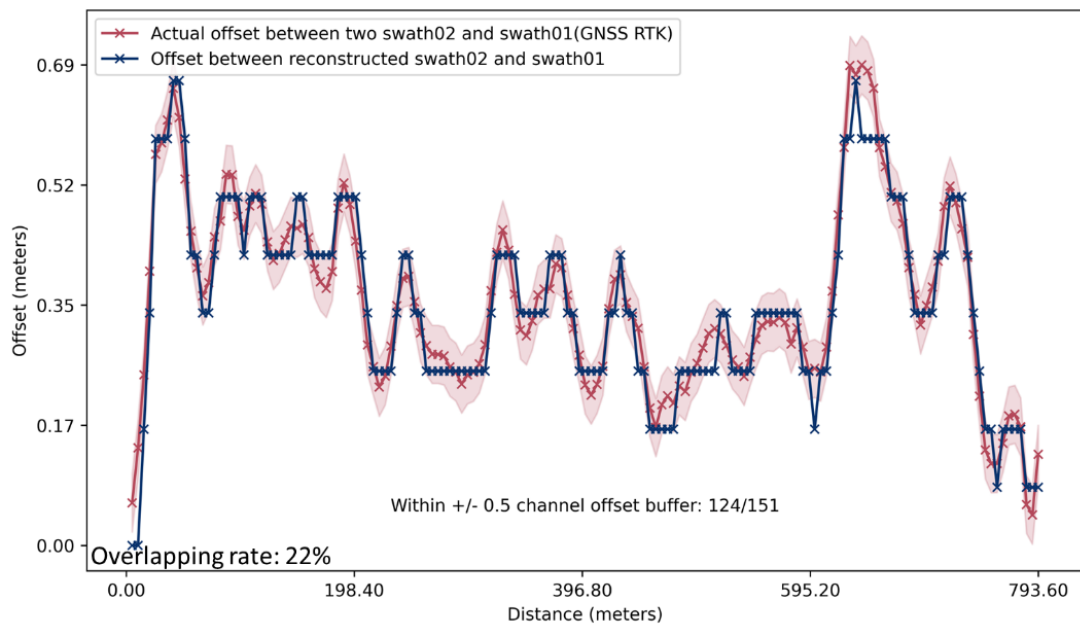
The overlapping rate between adjacent swaths is critical for accurate feature matching in the proposed method. To evaluate the impact of overlapping rates on positioning enhancement performance, the data from Site 1 was artificially degraded by masking channels, progressively reducing the overlap from 72% to 5% (equivalent to 1 of 18 channels). For instance, to simulate two swaths with a 50% overlap, the last four channels of data were masked, reducing the overlapping rate to 50%. Similarly, datasets with other overlapping percentages were created by masking additional channels. In the most

challenging case, the overlapping rate was reduced to 5%, leaving only one channel overlapping between adjacent swaths.

As illustrated in Figure 6-13, the reconstructed trajectories (blue lines) closely align with the ground truth (red buffers) across overlapping rates ranging from 72% to 5.6%. However, quantitative results in Table 6-2 demonstrate that the MAE increases as the overlapping rate decreases. For example, the MAE increases from 0.0286 m at a 72% overlap to 0.0337 m at a 5.6% overlap. While these errors remain within acceptable limits for real-world applications, challenges arise when the overlapping rate falls below 50%.

At low overlap rates, some swath regions lack sufficient coverage for effective feature matching, leading to invalid computations. For example, as marked by "a" in Figure 6-13, certain points fall outside the valid computation threshold due to insufficient overlap.





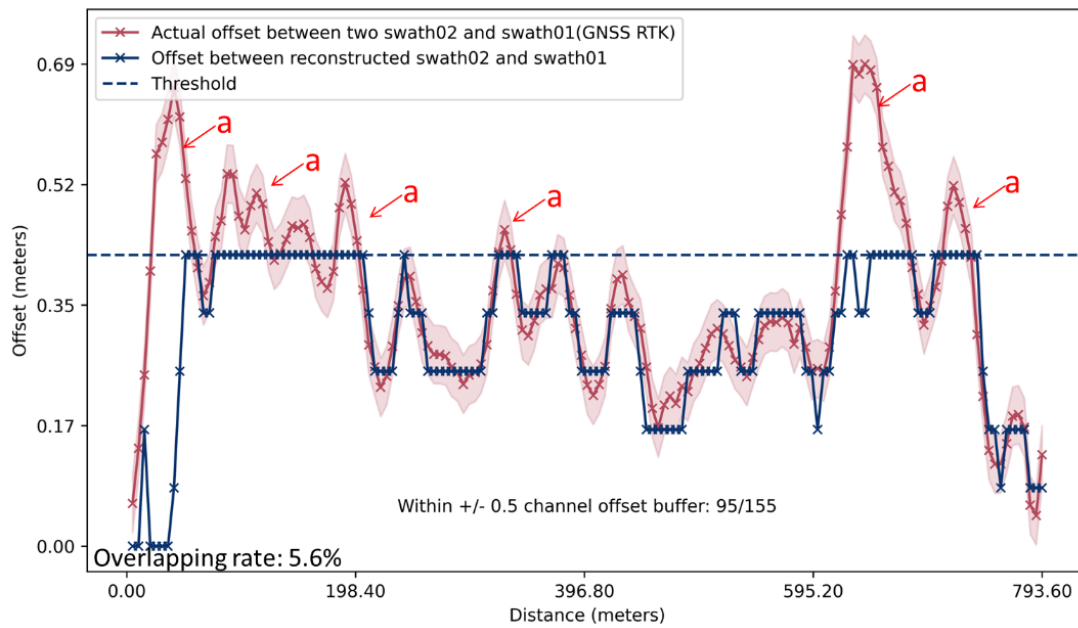


Figure 6-13 Impact of swath overlapping on trajectory reconstruction accuracy (Remark: Arrow 'a' marks the erroneous computations over the blue dotted lines in the last two cases)

Table 6-2 Results comparisons between different overlapping rates

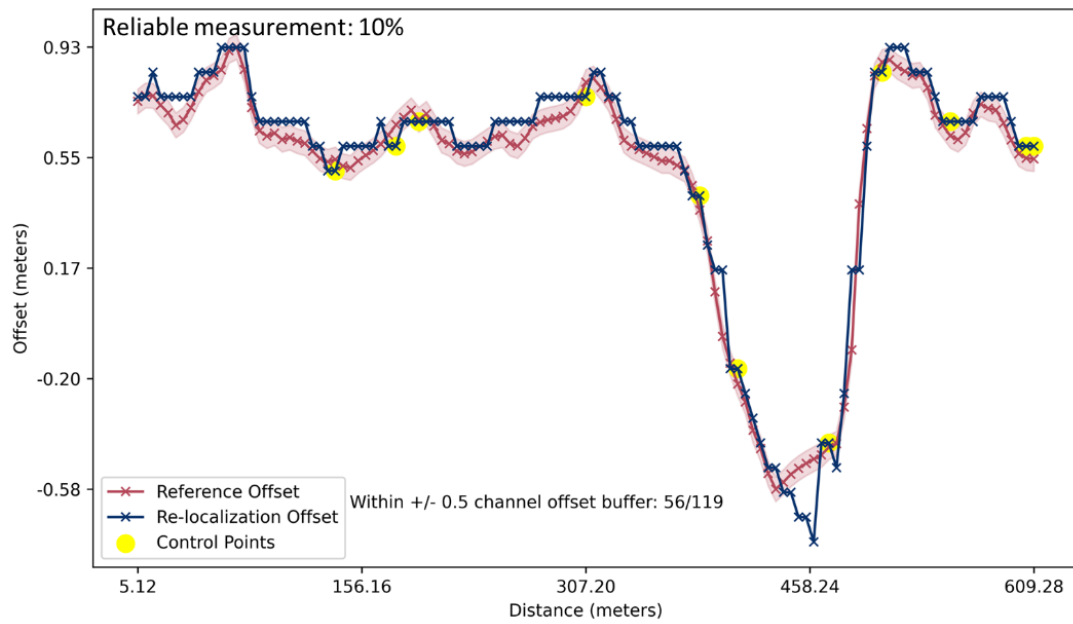
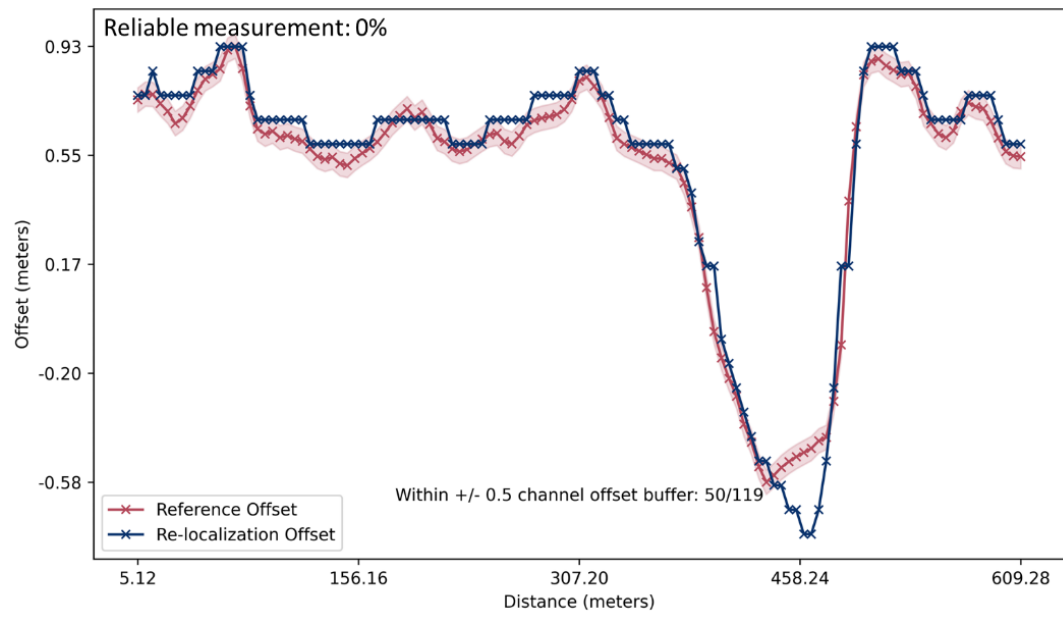
Overlapping Rate	Number of Samples	MAE (meter)	RMSE (meter)
72%	155	0.02863	0.03573
50%	155	0.02880	0.03592
22%	151	0.03221	0.05343
11%	135	0.03102	0.04877
5.6%	109	0.03366	0.05384

The findings demonstrate that reducing the overlapping rate decreases positioning accuracy, but the resulting degradation is generally within acceptable limits. For instance, the MAE increase from 3.86 cm to 3.37 cm is relatively minor and unlikely to significantly affect most practical applications. However, the larger relative increase in RMSE compared to MAE suggests that, while the majority of trajectory points maintain high accuracy, certain segments experience significant degradation under low-overlap conditions. This loss of overlap not only introduces localized errors but also allows these errors to propagate along the trajectory, ultimately reducing the overall quality of the reconstruction. Therefore, the

most critical takeaway is the necessity of maintaining sufficient overlapping regions—not only to ensure positioning accuracy at the current location but also to preserve the accuracy of adjacent points during trajectory reconstruction.

D. Impact of reliable positioning measurements in Site2

To evaluate the impact of incorporating reliable positioning measurements on positioning accuracy, a specialized dataset was designed to simulate challenging conditions. Unlike the datasets used in the results section—where all reliable positioning measurements were masked—Site 2’s data was processed with partial positioning coverage retained at varying levels (50%, 20%, and 10%) to emulate different scenarios with reliable measurements. This dataset represents a more challenging case, as it aims to verify whether the integration of reliable positioning points can effectively enhance performance in complex environments. Site 2 was selected due to its inherent complexity, characterized by sharp turning angles and higher baseline errors compared to other Sites. In such challenging settings, incorporating reliable positioning measurements proved to significantly improve trajectory reconstruction accuracy. As shown in Figure 6-14 and Table 6-3, the MAE for Site 2 decreased from 0.057 m to 0.035 m as the percentage of reliable positioning measurements increased from 0% to 50%. These reliable points, highlighted in yellow, were particularly effective when strategically placed in regions with significant errors, such as at $x = 458.24$, where they corrected trajectory trends and reduced overall errors. Conversely, control points placed in already accurate regions provided minimal benefits, underscoring the importance of strategic placement to maximize their impact.



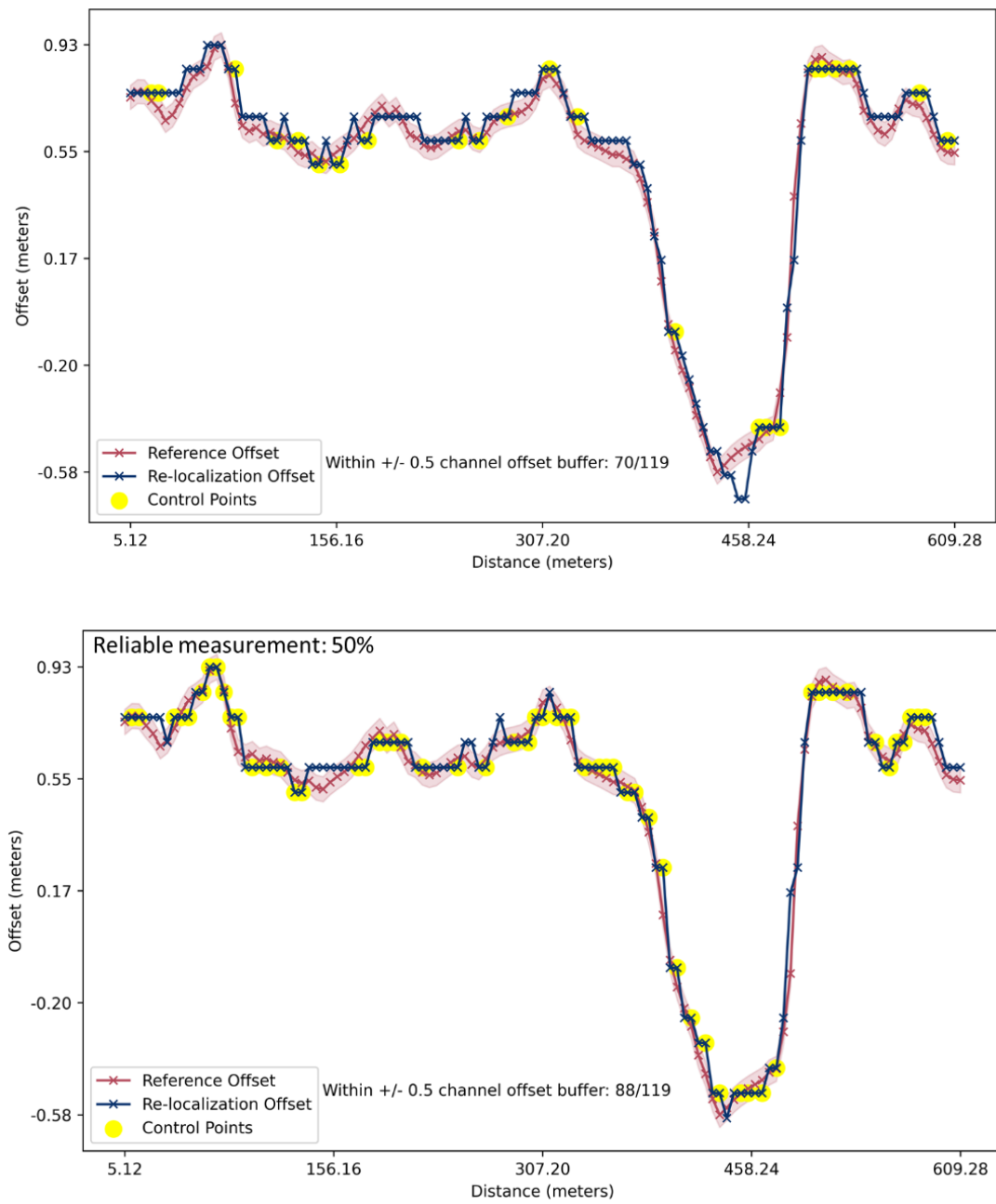


Figure 6-14 Impact of reliable positioning measurements on trajectory reconstruction accuracy

Table 6-3 Results comparisons between different reliable position measurements

Reliable Positioning (overlapping rate: 38%)	MAE (meter)	RMSE (meter)
0%	0.05711	0.07892
10%	0.05284	0.07084
20%	0.04231	0.05534
50%	0.03506	0.04969

Overall, these findings highlight that while the proposed method performs robustly even without additional reliable measurements, integrating such measurements can significantly enhance positioning accuracy in challenging environments. This improvement is especially critical when dealing with complex trajectories, as reliable measurements help correct localized errors and prevent error propagation throughout the trajectory.

6.1.3.3 Conclusion

Accurate, reliable, and robust positioning methods are essential for MCGPR surveys of road infrastructure, where precise trajectory information is critical for effective data analysis and interpretation. This study introduced a novel method to enhance positioning by leveraging cross-channel MCGPR data, marking the first attempt to use MCGPR data itself as an auxiliary source for positioning improvement. The method is designed to be compatible with various positioning systems, as it does not rely on specific positioning technologies but instead uses MCGPR data to enhance accuracy. Empirical validation at three real-world test sites demonstrated that the proposed method performs robustly under typical survey conditions (e.g., Sites 1 and 3) and offers significant accuracy improvements. However, Site 2, with its large turning angles and challenging trajectories, revealed key limitations that merit further investigation. Specifically, errors in Site 2 were linked to data distortion due to turning and systematic misalignment within transition zones. Additionally, the study highlighted the

influence of overlapping rates and reliable positioning measurements, showing that maintaining sufficient overlap and strategically incorporating reliable control points are critical for improving accuracy in complex scenarios.

Overall, the findings emphasize the proposed method's reliability and potential for enhancing MCGPR positioning in diverse road survey applications. Future research will focus on addressing the identified challenges, including reducing data distortion during turns and developing specialized correction methods for transition zones. Furthermore, optimizing overlapping rates and exploring advanced strategies for integrating reliable positioning measurements will further strengthen the method's applicability, particularly in challenging environments.

6.2 Harmonizing MCGPR Signals

Despite advancements in precise positioning enhancements for MCGPR systems, the collected GPR data often poses challenges in practical applications. One major issue is the presence of streaky or random noise in MCGPR-generated images. This noise predominantly arises from variations in antennas across individual channels due to hardware performance differences. These variations manifest as strip-like noise, especially in slices at shallow depths, which can be misinterpreted as subsurface features. To mitigate these issues, self-calibration techniques are employed, typically correcting the measured direct-coupling responses using S-parameters established during factory calibration (Cornick et al., 2016). Additionally, sparse blind deconvolution methods have been implemented to denoise MCGPR signals, ensuring consistent reflectivity across the array (Imai & Mizutani, 2023). However, limited research focuses on improving the quality of horizontal slices in MCGPR data volumes. This section introduces a specialized two-step post-processing workflow to harmonize MCGPR outputs. The proposed method addresses variations stemming from individual antennas and amplitude

inconsistencies between swathes.

6.2.1 MCGPR Harmonizing Method

Figure 6-15 illustrates a grayscale MCGPR amplitude slice at a depth of 0.24 meters (thickness: 0.1 meters). Clear strip patterns aligned with driving trajectories are visible, indicating variations caused by MCGPR antennas. These patterns are artefacts amplified by time-varying gain functions applied during data processing, mimicking linear underground features. The strip-like noise originates from several sources: 1. Inconsistencies in first arrival times: Time-zero points of air/ground wavelets can vary due to thermal drift, electronic instability, or antenna airgap changes. 2. Coupling effects: The interactions between antennas and their environment differ across the array. 3. Mechanical vibrations: Vehicle vibrations during data acquisition add further noise. 4. Antenna array dependencies: In array-based setups, discrepancies between individual antennas can exacerbate the noise due to mutual interdependence.

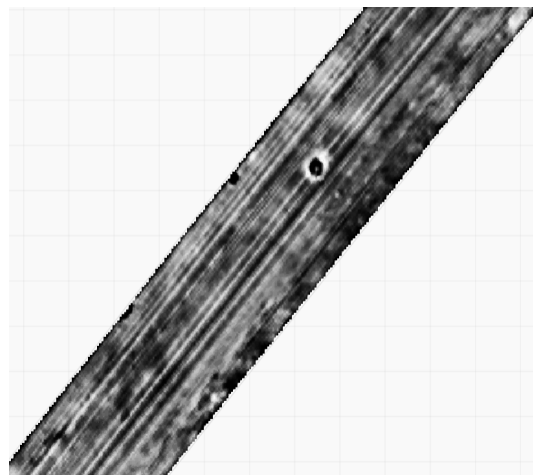


Figure 6-15 Example of typical strip-like noise comes from the variations of the MCGPR system in the shallow slice

This work proposes a two-step workflow for harmonizing the outputs of an 8-channel, 450 MHz Raptor pushcart MCGPR system, addressing variations caused by individual antenna differences and amplitude imbalances between swathes. For Stage 1 corrections, the MCGPR

system was oriented toward open air and operated in time mode to capture reflections from the air interface. This dataset was used to compute initial corrections, specifically focusing on amplitude normalization and time-zero alignment. These corrections are essential to standardize the signal baseline across all channels before further processing. Stage 2 corrections involved harmonizing the spectral and amplitude properties of signals across different channels. To evaluate the performance of these corrections, two controlled experiments were designed and executed. Figure 6-16 illustrates the Raptor45 system's antenna arrays, including transmitters, receivers, and their channel layout. The variations are categorized into antenna-specific differences, such as intrinsic discrepancies between antennas, and array-related variations stemming from relative positions in the antenna array (e.g., differences in cable lengths and surroundings). To harmonize the outputs, two controlled experiments were conducted: an air reflection experiment, where the MCGPR system was positioned upward to measure free-space reflections, and a metal plate experiment, where the system was positioned downward toward a metal plate to simulate full reflections. Each experiment was performed to characterize amplitude ranges for each antenna and time-zero shifts (Δt_c) resulting from both antenna-specific and array-related variations.

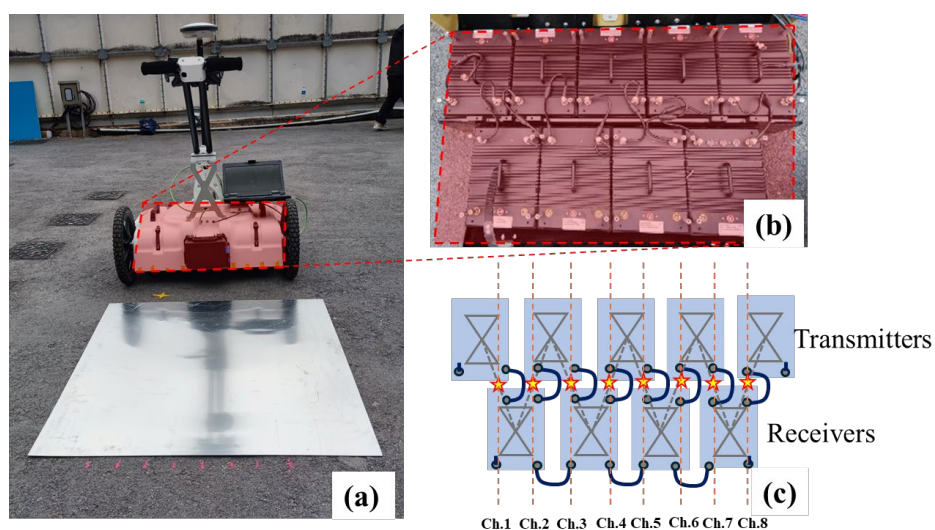


Figure 6-16 Photos of the raptor MCGPR system and the antenna arrays (a) and (b); (c) The illustration of the transmitters, receivers, and channel location of the raptor MCGPR system

The harmonization workflow, illustrated in Figure 6-17, comprises two stages:

- (1). Stage 1: Initial correction using calculated corrections from the controlled experiments and
- (2). Stage 2: Site-specific signals correction using GPR pairs.

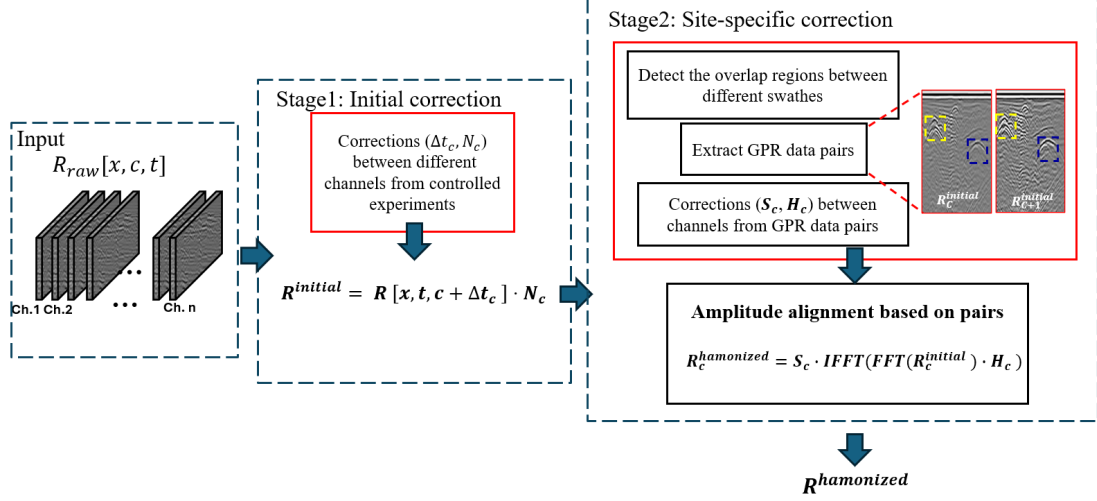


Figure 6-17 The whole workflow of the signal harmonization method

6.2.1.1 Stage 1: Initial Correction

Initial correction addresses time-zero shifts and inter-channel amplitude scale mismatches between swaths. It scales signals by aligning channel reflections using corrections obtained from controlled experiments. Given the received signals from MCGPR that need to be corrected as set R_{raw} . Stage1, the initial correction was represented as:

$$R^{initial}[x, c, t] = R_{raw}[x, c, t + \Delta t_c] \cdot N_c, \quad [6 - 12]$$

where x is the trace index in the scan direction, c is the channel number, t is the reflection time. And Δt_c is the time-zero corrections, N_c is the amplitude scale factor of each channel, represented as:

$$\Delta t_c = t_{base} - t_c, \quad [6 - 13]$$

$$N_c = \frac{A_{base}}{A_c}, \quad [6 - 14]$$

where t_c and t_{base} is the two-way travel time of the direct wave of channel c and the base

channel, respectively. Δt_c is related to the hardware and hardware installation. The variations are derived from thermal drift, electronic instability, cable length differences, and the installed positions in the antenna array. A_{base} is the positive peak of the direct wave's amplitude in the selected base channel, A_c is the positive peak of the direct wave's amplitude in channel c .

6.2.1.2 Stage 2: Site-specific Correction

Besides, Stage 2 aims to correct site-specific variations of the MCGPR survey. In this stage, the site-specific correction was represented as:

$$R_c^{hamonized} = S_c \cdot IFFT(FFT(R_c^{initial}) \cdot H_c). \quad [6 - 15]$$

In Stage 2, the overlapping regions of two different swathes were extracted. To keep consistent, most MCGPR surveys remain 10-30 percent overlap with the adjacent swathes. In these overlapping regions, GPR data pairs, which means GPR data collected by different antennas at the same places, were extracted. Given GPR data pairs $(P_{base}^{initial}, P_c^{initial})$, based on these data pairs, scale factor S_c and frequency filter H_c were represented as:

$$S_c = \frac{A_{P_{base}^{initial}}}{A_{P_c^{initial}}}, \quad [6 - 16]$$

$$H_c = \frac{FFT(P_{base}^{initial})}{FFT(P_c^{initial})}. \quad [6 - 17]$$

To quantify the similarity between two signals, wavelet similarity was used for accuracy assessment, represented as:

$$Wavelet\ similarity(S_i, S_j) = \frac{1}{K} \sum_{k=0}^K [1 - \frac{\|w_i^k - w_j^k\|_2}{\|w_i^k\|_2 + \|w_j^k\|_2}], \quad [6 - 18]$$

where w_i^k denotes the wavelet coefficients at level k for the signal S_i .

6.2.2 Experiments and Results

Several experiments were conducted in a controlled environment to facilitate signal

harmonization for MCGPR systems. These experiments were designed to address the corrections required for amplitude and time-zero alignment (Stage 1) and spectral and amplitude harmonization across multiple channels (Stage 2).

6.2.2.1 Experiment 1: Material-based Reflections in Time Mode

In the first experiment, the MCGPR system was positioned over various materials—brick, concrete, and asphalt—to capture reflections in time mode, illustrated in Figure 6-18 (b). The setup aimed to simulate typical subsurface conditions with varying electromagnetic properties. As shown in Figure 6-18 (a) and (b), the MCGPR antenna array (Ch.1 to Ch.8) was initially aligned at positions labelled a_1 to a_8 . After collecting data from this configuration, the antenna array was shifted laterally by one channel offset distance to the right, and an additional swath of data was collected. This process resulted in eight swathes of data, with varying numbers of GPR data pairs obtained from the positions a_1 to a_8 . The purpose of this experiment was to assess how well the Stage 2 corrections harmonized the signals across channels when the radar was exposed to different materials. The diverse materials allowed for a robust evaluation of the corrections under varying reflection conditions.

6.2.2.2 Experiment 2: Aligned Reflections in Distance Mode

The second experiment followed a similar setup to the first but used the distance mode of the MCGPR system instead of the time mode. In this experiment, the MCGPR collected data pairs aligned along survey lines, spanning positions from a to b , as shown in Figure 6-18 (c). Unlike the time mode, the distance mode reflects the most common operating scenario in real-world MCGPR surveys, where data is collected continuously along a survey path. The GPR data pairs acquired in distance mode were used to evaluate the effectiveness of the Stage 2 corrections in a more practical context, simulating typical survey conditions. By comparing the results of the two experiments, the robustness and applicability of the Stage 2 corrections across

varying acquisition modes and materials were thoroughly assessed.

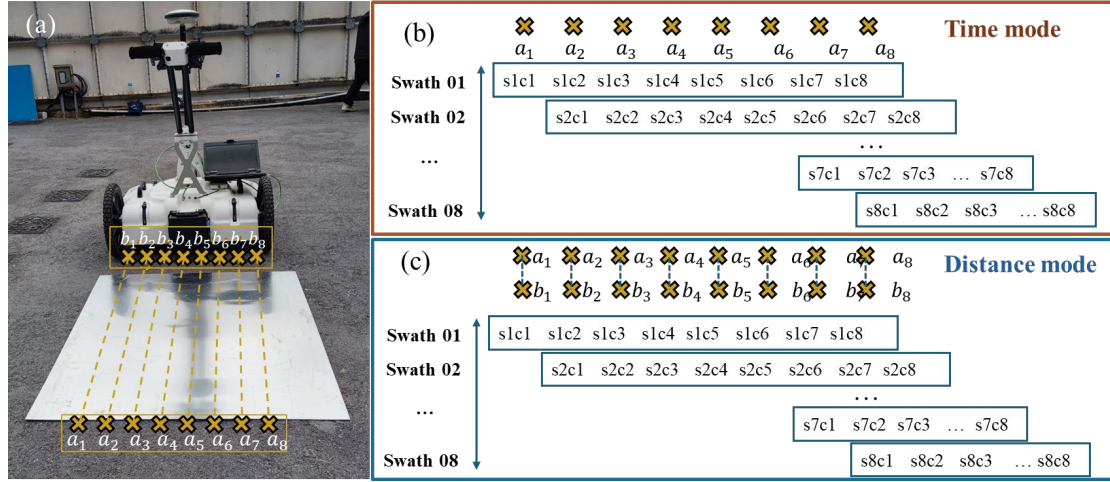


Figure 6-18 Photos of the experiments (a) Illustration of the time mode; (b) and distance mode; (c) data collection

Figure 6-19 illustrates the variations in antenna performance across all channels by plotting the average values of the positive and negative peaks, along with the two-way travel time at time-zero. The positive peak amplitudes (red bars) vary noticeably across channels, ranging from 18,464.7 (Ch.2) to 20,417.6 (Ch.3), with Ch.3 exhibiting the highest amplitude. Similarly, the negative peak amplitudes (blue bars) show significant variation, ranging from -13,909.7 (Ch.3) to -17,598.0 (Ch.8). Notably, the outermost channels (Ch.1 and Ch.8) display larger deviations in negative peak amplitudes compared to the inner channels, indicating greater signal inconsistencies at the edges of the antenna array.

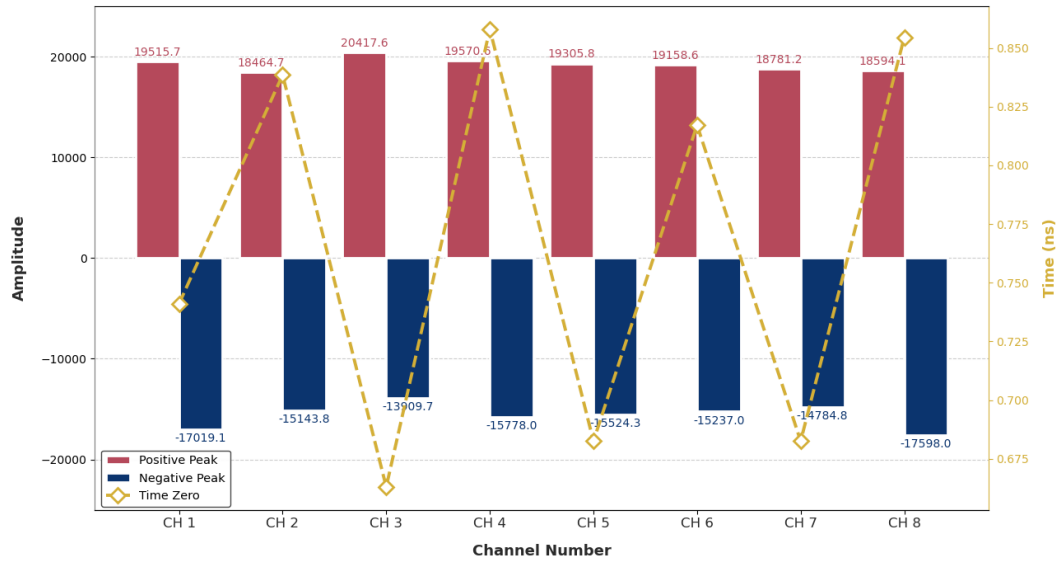
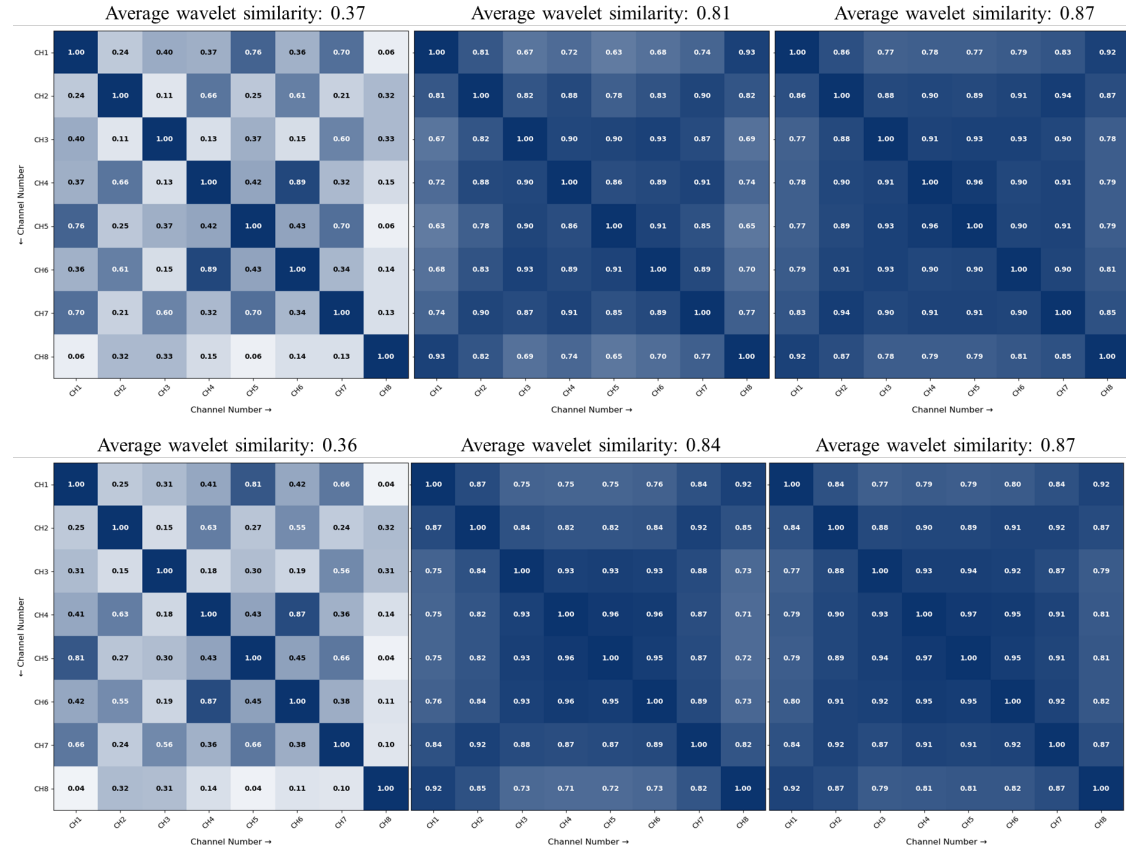


Figure 6-19 Peak to peaks amplitude and time-zero variations across channels

The two-way travel time at time-zero (yellow diamonds) shows a periodic fluctuation across the channels. This periodicity is likely due to the installation configuration, in which two adjacent antennas are housed together in a single box. The relative gap introduced by this arrangement may cause variations in the two-way travel time, resulting in the observed periodic fluctuations. Such inconsistencies in the time-zero calibration across channels can affect the alignment of signals during data processing. Additionally, gain curves applied in GPR data processing—which are typically time-dependent functions—can amplify these variations in time-zero and amplitude. As a result, strip-like noise artefacts may appear along the trajectories, as shown in Figure 6-15. The observed variations in both amplitude and time-zero underscore the necessity of harmonization corrections to achieve consistent signal quality and alignment across the entire antenna array.

Figure 6-20 illustrates the wavelet similarity matrices across channels for GPR data collected on three different materials (brick, concrete, and asphalt) and at three correction stages (raw, after Stage 1 correction, and after Stage 2 correction) using time mode to obtain GPR data at the same static place. The results reveal that the initial correction stage leads to a substantial improvement in wavelet similarity across channels, increasing from an average of 0.36 in the

raw data to 0.82 after the first correction. This significant enhancement demonstrates the effectiveness of the initial correction, which relies on adjustments derived from the air reflections of the MCGPR system. Beyond the corrections based on air reflections, the application of frequency filter corrections and scale factors calculated from cross-channel GPR data pairs (time mode) further improves the consistency of the signal outputs. These additional refinements, applied during the second correction stage, result in a further increase in wavelet similarity, from 0.82 to 0.85. The observed improvements across both correction stages highlight the robust performance of the two-stage correction process in enhancing signal uniformity and alignment across channels for various materials.



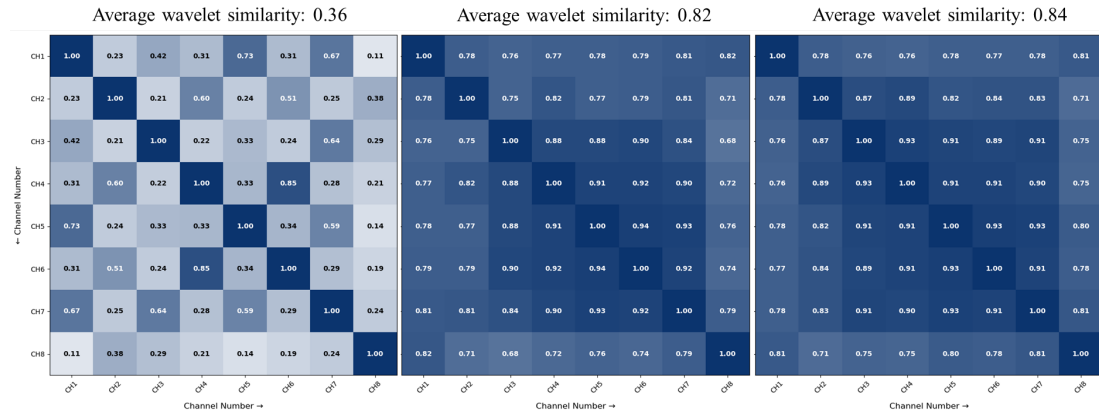


Figure 6-20 Wavelet similarity matrices across channels for GPR data, with different rows representing materials (brick, concrete, and asphalt) and different columns representing correction stages (raw, after Stage 1 correction, and after two-stage correction)

Besides the static GPR data acquisition shown in Figure 6-18 (a), the GPR data pairs obtained by distance mode (seeing Figure 6-18 (b) and (c)) were used for corrections in Stage 2. Distance mode is the more common mode when doing GPR data acquisition. Figure 6-21 presents the wavelet similarity across three different materials—brick, concrete, and asphalt—for eight channels (Ch.1 to Ch.8). Each bar represents the mean wavelet similarity of eight survey swathes for each material and channel. The data is categorized into three stages: raw data (grey shades), Stage 1 corrections (blue, orange, and green), and Stage 2 corrections (lighter shades of blue, orange, and green). Error bars indicate the variability of wavelet similarity across the survey swathes, while dashed horizontal lines mark the highest (0.90) and lowest (0.79) similarity values observed across all channels and materials. The raw data showcases relatively low similarity across channels for all materials, which ranges from 0.18 to 0.44. especially the mean similarities of Ch.1 to other channels are the lowest regardless of materials. These low similarity values show the inconsistency of the raw outputs. After Stage 1 corrections, all materials exhibit a significant improvement in similarity across channels. Stage 2 corrections, which incorporate GPR data pairs obtained using the distance mode, further refine the similarity values. This stage is particularly effective for brick and concrete, bringing their similarity values closer to the highest similarity threshold, while asphalt shows more

modest improvements in some channels. Additionally, the plot highlights the variation in improvements across different channels. Channels Ch.1, Ch.2, and Ch.3 generally show higher similarity improvements compared to channels Ch.7 and Ch.8, which exhibit greater variability and smaller gains. This indicates that the correction methods' effectiveness varies depending on the antenna's quality, likely due to differences in signal behaviour across the channel array.

In summary, the figure demonstrates the overall effectiveness of the correction stages in improving wavelet similarity across all materials, with Stage 2 corrections (distance mode) providing the most substantial enhancements. However, the variation in improvements across channels suggests the need for channel-specific considerations when applying these corrections.

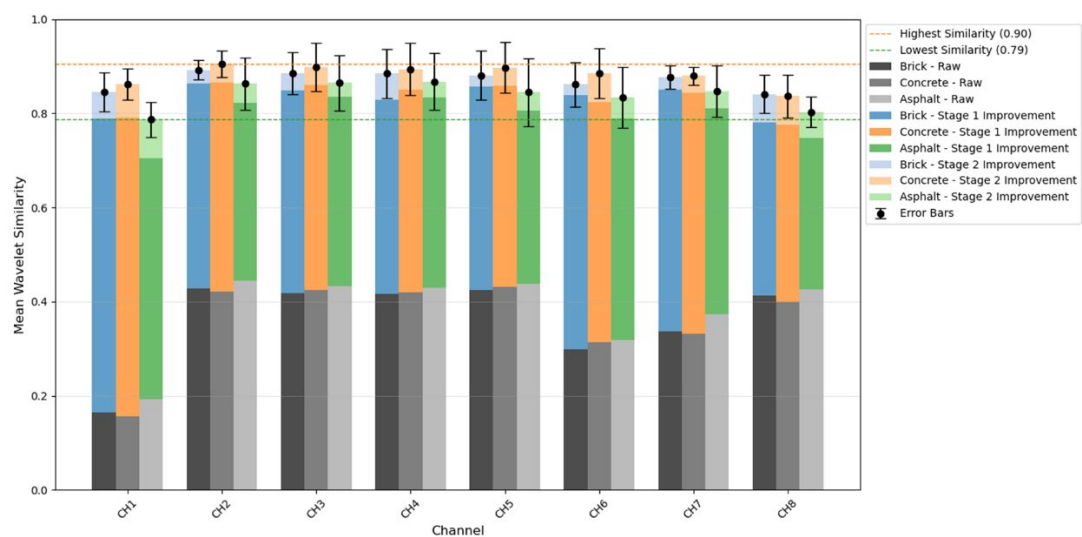


Figure 6-21 The signal harmonization performance of different stages in different pavement materials (brick, concrete, and asphalt)

To assess the effectiveness of the proposed harmonization workflow, Figure 6-22 shows the amplitude slices generated by MCGPR data with and without the proposed signal harmonization workflow.

The experimental results strongly validate the proposed workflow. Observations of time-zero indices (Figure 6-17) underscore the need to address both position-based and antenna-specific variations. Controlled experiment-derived corrections effectively normalized

the initial time-zero shifts, ensuring consistent timing across all channels. Additionally, the amplitude alignment method substantially reduced amplitude imbalances between swathes, as evidenced by the diminished swath boundaries and uniform reflectivity in the harmonized slices. These findings highlight the importance of comprehensive post-processing to enhance the interpretability of MCGPR data and pave the way for improved horizontal slice quality in practical applications.

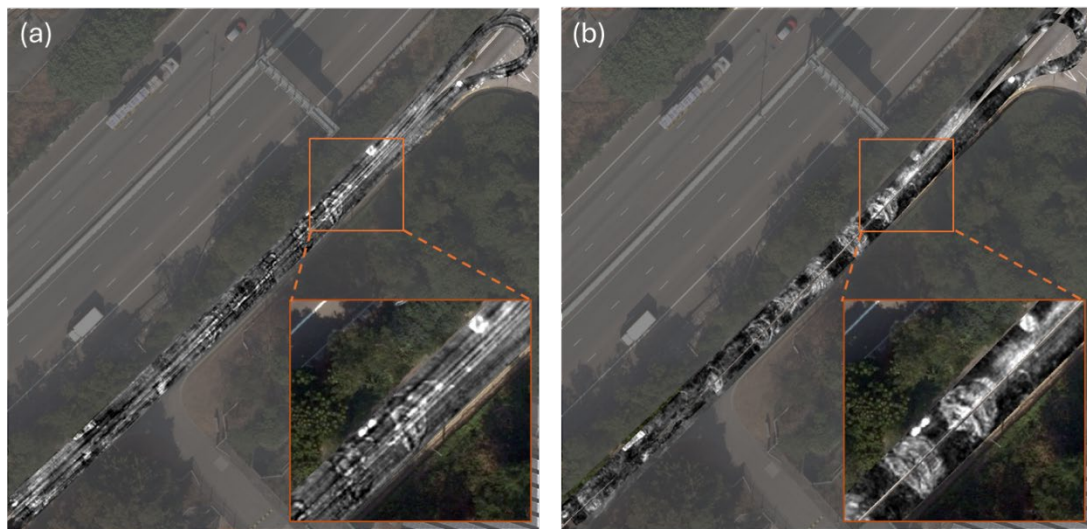


Figure 6-22 Comparison of amplitude slices generated (a) with and (b) without the proposed harmonization workflow

Figure 6-22 compares the amplitude slices generated from MCGPR data without the proposed signal harmonization workflow (a) and with the harmonization applied (b), demonstrating significant improvements in data quality. The harmonized data in (b) exhibits smoother transitions across swathes, with reduced visual discontinuities and artefacts compared to (a), indicating that the workflow effectively mitigates inconsistencies caused by variations in channel responses and environmental noise. Additionally, the zoomed-in inset in (b) reveals sharper and more defined underground features, whereas (a) shows blurred and fragmented details, highlighting the harmonization workflow's capacity to preserve and enhance fine-grained spatial details. The unharmonized data in (a) also displays irregular amplitude variations, obscuring the continuity of subsurface features, an issue significantly resolved in

(b), where the amplitude slices maintain greater uniformity, allowing for better interpretation of the GPR data. Overall, these improvements suggest that the proposed harmonization workflow enhances the consistency, clarity, and reliability of MCGPR amplitude slices, making them more suitable for accurate subsurface mapping in complex environments.

6.3 Chapter Summary and Contribution

This chapter introduced two critical advancements in time-series MCGPR data processing, addressing challenges essential for large-scale subsurface imaging from two perspectives.

Firstly, a novel positioning enhancement method was developed to migrate positioning errors (above-ground). This method improves trajectory accuracy in MCGPR surveys, particularly in urban environments where GNSS signals are prone to obstructions and multi-path errors. The method integrates advanced techniques such as HOG and 2D FFT for underground feature extraction, paired similarity-based alignment, and global optimization using dynamic programming. This multi-step process leverages MCGPR data as an external input to refine positioning accuracy, ensuring compatibility with diverse positioning systems and achieving robust trajectory alignment.

Secondly, a two-stage harmonization workflow was proposed to address signal variations caused by antenna-specific differences and array-related factors (underground). The first step involves corrections derived from controlled experiments to normalize time-zero indices and amplitude ranges, ensuring consistency across channels. The second step dynamically aligns amplitudes across swathes to eliminate striping artefacts in processed slices. This workflow was validated through experimental results, demonstrating its capability to produce uniform, high-quality subsurface images free from striping noise.

Together, these contributions establish a framework for improving the accuracy and

reliability of time-series MCGPR data. The proposed methods address fundamental challenges in positioning and signal harmonization, paving the way for more robust subsurface imaging. These advancements hold significant potential for real-world applications, particularly in urban geophysical surveys, and provide a strong methodological foundation for future research in large-scale MCGPR analysis.

Chapter 7. A Forensic Investigation in a Real-world Leakage Case

In this chapter, a real-world water leakage case was thoroughly investigated through the complete workflow illustrated in Figure 7-1. While MCGPR is a powerful tool for rapid and extensive screening, it cannot independently resolve leakage detection due to its inability to provide direct confirmation of leaks. Excavation, being costly, disruptive, and often restricted without compelling evidence, necessitates a cross-verification process before any ground is opened. The workflow begins with water flow monitoring, which identifies DMAs (District Metering Areas) with abnormal water losses, narrowing the search area to regions of interest. Within these flagged areas, MCGPR is employed to conduct fast and efficient road surveys, mapping clusters of potential changes indicative of leaks. The primary strength of MCGPR lies in its broad coverage and high efficiency, significantly reducing manpower requirements. However, MCGPR outputs only provide “suspected regions” rather than confirmed leak locations, necessitating further verification. To pinpoint exact leak locations, acoustic methods are employed. These methods excel at detecting the sound or vibration generated by leaks, offering high accuracy and reliability (Muggleton et al., 2011b). Tools such as ground microphones and noise loggers are capable of directly confirming leaks, even in complex environments. However, acoustic methods also have limitations. They are often labor-intensive and require human inspectors to manually operate equipment, such as listening sticks or attachable acoustic sensors (Duan et al., 2015; Gao et al., 2005). This dependency makes them time-consuming and costly for large-scale surveys (Yu et al., 2021). The complementary nature of MCGPR and acoustic methods is what makes the workflow effective. While MCGPR provides rapid, large-scale screening to identify suspected regions with minimal manpower, acoustic tools are essential for cross-verification and precise localization. This complementary

approach ensures that excavation is only authorized when both MCGPR and acoustic evidence align, minimizing unnecessary disruptions and costs while maintaining high accuracy. After excavation and rehabilitation, a post-repair MCGPR resurvey is performed to confirm that the anomaly has been resolved. This step provides a quick and non-intrusive check to ensure the effectiveness of the repair. As this thesis focuses on the second step of the workflow—proactive leakage detection using the MCGPR system—the following content will provide a detailed discussion of this process.

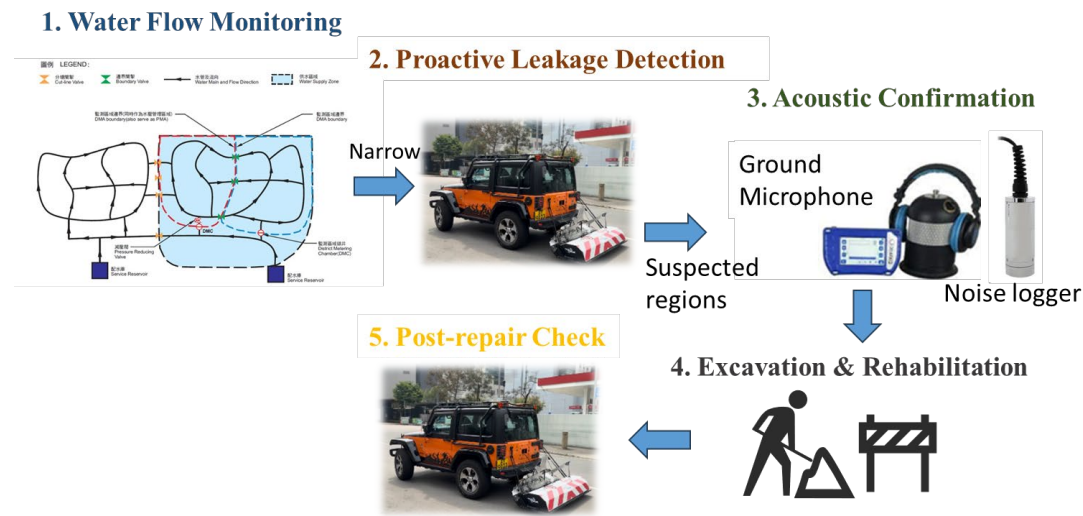


Figure 7-1 Workflow of water leakage detection

In proactive leakage detection, the methods proposed in the previous chapters were fully utilized to address this case in a structured four-step process. First, the positioning enhancement and signal harmonization methods were applied during the post-processing of the MCGPR data to improve its quality and reliability. Second, temporal AS and IFS volumes were generated based on the methodologies outlined in Chapter 4. Third, the change detection techniques introduced in Chapters 4 and Chapter 5 were employed to identify the leakage in the temporal measurements. Finally, the chapter concludes with a discussion on the application of temporal MCGPR measurement features in real-world cases and highlights the associated challenges.

7.1 The Site and MCGPR Temporal Measurement's Introduction

As illustrated in Figure 7-2, the global survey region spans a length of 5.5 km. After a preliminary analysis of MCGPR data, combined with water pressure anomaly records provided by our collaborators (WSD), a 1.6 km-long section (highlighted in green) was identified as the preliminary suspected region. Subsequently, three MCGPR surveys were conducted using the Raptor45 system between November 2023 and June 2024, as shown in the timeline at the bottom of Figure 7-2.

Through further analysis of the temporal MCGPR data and water pressure records during these surveys, a narrowed region of interest measuring 250 meters in length was selected for detailed investigation. This section, indicated by the red rectangle in Figure 7-2, provided critical data for analyzing subsurface conditions leading up to a major pipe burst event on November 3, 2024. These data enabled a comprehensive evaluation of the proposed change detection methods, focusing on temporal measurements.



Figure 7-2 Illustration of the research regions and the temporal MCGPR surveys have been conducted

The surveys were conducted using the Raptor45 system, an 18-channel MCGPR system with a channel offset of 8.4 cm, providing a total survey width of 1.4 m perpendicular to the heading direction. The trace increment was set to 4 cm in the heading direction, ensuring high spatial resolution for detecting and mapping subsurface features. Over the course of the three surveys, a total of 4.8 km of MCGPR data were collected per survey, resulting in a cumulative data volume of 86.4 km across all channels. This substantial dataset required extensive processing to detect and analyze subsurface anomalies effectively.

The underground utility records for the surveyed region were provided by the Underground Utilities Information System (UUIS), developed by the Lands Department of Hong Kong. These records indicated the presence of two critical underground utilities within

the ROI. The first was a 400 mm steel gas pipe, located at a depth of approximately 1.1 meters. This pipe is marked in green on the utility map in Figure 7-3 (a) and is further confirmed by the amplitude slice at the same depth in Figure 7-3 (b). The corresponding radargram, highlighted by the yellow rectangle in Figure 7-3 (d), shows a prominent horizontal layer with high reflectivity at a depth of approximately 20–30 ns, consistent with the gas pipe's location. The second utility was a 1200 mm concrete water main, buried at depths ranging from 1.6 meters to 1.9 meters. The presence of the water main is represented in the amplitude slice at 1.9 meters depth in Figure 7-3 (c). The radargram highlighted within the orange rectangle in Figure 7-3 (e) reveals a distinct reflective layer at a depth of approximately 30–40 ns, which aligns with the recorded location of the water main.

The consistency between the utility records and the detected subsurface features observed in the amplitude slices and radargrams highlights the reliability of the Raptor45 system for underground utility mapping. These pre-burst surveys, conducted before the pipe failure, provided critical baseline data on subsurface conditions under normal circumstances, forming the foundation for detecting early signs of pipe leakage.

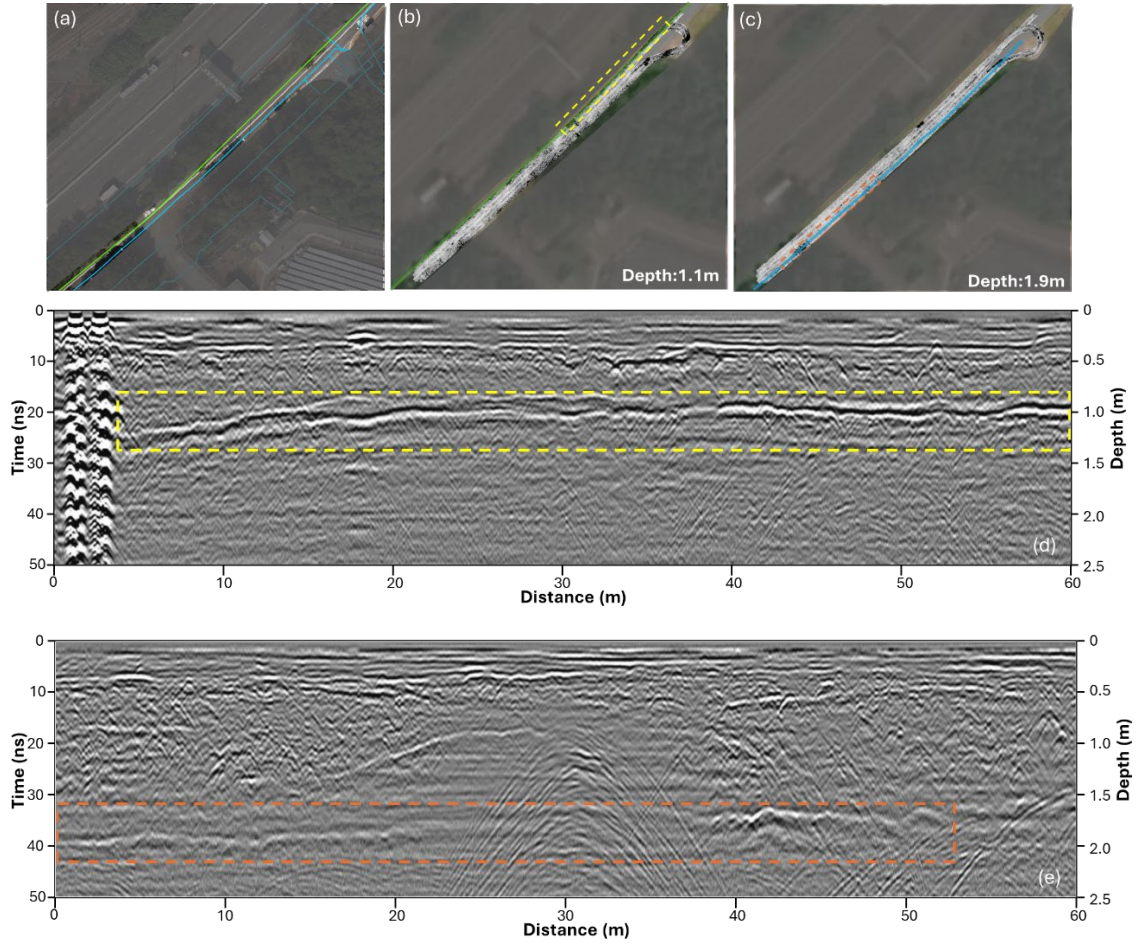


Figure 7-3 (a) Utility record plant overlaid on a satellite image; (b) Amplitude slice at 1.1-meter depth showing the gas pipe; (c) Amplitude slice at 1.9-meter depth showing the water main; (d) Radargram corresponding to the yellow rectangle in (b); (e) Radargram corresponding to the orange rectangle in (c)

The pre-burst surveys were particularly valuable in identifying gradual changes in subsurface conditions, such as early leakage-related anomalies, which may occur before the actual pipe burst. This data enabled a detailed analysis of the progression of subsurface failure, supporting the evaluation of temporal change detection methods and providing insights into the stages of pipe leakage leading up to the burst event.

On November 3, 2024, a significant pipe burst occurred in the target region (see Figure 7-4). The burst caused extensive damage, disrupting water supply to approximately 120,000 residents for 12 hours. The incident created a unique opportunity to analyze the temporal

MCGPR data collected over time. By comparing multiple measurements before the burst, the study aimed to evaluate the effectiveness of the proposed methods for detecting and characterizing invisible hazards in subsurface conditions caused by the water main failure.



Figure 7-4 News and on-site photo of the pipe bursts (Emily Hung et al., 2024)

7.2 Position Enhancement and MCGPR Signal Harmonization

Accurate positioning and signal harmonization are critical for aligning temporal measurements, ensuring consistent spatial references, and improving data quality for reliable change detection. Real-world MCGPR surveys often face challenges in obtaining a continuous trajectory with high positioning accuracy due to GNSS signal degradation caused by environmental factors. At the survey site, tree canopies and tall buildings obstruct GNSS signals, leading to positioning inaccuracies. Figure 7-5 provides an on-site photo of the target region,

where tree canopies on both sides of the road create signal blockages. GNSS signal quality frequently drops from RTK-Fixed to differentially corrected levels near the tree canopy, resulting in GNSS drift, marked by red lines in Figure 7-6 (a). This GNSS drift represents a common issue in road-based MCGPR surveys, causing misalignments in data and making subsurface feature detection impossible because of a lack of image integrity.



Figure 7-5 *On-site photo of the target region*

Unlike jumping GNSS noise, which is easy to identify and correct, GNSS drift introduces systematic offsets that occur alongside the true trajectory. This drift manifests in the MCGPR amplitude slices as misalignments of subsurface features between survey lines. For instance, Figure 7-5 (b) highlights these misalignments near the tree canopy, where subsurface objects such as manhole reflections are offset across different survey swaths.

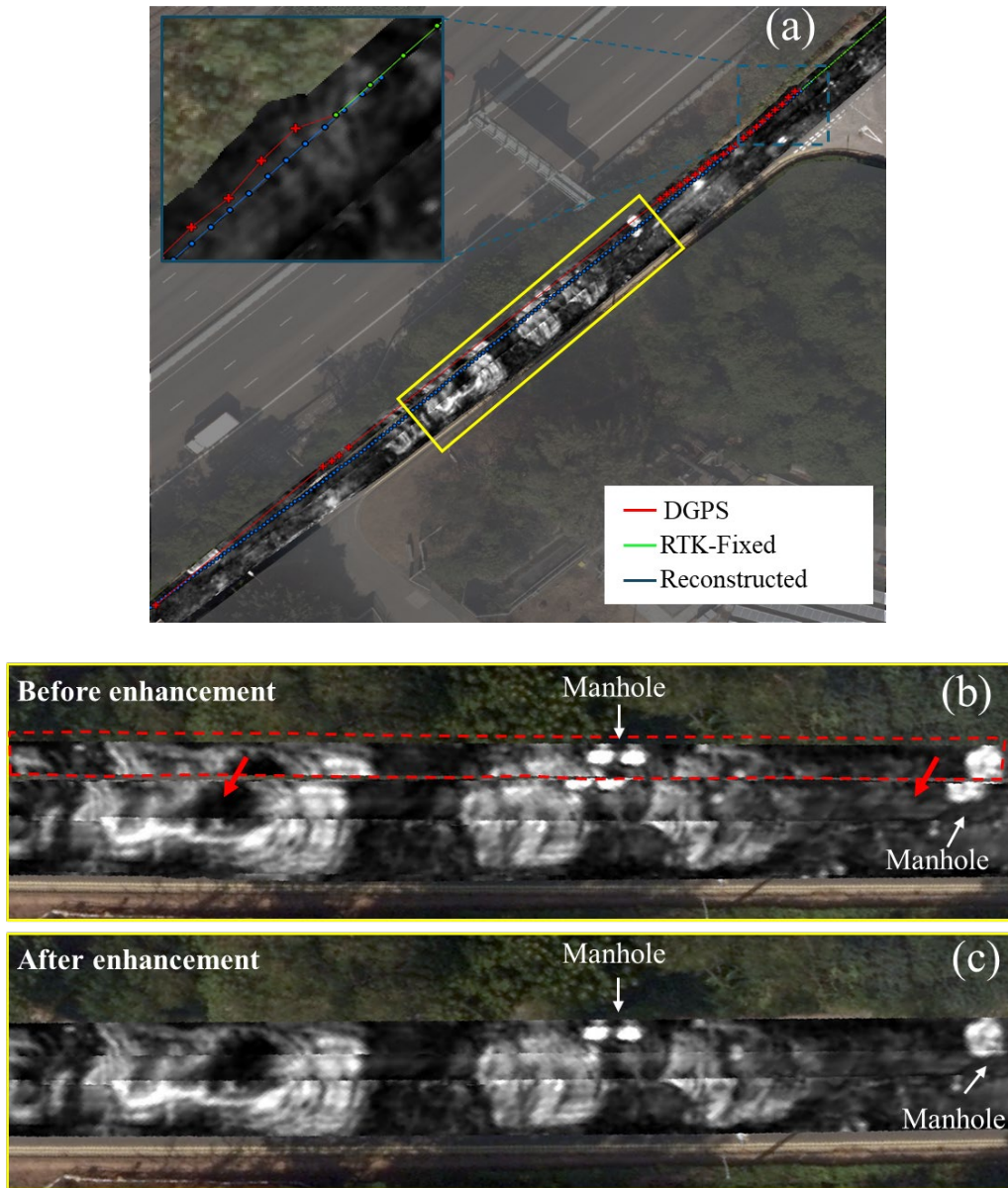


Figure 7-6 (a) GNSS trajectories of the target region; Amplitude slices (b) before, and (c) after positioning enhancement

To overcome this challenge, the positioning enhancement method proposed in Chapter 6 was applied to correct GNSS drift. After processing, the corrected amplitude slices, shown in Figures 7-6 (c), demonstrate that features such as manhole reflections are now properly aligned across survey swaths. This correction highlights the effectiveness of the proposed method in mitigating GNSS drift, ensuring consistent spatial references, and significantly enhancing the reliability of temporal MCGPR data.

Another major challenge in real-world MCGPR surveys is the inconsistency in signal amplitudes caused by variations in antenna responses and environmental noise. Before harmonization, the amplitude slices revealed striping artefacts due to unnormalized antenna responses, as shown in Figure 7-7 (a). (This figure is identical to Figure 6-22 and is reproduced here for convenience.) These artefacts obscure the true strength of subsurface reflections, complicating data interpretation and reducing the reliability of temporal analysis.

To address this, the signal harmonization method proposed in Chapter 6 was applied. This method normalizes the amplitude levels across all antennas, eliminating inconsistencies and enhancing data quality. The harmonized amplitude slice, shown in Figure 7-7 (b), reflects subsurface reflections accurately, free from striping artefacts.

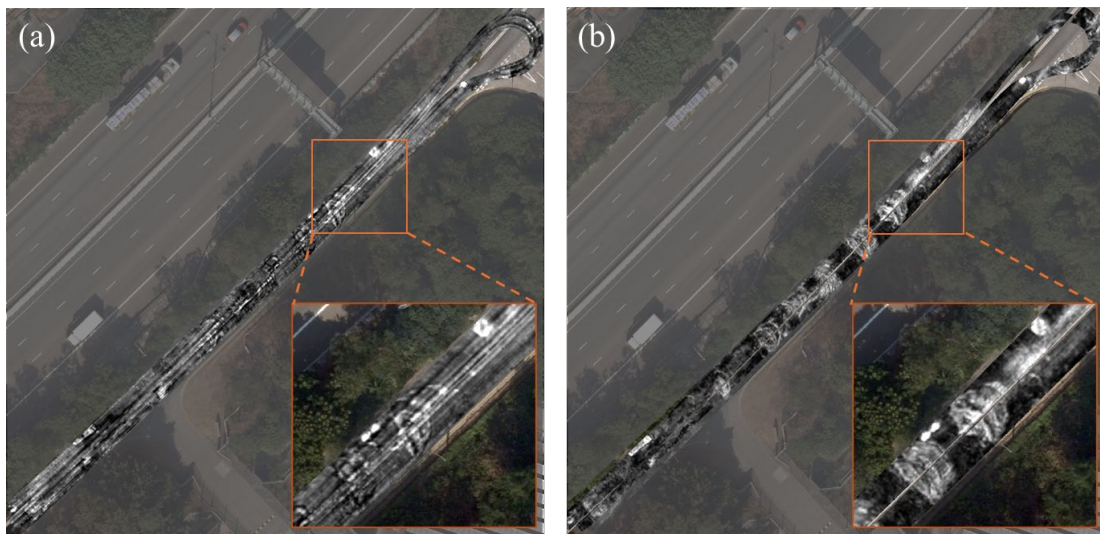


Figure 7-7 Amplitude slice before and after signal harmonization (same as **Figure 6-22**)

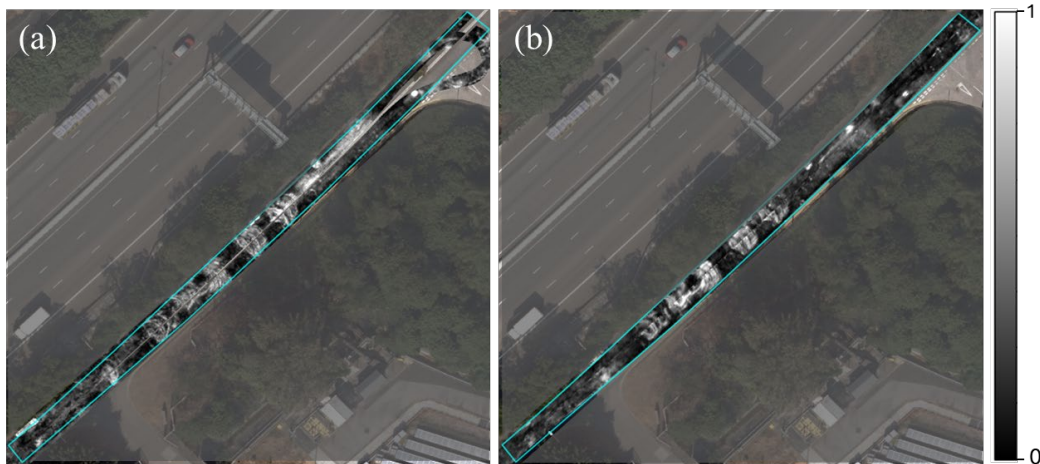
By removing noise and standardizing the amplitude levels, signal harmonization improves the interpretability of the MCGPR data. This step is particularly important for temporal measurements, where consistent signal quality across surveys is essential for accurate change detection.

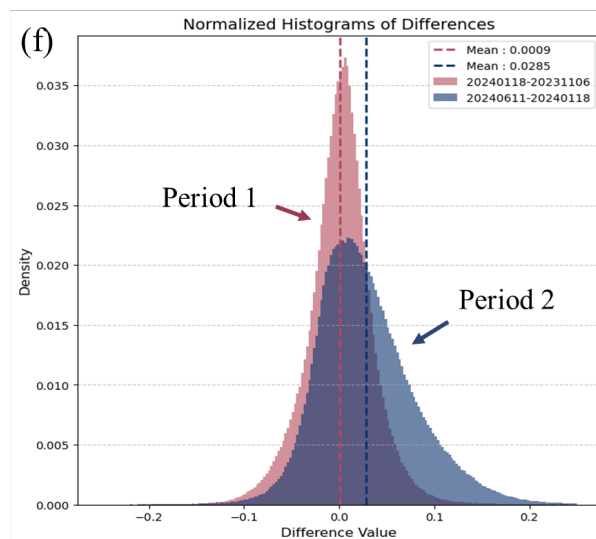
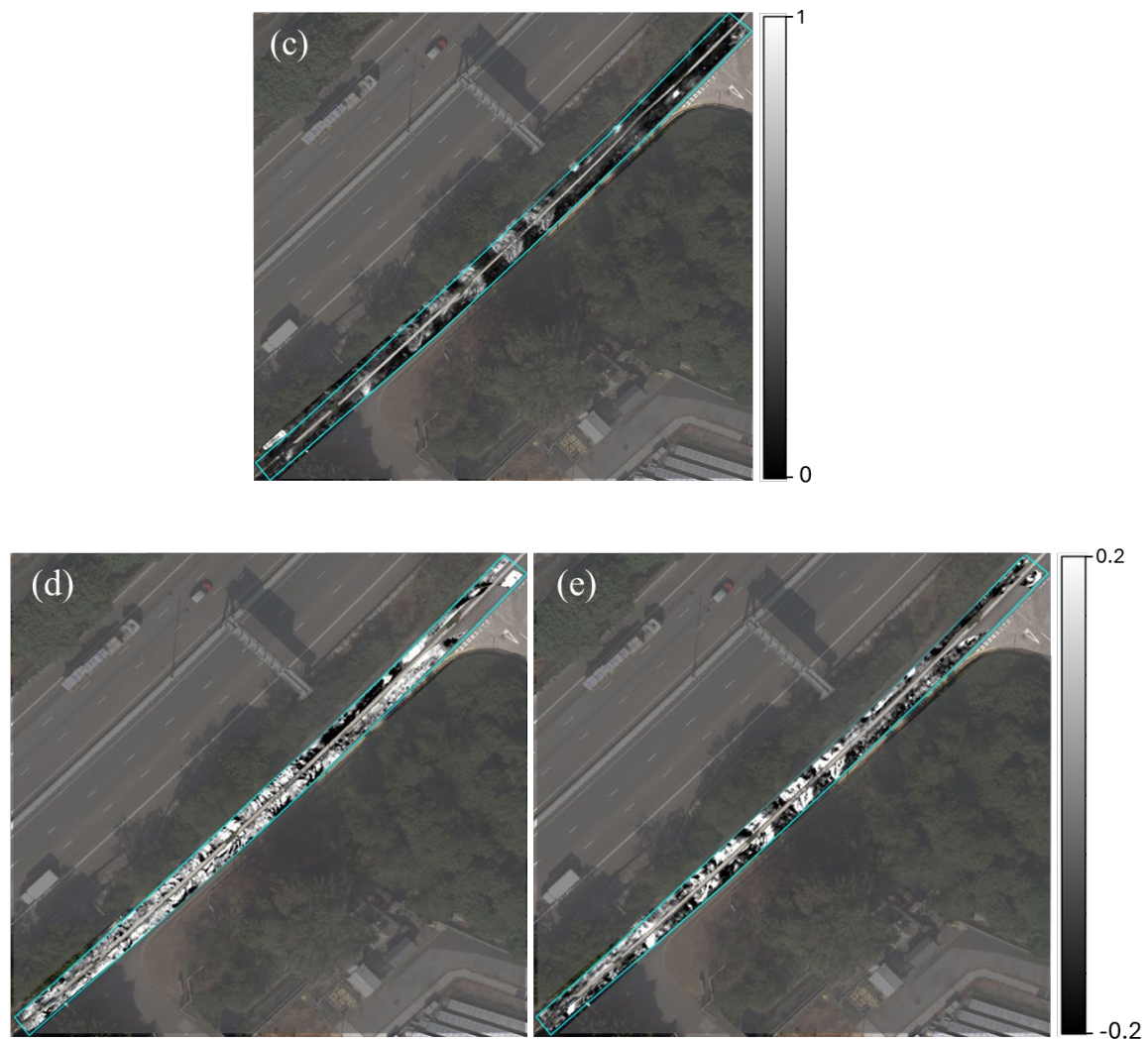
7.3 Change Detection Results

In this section, three MCGPR surveys were divided into two periods utilizing the proposed leakage detection method in Chapter 5 which integrates both amplitude and instantaneous frequency changes to extract the suspected leakage regions. MCGPR measurements were conducted on 2023-11-06, 2024-01-18, and 2024-06-11. For the difference image generation, the first-time interval (2023-11-06 to 2024-01-18) is referred to as Period 1, and the second interval (2024-01-18 to 2024-06-11) is referred to as Period 2. Two leakage maps were generated to highlight regions with high suspicion.

7.3.1 Amplitude Slices

Figure 7-8 illustrates the temporal analysis of normalized amplitude slices and their corresponding amplitude difference images between consecutive surveys, along with histograms of the amplitude differences for two distinct periods. Three normalized amplitude slices, corresponding to surveys conducted on November 6, 2023, January 18, 2024, and June 11, 2024, are presented in Figures 7-8 (a), 7-8 (b), and 7-8 (c), respectively. These slices, extracted at a depth of 20–32.5 cm, highlight high-reflection features such as manholes and clusters located in the middle of the road. These clusters may indicate the separation between pavement and soil layers or reflect underlying subsurface anomalies.





(Figure 7-8 Caption on the next page)

Figure 7-8 Temporal analysis of normalized amplitude slices (AS) and difference images: (a) Normalized amplitude slice (November 6, 2023); (b) Normalized amplitude slice (January 18, 2024); (c) Normalized amplitude slice (June 11, 2024); (d) Amplitude difference image (Period 1); (e) Amplitude difference image (Period 2); (f) Histograms of amplitude differences for Period 1 and Period 2

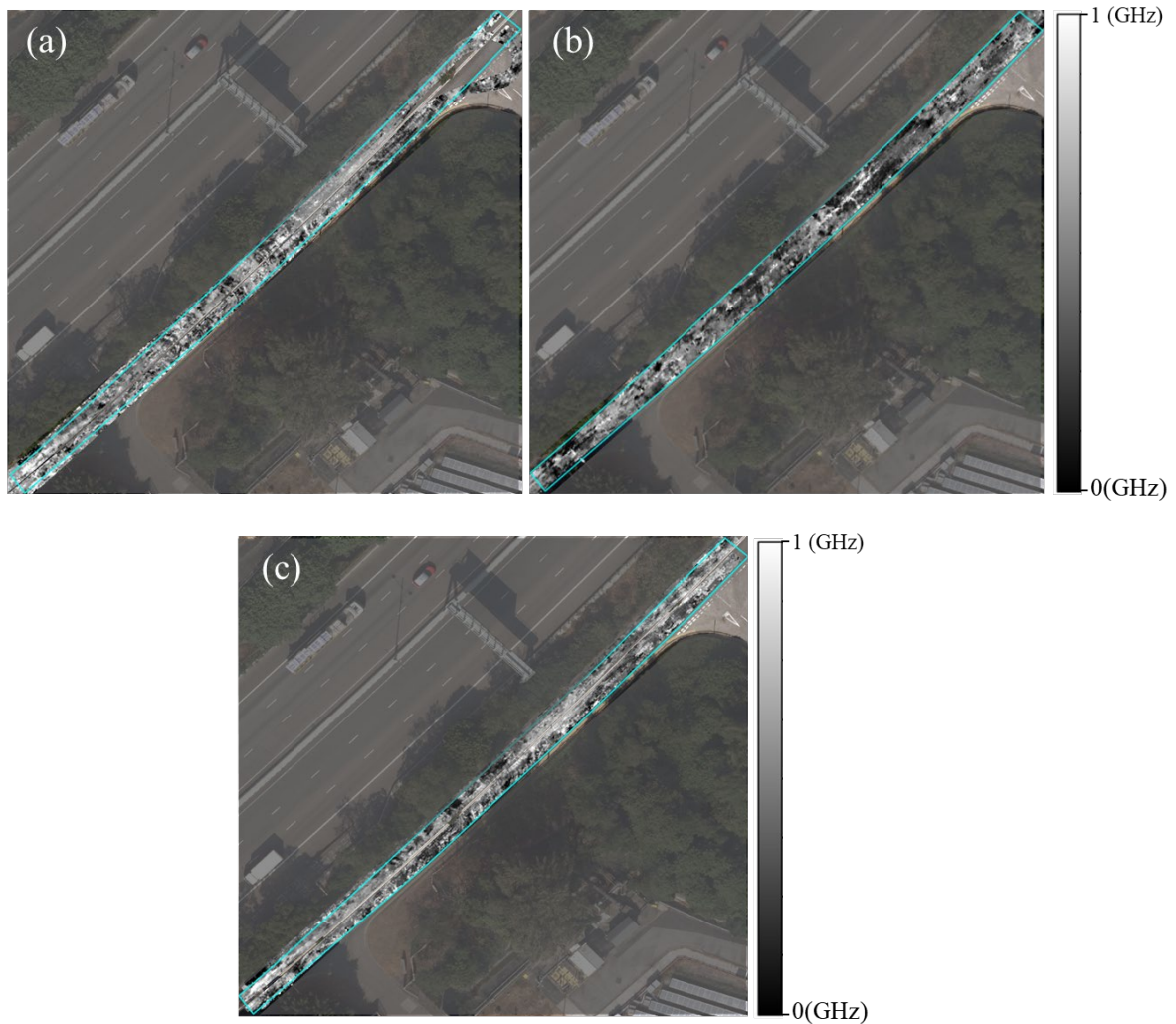
To examine temporal changes, amplitude difference images were calculated between consecutive surveys. Figure 7-8 (d) represents the difference image between the first two surveys (Period 1), while Figure 7-8 (e) shows the difference image between the last two surveys (Period 2). While some patterns in the AS are straightforward to interpret, the difference images reveal spatial clusters of changes that are less intuitive to analyze visually. Compared to the original amplitude slices, the difference images are more challenging to interpret directly due to their abstract representation of temporal changes. To gain a clearer understanding of the overall trends, histograms of the amplitude difference values for the two periods were generated, as shown in Figure 7-8 (f).

During Period 1, the mean amplitude difference value was approximately 0.0009, which is very close to zero. The histogram for this period, represented by the red curve, is narrow and symmetrical, resembling a normal distribution. This indicates that most amplitude differences during Period 1 were small and localized. These minimal changes are likely attributable to minor environmental factors such as soil compaction, seasonal variations, or slight differences in survey conditions. In contrast, Period 2 exhibited a mean amplitude difference value of 0.0285, reflecting more significant changes in subsurface conditions compared to Period 1. The histogram for Period 2, represented by the blue curve, is wider and slightly skewed, suggesting larger and more unevenly distributed amplitude differences. These observations point to more dynamic and non-uniform subsurface changes during this period, which could be associated with water leakage, increased soil saturation, or structural shifts near buried utilities. While AS depicts the horizontal distribution of reflected energy, it primarily describes variations in the strength of reflections. In comparison, IF slices provide an alternative

perspective by describing the frequency characteristics of the GPR signal.

7.3.2 Instantaneous Frequency Slices

Similar to the amplitude slices, IF slices at a depth of 20–35.25 cm were generated for the three MCGPR measurements (Figures 7-9 (a–c)). Unlike amplitude slices, IF slices do not require normalization before generating difference images. The IF difference images for the two periods are shown in Figures 7-9 (d) (Period 1) and 7-9 (e) (Period 2). To provide a global perspective on the temporal variations, the histogram of the IF difference cube for both periods is shown in Figure 7-9 (f).



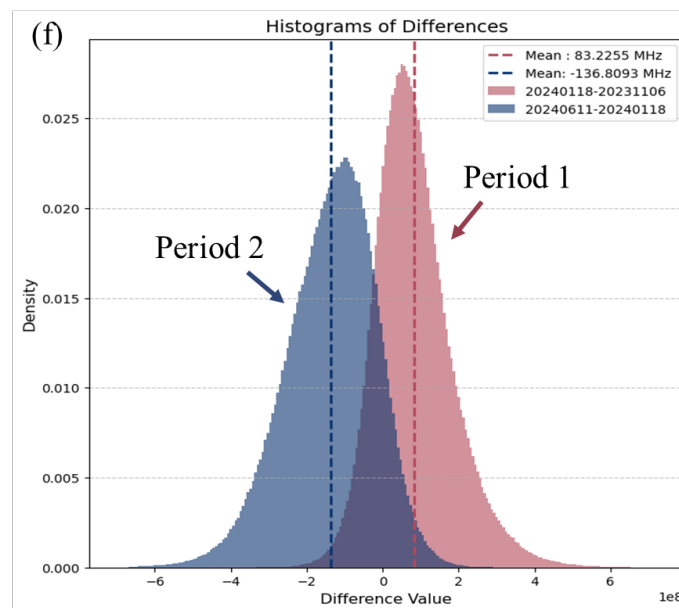
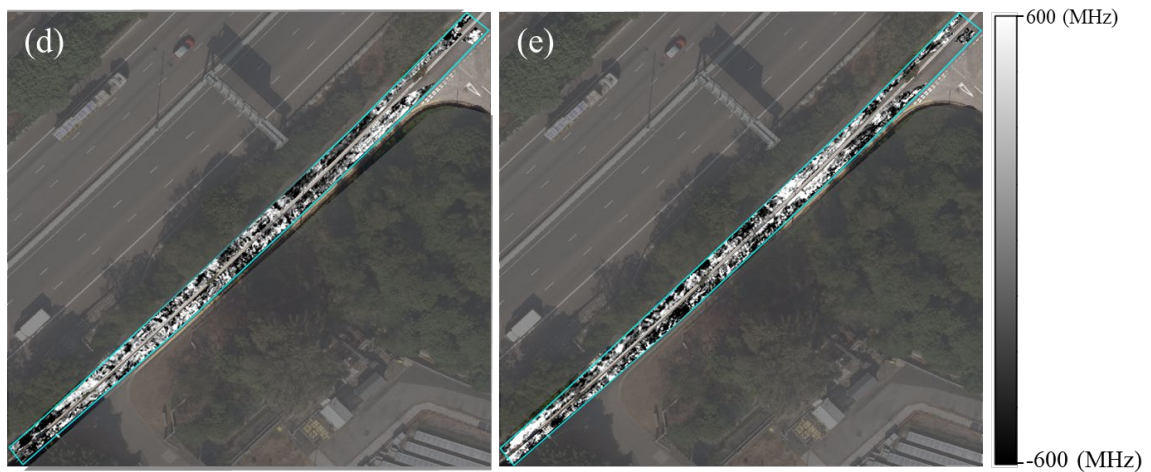


Figure 7-9 Temporal analysis of IFS and difference images: (a) IFS (November 6, 2023); (b) IFS (January 18, 2024); (c) IFS (June 11, 2024); (d) IF difference image (Period 1); (e) IF difference image (Period 2); (f) Histograms of IF differences for Period 1 and Period 2

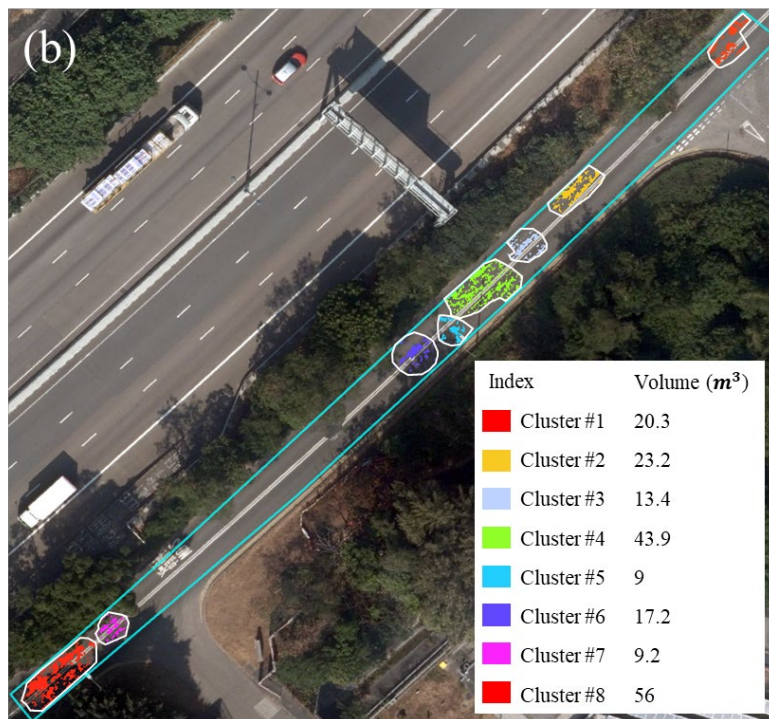
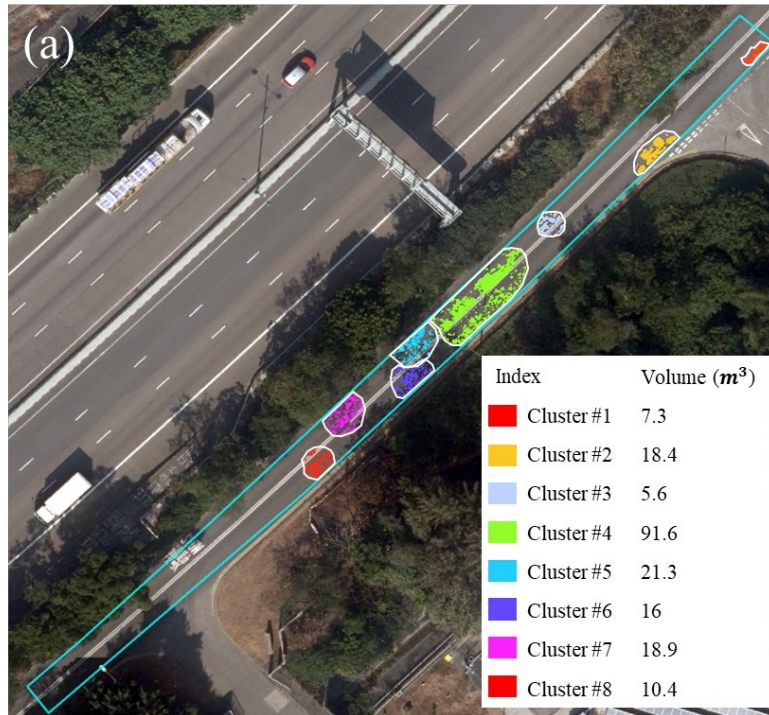
Key observations from the analysis reveal distinct characteristics for the two periods. For Period 1, the mean IF difference value is approximately 83.23 MHz, indicating relatively small changes in the frequency spectrum. The blue histogram curve is narrow and concentrated around the mean, reflecting minimal variations in the subsurface frequency domain. These minor changes are likely caused by environmental factors such as seasonal soil compaction or slight variations in survey conditions. In contrast, Period 2 shows a mean IF difference value of -136.81 MHz, indicating a significant negative shift in the frequency spectrum. The red

histogram curve is broader and skewed toward negative values, highlighting more substantial and unevenly distributed changes in the frequency domain. These changes are likely attributable to water-related dispersion effects, such as attenuation of high-frequency components caused by increased soil moisture content or leakage.

The increasing magnitude of IF differences between Periods 1 and 2 aligns with the amplitude changes observed in the corresponding analyses. This provides additional evidence of evolving subsurface conditions, such as rising water content, soil saturation, or structural disturbances near buried utilities. However, an unexpected finding in this case is that the IF values appear to be overestimated, reaching GHz levels despite the system operating at a fundamental frequency of 450 MHz, see Figure 7-9. These artefacts were introduced by the background removal filter applied during the post-processing stage before IF generation. The influence of this filter on the IF results will be further discussed in the discussion section.

7.3.3 Change Map Combining AS and IFS

After analyzing the global trends of amplitude and instantaneous frequency (IF) changes during the two periods, the change detection methods proposed in Chapter 5 were applied to identify regions with high susceptibility to subsurface anomalies. To identify potential leakage regions, change maps were generated using amplitude and IF difference images. A size filter was applied to reduce random scattering noise, retaining only clusters with a spread size greater than 2×2 meters and a depth greater than 0.375 meters (spanning at least three slices of 0.125-meter thickness). Figure 7-10 shows the change clusters identified using amplitude difference images (Figure 7-10 (a)), IF difference images (Figure 7-10 (b)), and the intersection of both (Figure 7-10 (c)) during Period 1.



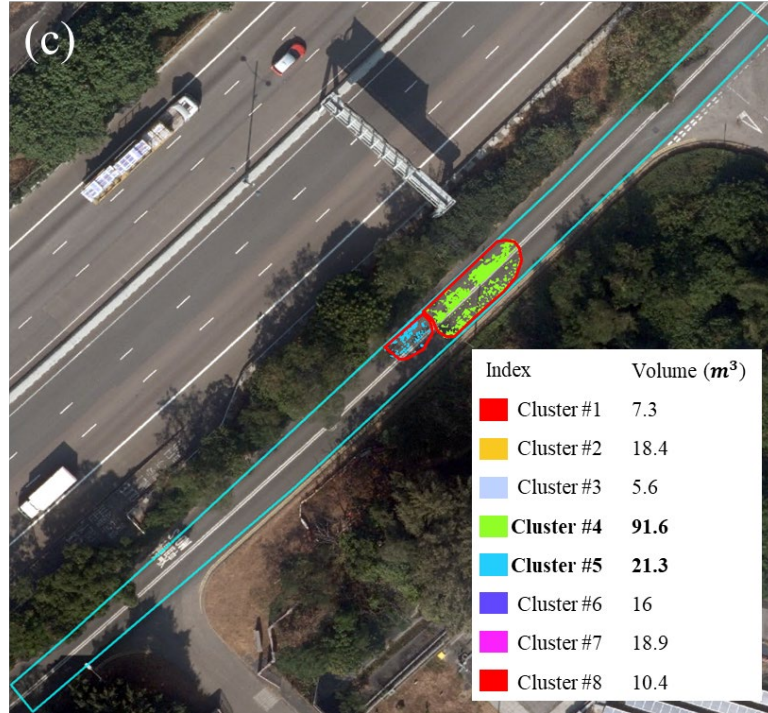


Figure 7-10 Illustration of change clusters identified for Period 1 (a) Change clusters in amplitude difference image slices; (b) Change clusters in IF difference image slices; (c) The intersection of change clusters in amplitude and IF difference images

Histograms of amplitude and IF changes within clusters #4 and #5 are shown in Figure 7-11. For cluster #4, negative amplitude changes dominate, with fewer positive amplitude changes, indicating a decrease in reflected energy during Period 1. In contrast, cluster #5 exhibits primarily positive amplitude changes, suggesting an increase in reflected energy. As discussed in Chapter 4, changes in water content can influence EM wave propagation in two primary ways: by increasing reflections, which result in positive amplitude changes, and by increasing attenuation, which causes negative amplitude changes. Regarding IF changes, both clusters show consistent negative changes, approximately -300 MHz, which align with the dispersion effect discussed in Chapters 4 and 5. Dispersion becomes more pronounced as water content increases, further supporting the interpretation of water-related changes within these clusters.

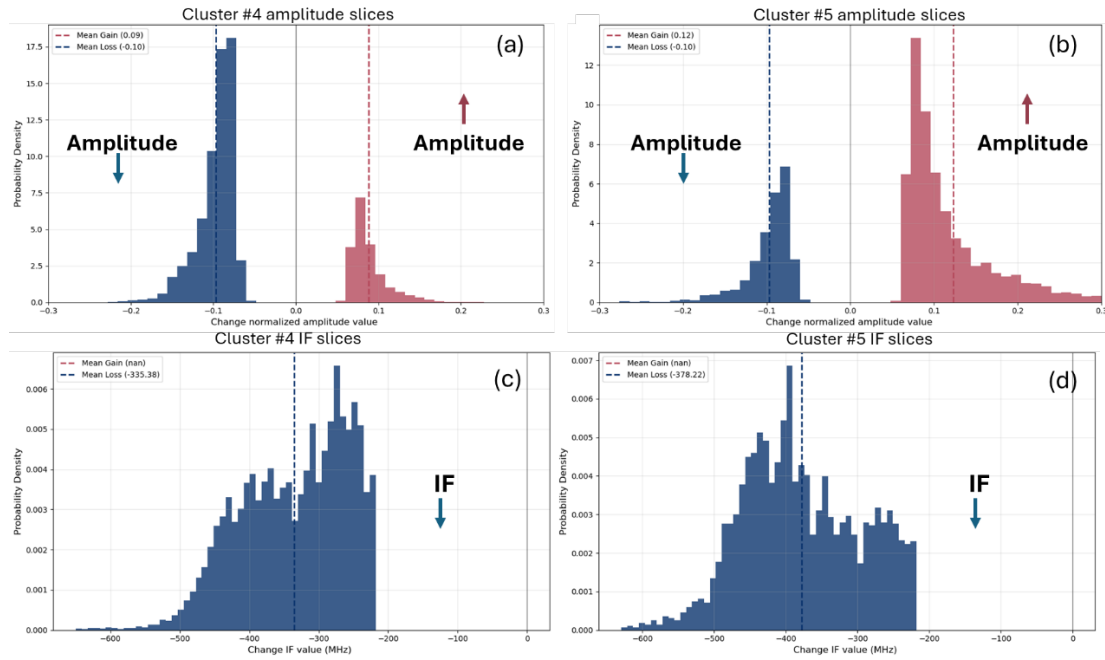


Figure 7-11 (a) Histogram of the amplitude difference volume within cluster#4; (b) Histogram of the amplitude difference volume within cluster#5; (c) Histogram of the IF difference volume within cluster#4; (d) Histogram of the IF difference volume within cluster#5

A similar analysis was conducted for Period 2, with results presented in Figure 7-12. A prominent cluster with significant amplitude and IF changes was identified, as shown in Figures 7-12 (a–c). The large volume of this cluster suggests substantial underground changes, potentially associated with water leakage. Histograms of amplitude and IF differences within this cluster, shown in Figure 7-13, reveal clear attenuation and dispersion effects, which further reinforce the likelihood of leakage.

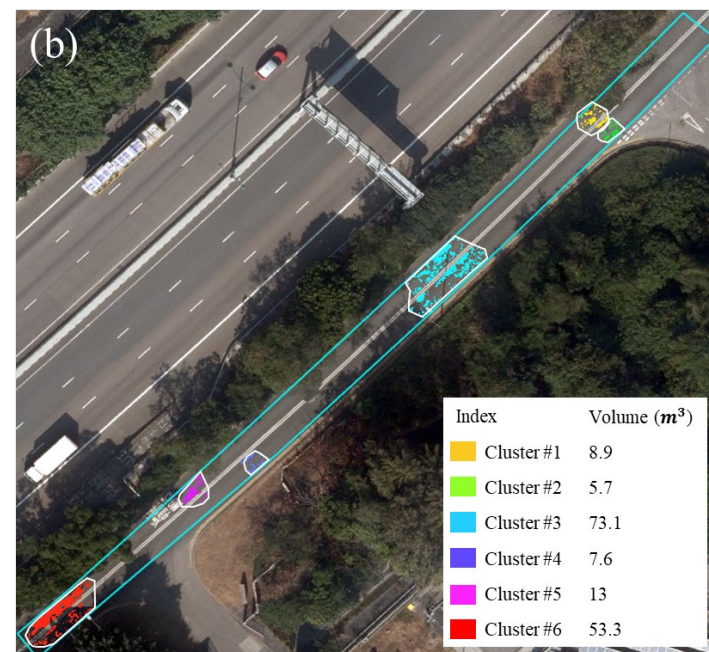
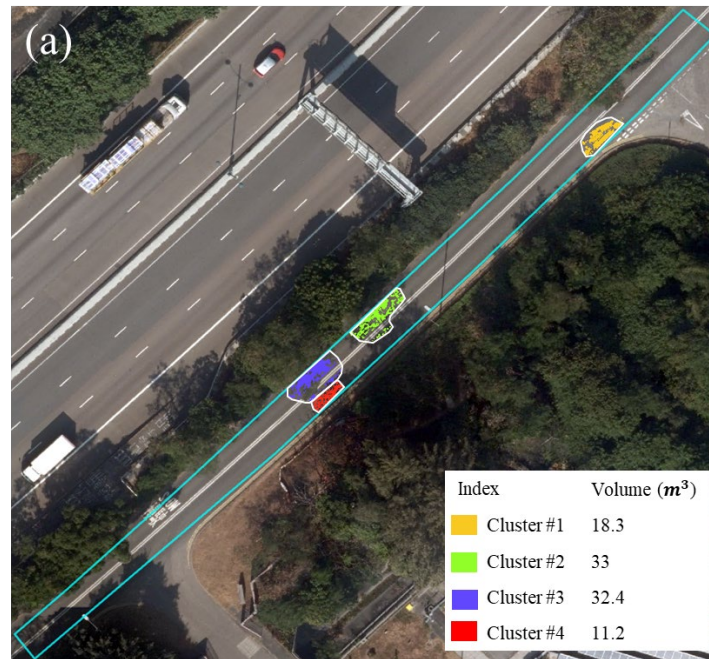




Figure 7-12 Illustration of change clusters identified for Period 2 (a) Change clusters in amplitude difference image slices; (b) Change clusters in IF difference image slices; (c) The intersection of change clusters in amplitude and IF difference images

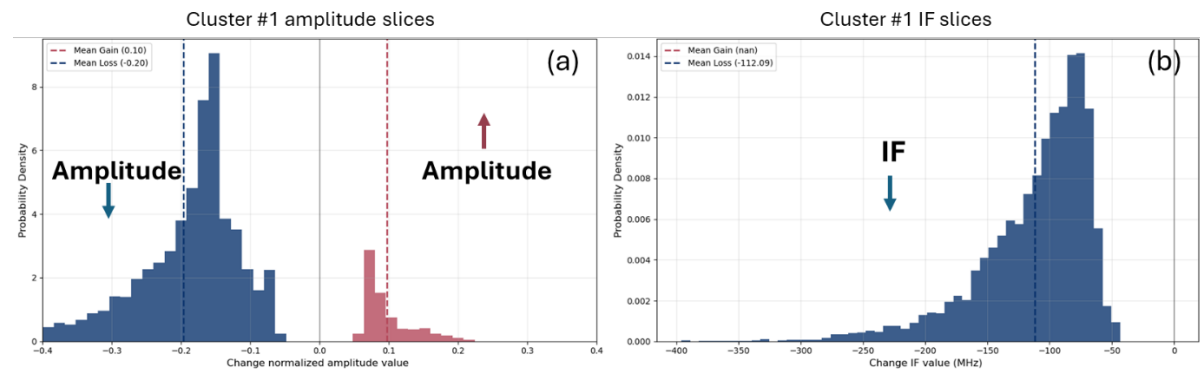


Figure 7-13 (a) Histogram of the amplitude difference volume within cluster#1; (b) Histogram of the IF difference volume within cluster#1

Figure 7-14 shows the confirmed location of a leak point overlaid with the suspected leakage cluster identified during Period 2. The leak point is located at a 1200 mm concrete pipe, buried with 1 meter of cover, near a pipe joint. The alignment of the leak point with the suspected leakage region demonstrates the effectiveness of the proposed method in detecting leakage through MCGPR-based change detection.

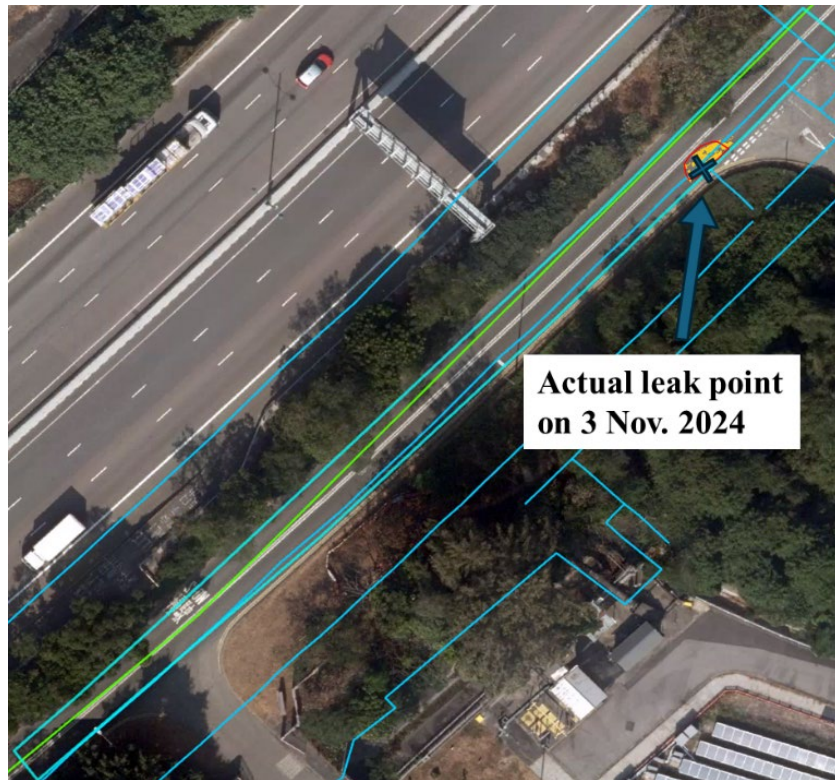


Figure 7-14 The location of confirmed leak points overlaid with the suspected leakage clusters

The results demonstrate the capability of IF slices to reveal signal dispersion associated with water content changes. However, in this case, the IF values were unexpectedly overestimated. This overestimation originates from the post-processing stage, where a background removal function was applied before IF generation. While the background removal filter is crucial for improving slice consistency and isolating meaningful subsurface features, it also introduced artefacts that compromised the accuracy of the IF values. The specific impact of this filter on the reliability of the IF results will be explored in greater detail in the following section.

7.4 Discussion and Conclusion

Unlike single-channel GPR systems, MCGPR introduces additional challenges due to channel-specific variations in the output signals. For example, the dynamic value ranges of each channel result in different amplitude responses to the same host material, as discussed in

Chapter 6. Besides, the variation of the reflected signals' dynamic range the ranging effect is another noise source. This variability arises because GPR antennas emit signals not only into the ground but also into their surroundings. If simplified using a ray-path model, the installation of MCGPR antennas creates unwanted ray paths, which manifest as horizontal "ringing" noise on radargrams. While background removal filters are commonly used to suppress this noise in GPR systems by subtracting the mean or median value of each row to keep consistency when generating amplitude slices, both for single channel GPR and MCGPR. However, when generating IF slices, the situation becomes more complex.

In single-channel systems, the ringing noise is consistent across measurements, allowing it to be reduced when generating IF difference images. This minimizes its impact on change detection tasks. Besides, this function has a risk of changing the signal phase structure, so no background removal filter has been applied in the proposed method in Chapter 4 when generating IF. In the MCGPR system, each channel exhibits unique ringing noise due to variations in antenna design and positioning, which introduces additional variability in IF slices. As shown in Figure 7-15 (a), IF slices generated without background removal filters display line-pattern strip noise caused by channel-specific ringing effects. Applying a background removal filter, as seen in Figure 7-15 (b), reduces this strip noise and improves consistency across channels. However, the filtered IF values are overestimated, with ranges exceeding the expected operating frequency of the MCGPR system (450 MHz). In contrast, the unfiltered IF values remain within a more physically reasonable range but retain some visible noise. These results demonstrate that while background removal reduces strip noise, it also overestimates IF values, potentially compromising the physical interpretation of IF changes.

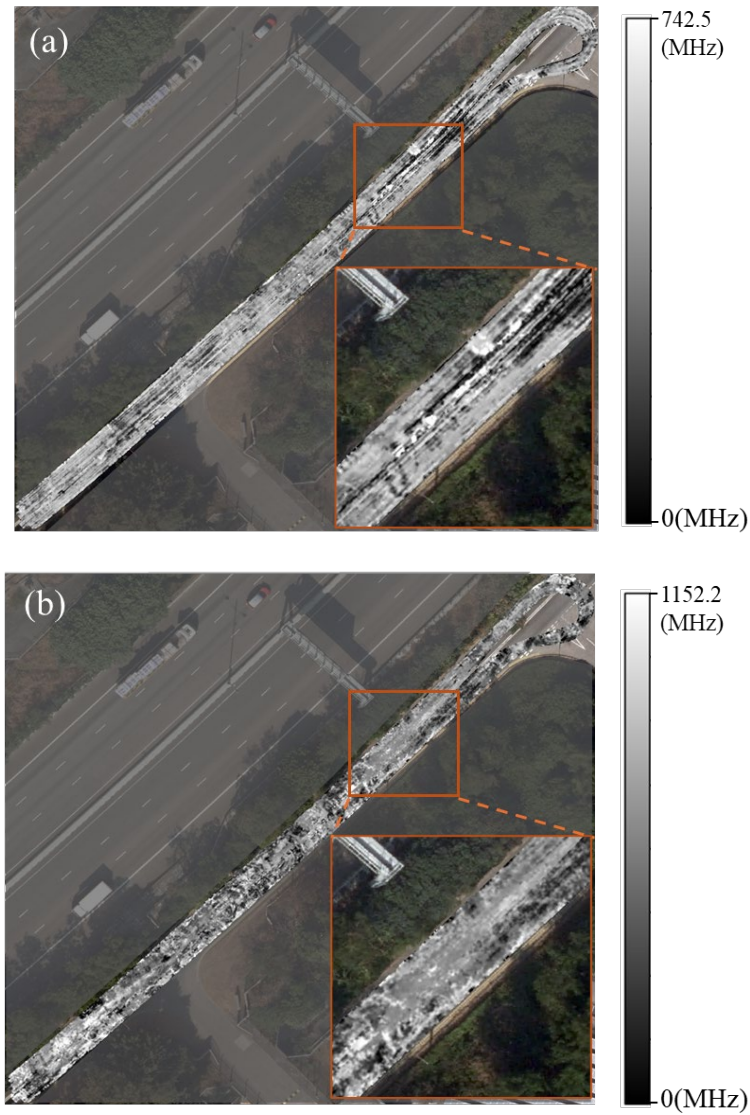


Figure 7-15 (a) Example IF slice without background removal filter. (b) Example of an IF slice with background removal filter

To further investigate the variability of ringing effects, a wavelet similarity matrix was computed for the removed ringing signals of all channel pairs, as shown in Figure 7-16. The similarity values, ranging from 0.59 to 0.91, reveal significant differences in ringing effects between channels, underscoring the need for effective noise reduction techniques when generating IF slices. These variations in ringing effects directly impact the consistency of IF slices.

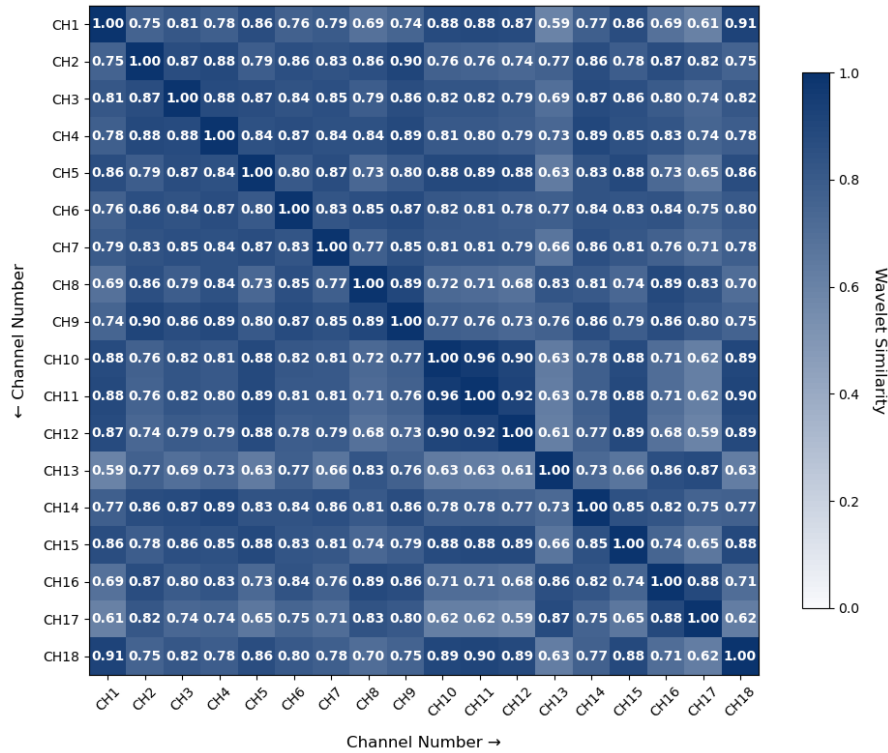


Figure 7-16 Wavelet similarity matrix of each channel pair

In the Chapter 6.2 MCGPR harmonization section, a dataset with eight different channels scans the same survey line was collected. To further investigate the impact of background removal on IF consistency, this dataset, where eight channels of MCGPR scanned the same survey line under identical subsurface conditions, was used to analyze. For shallow depths (0–15 ns), Figure 7-17 (a) shows significant variability in unfiltered IF values due to channel-specific ringing effects, although the values remain within a reasonable range. After applying a background removal filter, as shown in Figure 7-17 (b), consistency improves significantly, but at the cost of overestimating IF values and potentially losing meaningful signal information. For deeper depths (15–45 ns), the variability across channels is less prominent in the unfiltered data, and the impact of the background removal filter is reduced. At these depths, the filtered IF values exhibit a smaller degree of overestimation and better consistency.

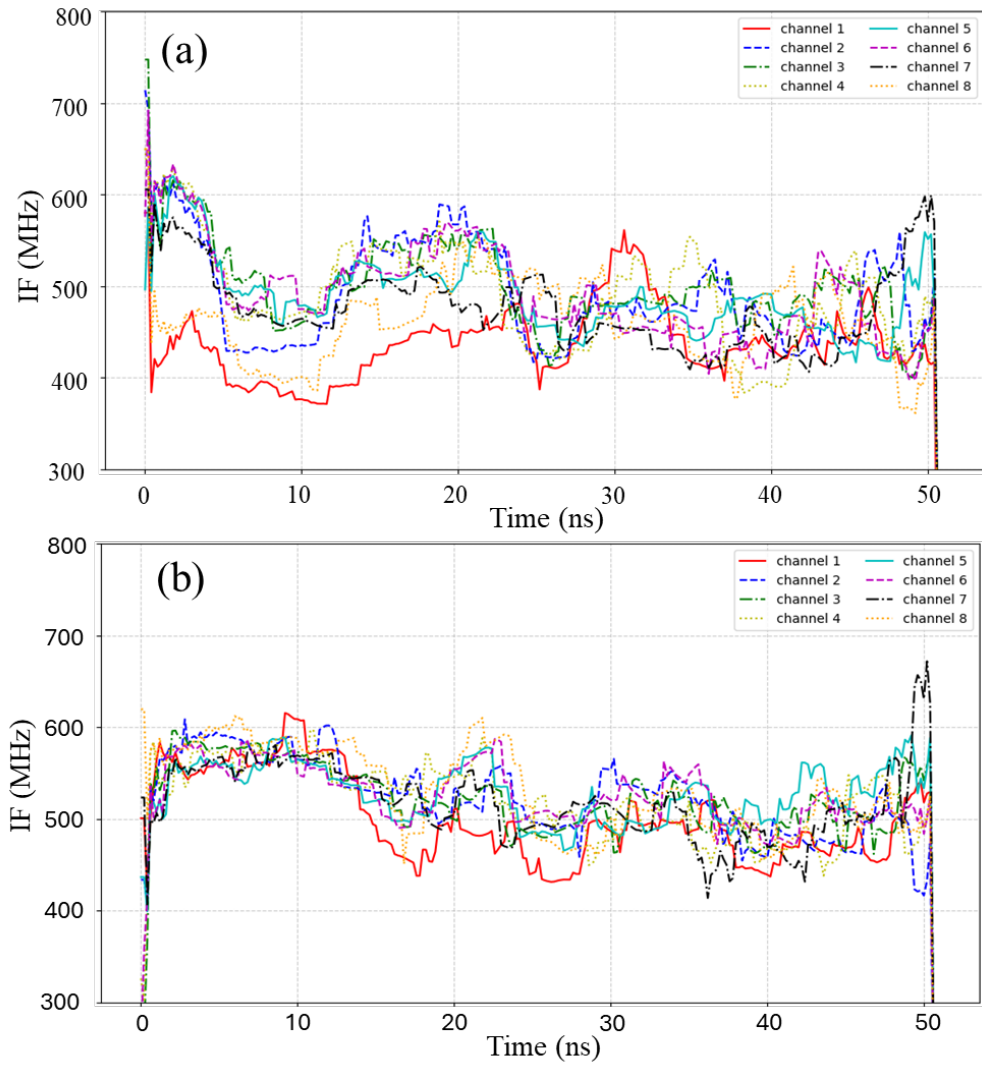


Figure 7-17 (a) IF values of different channels target the same objects without a background removal filter; (b) IF values of different channels target the same objects with a background removal filter

The depth-dependent effects of background removal were further analyzed by comparing the average IF values with and without filtering, as shown in Figure 7-18. At shallow depths (0–15 ns), the overestimation effect is most significant, reaching up to 120 MHz, indicating that background removal has a greater impact on shallow signals. However, as signal travel time increases, the overestimation effect diminishes significantly, reducing to approximately 30 MHz at greater depths (20–45 ns). This trend demonstrates that the influence of background removal decreases with depth, stabilizing at deeper regions where the relative differences between filtered and unfiltered values become stable. For change detection tasks,

the impact of overestimation caused by background removal is less critical for two main reasons. First, the overestimation effect is concentrated primarily in the shallow regions (before 15 ns in this case), where it is most pronounced. Deeper regions, which are often more relevant for detecting meaningful subsurface changes, experience minimal background removal influence. Second, when generating difference images, the stable overestimation effect at deeper depths is effectively cancelled out during the subtraction process. As a result, the difference images primarily reflect actual subsurface changes with acceptable artefacts introduced by background removal toward leakage detection tasks.

Overall, while background removal introduces some overestimation at shallow depths, its influence on change detection is minimal due to its depth-dependent nature and the subtraction process used in generating difference images.

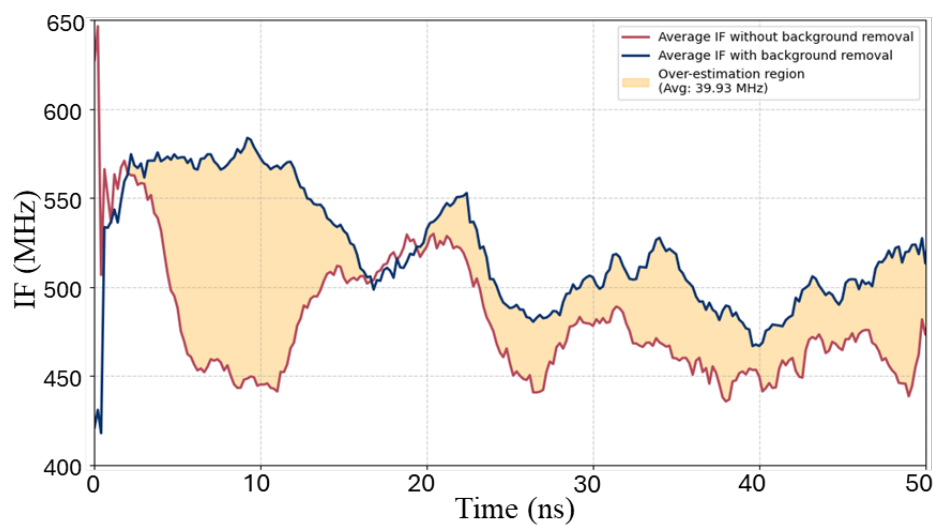


Figure 7-18 Average of the MCGPR overestimated IF values derived from the background removal filter

The proposed leakage detection method based on MCGPR change detection demonstrates its effectiveness in identifying suspected leakage regions. The results from Period 2 validate its practical potential, as the identified regions align well with the actual leakage points. However, translating this method to real-world applications introduces additional complexities. In this case, significant effort was required for essential post-processing tasks,

including geolocation enhancement and output harmonization, to produce consistent and reasonable results.

While the method delivers promising results and demonstrates robustness, it is important to recognize its limitations. One notable limitation is the background removal filter used during IF slice generation. While essential for improving slice consistency, this filter can overestimate IF values, particularly at shallow depths. For tasks relying on single IF slices rather than change detection, the impact of this overestimation becomes more critical, potentially leading to inaccuracies in subsurface interpretation. However, for change detection tasks, as highlighted earlier, this limitation is mitigated through the subtraction process.

Translating this method to real-world applications also introduces additional complexities. Significant effort was required in this study for key post-processing tasks, such as geolocation enhancement and output harmonization, to ensure consistent and realistic results. Despite these challenges, the study demonstrates the practical viability of MCGPR-based leakage detection in addressing real-world subsurface problems. Bridging the gap between controlled experiments and real-world applications remains a critical direction for future research. Efforts should focus on overcoming post-processing challenges, such as reducing the influence of channel-specific ringing noise, and further improving the accuracy and consistency of IF slices.

Ultimately, the success of MCGPR-based leakage detection in practical applications will depend on refining the methodology to address these challenges, reducing uncertainties, and enhancing system robustness. Future work should aim to develop more sophisticated post-processing methods and adaptive filtering techniques to minimize unexpected issues and ensure reliable results across diverse field conditions.

7.5 Chapter Summary and Contribution

This chapter focuses on the application of the proposed MCGPR-based methods to a real-world scenario involving a significant pipe burst. The affected area was surveyed multiple times using MCGPR at different time intervals, enabling the analysis of both temporal and spatial changes in subsurface conditions. As is often the case with underground hazards such as leakage, obtaining temporal MCGPR measurements before the observed burst event can be challenging. However, in this study, early signs of leakage were identified through water pressure anomalies, which provided crucial information for narrowing down the initial survey region and identifying preliminary suspected areas. This integration of indirect indicators with MCGPR surveys highlights a practical and effective approach for detecting underground hazards in their early stages.

Despite this uncertainty, the series of methods proposed in previous chapters was applied to detect spatial clusters characterized by significant amplitude and IF changes. While further verification using additional real-world cases is required to fully establish the robustness and effectiveness of these methods, the results demonstrate promising potential. Specifically, the detected leakage regions during Period 2 align well with the actual pipe burst location, providing partial validation of the entire post-processing workflow, including amplitude and IF slice generation, as well as change detection methods.

One of the greatest challenges encountered during this application was the variability inherent in the MCGPR system. Considerable effort was devoted to addressing issues such as geolocation misalignments and inconsistencies in amplitude dynamic ranges. Signal harmonization techniques were implemented to ensure the generated slices were comparable across different channels and surveys. While the background removal filter introduced some artefacts during IF slice generation, its impact on the change detection task was deemed

acceptable, as discussed earlier in this chapter. Nonetheless, future research will aim to develop improved methods for eliminating unwanted ringing effects while preserving signal integrity.

In summary, this chapter demonstrates the practical application of the proposed MCGPR-based methods for leakage detection. It highlights their ability to identify subsurface anomalies even under challenging conditions, such as system variability and post-processing complexities. The successful alignment of detected leakage regions with the actual pipe burst location underscores the potential of these methods for real-world applications. However, further work is needed to refine the methodology, improve robustness, and address the limitations identified during this study.

Chapter 8. Conclusion

This research systematically explored GPR imaging techniques and their application in diagnosing large-scale underground hazards using change detection strategies. It also introduced novel post-processing methods tailored to MCGPR systems, addressing key challenges in adapting these systems for practical applications. Through comprehensive laboratory experiments and real-world demonstrations, the study successfully showcased the potential of these methods in detecting primary underground hazards, marking it as the first study of its kind to address such challenges in this context.

8.1 Summary of Major Findings and Contributions to Knowledge

All objectives in Chapter 1 were accomplished. A summary of findings and the corresponding contributions to knowledge are presented as follows:

(1). **Objective 1:** Establish a Leakage **Characterization** Framework (Chapter 4).

A leakage signature database was developed, encompassing 78 leakage scenarios that consider various pavement materials and pipe attributes. This comprehensive dataset bridges the gap between simulated and real-world leakage scenarios, providing a valuable resource for the development and validation of advanced detection methods. By offering a structured and diverse collection of leakage scenarios, the database facilitates the design of methods that can adapt to real-world complexities.

To further enhance leakage characterization, IFS was introduced as a transformative complement to conventional amplitude-based methods. While AS is effective for initial anomaly detection, it faces significant limitations: its reliance on relative signal strength thresholds complicates the quantification of leakage severity, and it struggles to differentiate

between water-related changes and other subsurface anomalies. IFS addresses these challenges by extending the analysis into the time-frequency domain, leveraging the physical phenomenon of EM wave dispersion to highlight patterns uniquely associated with water leakage.

This physics-based approach distinguishes water-induced dispersion (evident in negative frequency shifts) from non-hydraulic anomalies such as voids, providing a clearer classification of detected changes. Moreover, IFS enables the quantification of leakage severity, offering deeper insights into the nature of detected change clusters.

(2). **Objective 2:** Develop MCGPR-Specific **Error Mitigation** Workflows (Chapter 6).

To improve the positioning performance of MCGPR road surveys, a post-processing method was introduced to mitigate positional distortions and misalignments, which can lead to significant data misinterpretation in temporal surveys. This method involves three key steps: (1) underground features are extracted from quality-assessed MCGPR data using advanced techniques such as the HOG and 2D FFT; (2) these features are paired based on their similarity to reconstruct relative positions within the antenna array; and (3) the entire survey trajectory is refined through global optimization using dynamic programming with constraints informed by driving behaviours. A significant advantage of this method is its ability to use MCGPR data as external information to enhance positional accuracy, ensuring compatibility with various positioning systems, particularly in urban environments.

The second algorithm, a signal harmonization method, was proposed to register signal variations within the MCGPR antenna array. Differences between antennas often introduce noise, distorting data quality and interpretation. This two-step harmonization process involves initial correction and site-specific correction, ensuring consistent amplitude and frequency spectrum reflections across the antenna array.

(3). **Objective 3:** Develop Change Detection **Algorithms** for Urban Underground Hazards (Chapter 5 and Chapter 7).

For void detection, a method combining FCM and MRF was proposed. This method balances amplitude changes and the spatial size of detected objects, ensuring the robustness of void detection. The method targets amplitude changes caused by voids, which disrupt signal propagation and cause reverberations and scattering. Laboratory experiments demonstrated its performance in identifying voids, including various sizes and shapes.

For leak detection, a hierarchical approach was introduced, combining amplitude changes with IFS to provide a more refined detection framework. Amplitude changes capture a wide range of signal variations, while IFS narrows these changes specifically to those caused by water leakage. This method was validated using the leakage signature database in Chapter 4 and a forensic investigation in a real-world leakage case in Chapter 7.

The integrated research contributions were applied and validated through a real-world forensic investigation of a leakage incident. The study began with an MCGPR survey across a 5.5 km road segment, narrowing down to a 1.6 km subsection based on water pressure anomaly records. Three subsequent MCGPR surveys were conducted, and the proposed error mitigation workflows and change detection algorithms were deployed to analyze data collected during two distinct periods (Period 1 and Period 2). The analysis successfully identified suspected leakage regions during Period 2, which closely corresponded with the actual location of the pipe burst, confirming the practical utility and accuracy of the developed methods.

MCGPR, as the most recent development in the GPR community, is not merely an antenna array stacked and towed behind a vehicle. In reality, its adoption and change-detection diagnosis of underground requires a series of fundamental research and development. This research advances underground hazard characterization by developing and validating a series

of scientific approaches. The contributions of this work are believed to benefit engineering, surveying, and infrastructure management by offering innovative tools for the non-invasive detection and prediction of underground hazards.

8.2 Challenges and Recommendations on Future Work

This research has focused on establishing a database, developing algorithms, and implementing error mitigation techniques, particularly in their effects on C-scan representations—semantic visualizations of GPR data designed to highlight key underground hazards. However, this approach may inadvertently overlook valuable information, such as the detailed features of signal propagation. Additionally, challenges remain in generating IF values from MCGPR data, particularly in addressing channel-specific ringing noise and overestimation due to background removal processes. To further enhance GPR-based diagnostics and forensic investigations, the following four key recommendations are proposed:

(1) Expand GPR Data Representations to 3D Data Volumes

With advancements in computing power, analyzing large-scale datasets has become more feasible, significantly reducing the time and labour traditionally required. While C-scans efficiently summarize specific properties of GPR data, they inherently result in information loss by compressing the dataset. Future research should explore the processing and analysis of full 3D data volumes, retaining the complete richness of the dataset. This shift would require adapting or redesigning existing processing methods to handle the challenges of large-scale data, including big data processing and visualization. Additionally, robust data standardization protocols must be established to ensure compatibility across different GPR systems, enabling scalability and reproducibility in research and practical applications.

(2) Enrich the Signature Database Through Ground Truth Validations

The availability of high-quality, validated datasets is crucial for advancing research in GPR-based diagnostics. While this study introduced a water leakage signature database, it is currently limited by the availability of real-world leakage cases with ground truth validations. Real leakage scenarios are often far more complex than simulated ones, and their inclusion would enable the development of more advanced detection methods. Collaborative efforts in data sharing within the research and engineering communities could facilitate the creation of comprehensive databases for various underground hazards. Expanding the signature database with real-world ground truth data would significantly enhance its utility, enabling more accurate model training, validation, and broader applications.

(3) Integrate Learning-Based Methods

The complexity of subsurface environments poses significant challenges to traditional GPR analysis, as reflections are influenced by numerous factors. Once a comprehensive signature database and standardized data representations are established, learning-based AI methods—such as machine learning and deep learning—can be integrated to revolutionize GPR data analysis. These methods could automate the classification of signal features, optimize representation models, and extract meaningful insights from large datasets. With a sufficiently large and validated database, AI could enable fully automated analysis, reducing reliance on human interpretation and significantly improving diagnostic efficiency. This would pave the way for next-generation GPR tools, offering scalable, data-driven solutions for underground hazard detection and characterization in the coming decades.

(4) Analyze and Refine the IF Generation Process

As highlighted in this study, the current methods for IF generation using MCGPR data face limitations, particularly in the accuracy of background removal processes, which often

lead to the overestimation of IF values. This distortion impacts the reliability of IF-based analyses, including water leakage detection. Future research should prioritize the development of improved correction and processing techniques to address these challenges, such as advanced background subtraction algorithms or adaptive filtering methods. Enhancing the accuracy of the IF generation process would not only improve water leakage detection but also expand the applicability of frequency-domain analyses to other underground hazards, strengthening the overall diagnostic capability of MCGPR systems.

In conclusion, this study has made significant strides in advancing the use of GPR imaging techniques for large-scale underground hazard detection. By developing innovative methods such as the leakage signature database, post-processing techniques for MCGPR data registration, and robust change detection frameworks for voids and leaks, this research provides a strong foundation for practical applications in infrastructure maintenance and hazard mitigation. However, challenges remain, particularly in fully leveraging the richness of GPR data and addressing limitations in IF generation for change detection. Future work should focus on expanding data representations to 3D volumes, enriching the signature database with real-world validations, integrating learning-based methods, and improving the IF generation process.

Addressing these challenges will unlock the full potential of GPR as a data-driven tool for underground infrastructure management and hazard mitigation. By building on the foundation established in this research, future advancements will enable more accurate, efficient, and scalable solutions for detecting, monitoring, and preventing subsurface anomalies, contributing to the long-term sustainability and resilience of urban infrastructure.

Reference

- Abbasov, M. E., & Gorbunova, A. A. (2025). *Dynamic Programming Algorithms for Finding Cost-Optimal Trajectory on the Terrain* (No. arXiv:2503.10922). arXiv. <https://doi.org/10.48550/arXiv.2503.10922>
- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), 2274–2282.
- Alani, A. M., Aboutalebi, M., & Kilic, G. (2013). Applications of ground penetrating radar (GPR) in bridge deck monitoring and assessment. *Journal of Applied Geophysics*, 97, 45–54.
- Ali, H., & Choi, J. (2019). A Review of Underground Pipeline Leakage and Sinkhole Monitoring Methods Based on Wireless Sensor Networking. *Sustainability*, 11(15), 4007. <https://doi.org/10.3390/su11154007>
- Angadi, S. A., & Kagawade, V. C. (2017). A robust face recognition approach through symbolic modeling of Polar FFT features. *Pattern Recognition*, 71, 235–248. <https://doi.org/10.1016/j.patcog.2017.06.014>
- Astm, D. (2011). 6432, Standard Guide for using the Surface Ground Penetrating Radar Method for Subsurface Investigation. *Annual Book of ASTM Standards*.
- Baker, J. A., Anderson, N. L., & Pilles, P. J. (1997). Ground-penetrating radar surveying in support of archeological site investigations. *Computers & Geosciences*, 23(10), 1093–1099.
- Barkataki, N., Tiru, B., & Sarma, U. (2023). Size estimation of underground targets from GPR frequency spectra: A deep learning approach. *Journal of Applied Geophysics*, 213, 105009. <https://doi.org/10.1016/j.jappgeo.2023.105009>
- Barzaghi, R., Cazzaniga, N., Pagliari, D., & Pinto, L. (2016). Vision-Based Georeferencing of GPR in Urban Areas. *Sensors*, 16(1), 132. <https://doi.org/10.3390/s16010132>
- Benedetto, A. (2010a). Water content evaluation in unsaturated soil using GPR signal analysis in the frequency domain. *Journal of Applied Geophysics*, 71(1), 26–35. <https://doi.org/10.1016/j.jappgeo.2010.03.001>
- Benedetto, A. (2010b). Water content evaluation in unsaturated soil using GPR signal analysis in the frequency domain. *Journal of Applied Geophysics*, 71(1), 26–35. <https://doi.org/10.1016/j.jappgeo.2010.03.001>
- Bernabini, M., Pettinelli, E., Pierdicca, N., Piro, S., & Versino, L. (1995). Field experiments for

- characterization of GPR antenna and pulse propagation. *Journal of Applied Geophysics*, 33(1–3), 63–76.
- Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2–3), 191–203.
- Birken, R., & Oristaglio, M. (2022). Mapping subsurface utilities with mobile electromagnetic geophysical sensor arrays. *Sensor Technologies for Civil Infrastructures*, 175–231.
- Boniger, U., & Tronicke, J. (2010). On the Potential of Kinematic GPR Surveying Using a Self-Tracking Total Station: Evaluating System Crosstalk and Latency. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10), 3792–3798. <https://doi.org/10.1109/TGRS.2010.2048332>
- Bordón, P., Bonomo, N., & Martinelli, P. (2019). Automatic detection of pipe-flange reflections in GPR data sections using supervised learning. *Journal of Applied Geophysics*, 170, 103856. <https://doi.org/10.1016/j.jappgeo.2019.103856>
- Briaud, J.-L., Ting, F., Chen, H., Cao, Y., Han, S., & Kwak, K. (2001). Erosion function apparatus for scour rate predictions. *Journal of Geotechnical and Geoenvironmental Engineering*, 127(2), 105–113.
- Broere, W. (2016). Urban underground space: Solving the problems of today's cities. *Tunnelling and Underground Space Technology*, 55, 245–248. <https://doi.org/10.1016/j.tust.2015.11.012>
- Cataldo, A., Persico, R., Leucci, G., De Benedetto, E., Cannazza, G., Matera, L., & De Giorgi, L. (2014). Time domain reflectometry, ground penetrating radar and electrical resistivity tomography: A comparative analysis of alternative approaches for leak detection in underground pipes. *NDT & E International*, 62, 14–28. <https://doi.org/10.1016/j.ndteint.2013.10.007>
- Chang, D. S., & Zhang, L. M. (2013). Extended internal stability criteria for soils under seepage. *Soils and Foundations*, 53(4), 569–583. <https://doi.org/10.1016/j.sandf.2013.06.008>
- Chen, C., Mei, S., Chen, S., Tang, Y., & Wan, C. (2022). Laboratory investigation of erosion behavior at the soil–structure interface affected by various structural factors. *Natural Hazards*, 1–20.
- Chen, G., Bai, X., Wang, G., Wang, L., Luo, X., Ji, M., Feng, P., & Zhang, Y. (2021). Subsurface Voids Detection from Limited Ground Penetrating Radar Data Using Generative Adversarial Network and YOLOV5. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 8600–8603. <https://doi.org/10.1109/IGARSS47720.2021.9554954>

- Cheung, B. W.-Y., & Lai, W. W.-L. (2019). Field validation of water-pipe leakage detection through spatial and time-lapse analysis of GPR wave velocity. *Near Surface Geophysics*, 17(3), 231–246. <https://doi.org/10.1002/nsg.12041>
- Chi, Y. M., Mallik, U., Clapp, M. A., Choi, E., Cauwenberghs, G., & Etienne-Cummings, R. (2007). CMOS camera with in-pixel temporal change detection and ADC. *IEEE Journal of Solid-State Circuits*, 42(10), 2187–2196.
- Ching, G. P. H., Chang, R. K. W., Luo, T. X. H., & Lai, W. W. L. (2021). GPR Virtual Guidance System for Subsurface 3D Imaging. *Remote Sensing*, 13(11), 2154. <https://doi.org/10.3390/rs13112154>
- Ciampoli, L. B., Calvi, A., Benedetto, A. D., Fiani, M., & Gagliardi, V. (2021). Ground Penetrating Radar (GPR) and Mobile Laser Scanner (MLS) technologies for non-destructive analysis of transport infrastructures. *Earth Resources and Environmental Remote Sensing/GIS Applications XII*, 11863, 166–174. <https://doi.org/10.1117/12.2599283>
- Cornick, M., Koechling, J., Stanley, B., & Zhang, B. (2016). Localizing Ground Penetrating RADAR: A Step Toward Robust Autonomous Ground Vehicle Localization. *Journal of Field Robotics*, 33(1), 82–102. <https://doi.org/10.1002/rob.21605>
- Crocco, L., Prisco, G., Soldovieri, F., & Cassidy, N. J. (2009). Early-stage leaking pipes GPR monitoring via microwave tomographic inversion. *Journal of Applied Geophysics*, 67(4), 270–277. <https://doi.org/10.1016/j.jappgeo.2008.09.006>
- Cui, J., Broere, W., & Lin, D. (2021). Underground space utilisation for urban renewal. *Tunnelling and Underground Space Technology*, 108, 103726. <https://doi.org/10.1016/j.tust.2020.103726>
- Daniels, D. J. (1996). Surface-penetrating radar. *Electronics & Communication Engineering Journal*, 8(4), 165–182.
- Daniels, D. J. (2004). *Ground penetrating radar* (Vol. 1). Iet.
- Dao, V. H., Ryu, H. K., & Yoon, K. B. (2021). Leak failure at the TP316L welds of a water pipe caused by microbiologically influenced corrosion. *Engineering Failure Analysis*, 122, 105244. <https://doi.org/10.1016/j.engfailanal.2021.105244>
- De Loor, G. P. (1983). The Dielectric Properties of Wet Materials. *IEEE Transactions on Geoscience and Remote Sensing*, GE-21(3), 364–369. *IEEE Transactions on Geoscience and Remote Sensing*. <https://doi.org/10.1109/TGRS.1983.350567>
- Desclée, B., Bogaert, P., & Defourny, P. (2006). Forest change detection by statistical object-based method. *Remote Sensing of Environment*, 102(1–2), 1–11.

- Doolittle, J. A., & Butnor, J. R. (2009). Soils, Peatlands, and Biomonitoring. In *Ground Penetrating Radar Theory and Applications* (pp. 177–202). Elsevier. <https://doi.org/10.1016/B978-0-444-53348-7.00006-5>
- Dragomiretskiy, K., & Zosso, D. (2014). Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, 62(3), 531–544. <https://doi.org/10.1109/TSP.2013.2288675>
- Duan, W., Kirby, R., Prisutova, J., & Horoshenkov, K. V. (2015). On the use of power reflection ratio and phase change to determine the geometry of a blockage in a pipe. *Applied Acoustics*, 87, 190–197. <https://doi.org/10.1016/j.apacoust.2014.07.002>
- Eaves, J., & Reedy, E. (2012). *Principles of modern radar*. Springer Science & Business Media. <https://books.google.com/books?hl=en&lr=&id=NHXjBwAAQBAJ&oi=fnd&pg=PA1&dq=Eaves,+J.L.+and+Reedy,+E.K.,+1987,+Principles+of+Modern+Radar,+Van+Nostrand+Reinhold+Company,+New+York,&ots=j11CiS3mJy&sig=7YRbKig1oLJWHtObhr2HU59Fn6w>
- El Amine Ben Seghier, M., Keshtegar, B., Tee, K. F., Zayed, T., Abbassi, R., & Trung, N. T. (2020). Prediction of maximum pitting corrosion depth in oil and gas pipelines. *Engineering Failure Analysis*, 112, 104505. <https://doi.org/10.1016/j.engfailanal.2020.104505>
- Elvis Peter Joseph. (2023). *Grade VS Stage In Cancer*. <https://www.manipalhospitals.com/millersroad/blog/grade-vs-stage-in-cancer/>
- Emily Hung, Oscar Liu, & Denise Tsang. (2024). Tens of thousands in Hong Kong's Tung Chung left without water supply after pipe bursts. *South China Morning Post*.
- Fell, R., Wan, C. F., Cyganiewicz, J., & Foster, M. (2003). Time for development of internal erosion and piping in embankment dams. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(4), 307–314.
- Forero-Ortiz, E., & Martínez-Gomariz, E. (2020). Hazards threatening underground transport systems. *Natural Hazards*, 100(3), 1243–1261. <https://doi.org/10.1007/s11069-020-03860-w>
- Foster, M., & Fell, R. (2001). Assessing embankment dam filters that do not satisfy design criteria. *Journal of Geotechnical and Geoenvironmental Engineering*, 127(5), 398–407.
- Foster, M., Fell, R., & Spannagle, M. (2000). The statistics of embankment dam failures and accidents. *Canadian Geotechnical Journal*, 37(5), 1000–1024.
- Gabryś, M., & Ortyl, Ł. (2020). Georeferencing of Multi-Channel GPR—Accuracy and Efficiency of Mapping of Underground Utility Networks. *Remote Sensing*, 12(18), Article 18. <https://doi.org/10.3390/rs12182945>

- Gan, G., & Cheng, J. (2011). Pedestrian Detection Based on HOG-LBP Feature. *2011 Seventh International Conference on Computational Intelligence and Security*, 1184–1187. <https://doi.org/10.1109/CIS.2011.262>
- Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., & Hunaidi, O. (2005). On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. *Journal of Sound and Vibration*, 283(3–5), 927–941. <https://doi.org/10.1016/j.jsv.2004.05.004>
- GDAL/OGR contributors. (2025). *GDAL/OGR Geospatial Data Abstraction software Library*. Open Source Geospatial Foundation. <https://doi.org/10.5281/zenodo.5884351>
- Giannopoulos, A. (2005). Modelling ground penetrating radar by GprMax. *Construction and Building Materials*, 19(10), 755–762. <https://doi.org/10.1016/j.conbuildmat.2005.06.007>
- Goodman, D., Piro, S., & others. (2013). *GPR remote sensing in archaeology* (Vol. 9). Springer.
- Grasmueck, M., & Viggiano, D. A. (2007). Integration of Ground-Penetrating Radar and Laser Position Sensors for Real-Time 3-D Data Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 45(1), 130–137. <https://doi.org/10.1109/TGRS.2006.882253>
- Gupta, S. K., & Gupta, B. K. (1979). The critical soil moisture content in the underground corrosion of mild steel. *Corrosion Science*, 19(3), 171–178. [https://doi.org/10.1016/0010-938X\(79\)90015-5](https://doi.org/10.1016/0010-938X(79)90015-5)
- Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122, 66–74.
- Harada, A., & Miyazawa, Y. (2013). Dynamic Programming Applications to Flight Trajectory Optimization. *IFAC Proceedings Volumes*, 46(19), 441–446. <https://doi.org/10.3182/20130902-5-DE-2040.00145>
- He, B., Han, P., Lu, C., & Bai, X. (2015). Effect of soil particle size on the corrosion behavior of natural gas pipeline. *Engineering Failure Analysis*, 58, 19–30. <https://doi.org/10.1016/j.engfailanal.2015.08.027>
- He, W., Hao, T., Ke, H., Zheng, W., & Lin, K. (2020). Joint time-frequency analysis of ground penetrating radar data based on variational mode decomposition. *Journal of Applied Geophysics*, 181, 104146. <https://doi.org/10.1016/j.jappgeo.2020.104146>
- He, W., & Lai, W. W. L. (2024). Unified optimization-based analysis of GPR hyperbolic fitting models. *Tunnelling and Underground Space Technology*, 146, 105633. <https://doi.org/10.1016/j.tust.2024.105633>
- Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012

- using Landsat-derived time-series metrics. *Remote Sensing of Environment*, 170, 121–132.
- Hong, S., Wiggerhauser, H., Helmerich, R., Dong, B., Dong, P., & Xing, F. (2017). Long-term monitoring of reinforcement corrosion in concrete using ground penetrating radar. *Corrosion Science*, 114, 123–132. <https://doi.org/10.1016/j.corsci.2016.11.003>
- Hong, W.-T., & Lee, J.-S. (2018). Estimation of ground cavity configurations using ground penetrating radar and time domain reflectometry. *Natural Hazards*, 92(3), 1789–1807. <https://doi.org/10.1007/s11069-018-3278-z>
- Hou, F., Liu, X., Fan, X., & Guo, Y. (2022). DL-Aided Underground Cavity Morphology Recognition Based on 3D GPR Data. *Mathematics*, 10(15), Article 15. <https://doi.org/10.3390/math10152806>
- Hu, Q., Zheng, Z., Liu, H., & Chen, B. (2020). Application of 3D Ground Penetrating Radar to Leakage Detection of Urban Underground Pipes. *Journal of Tongji University*, 48(7), 972–981. <https://doi.org/10.11908/j.issn.0253-374x.19395>
- Huber, E., & Hans, G. (2018). RGPR — An open-source package to process and visualize GPR data. *2018 17th International Conference on Ground Penetrating Radar (GPR)*, 1–4. <https://doi.org/10.1109/ICGPR.2018.8441658>
- Hui, C. K., Luo, T. X., Lai, W. W., & Chang, R. K. (2022). GPR mapping with mobile mapping sensing and tracking technologies. *Tunnelling and Underground Space Technology*, 122, 104362. <https://doi.org/10.1016/j.tust.2022.104362>
- Huisman, J. A., Hubbard, S. S., Redman, J. D., & Annan, A. P. (2003). Measuring Soil Water Content with Ground Penetrating Radar: A Review. *Vadose Zone Journal*, 2.
- Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, 91–106.
- Hussein Farh, H. M., Ben Seghier, M. E. A., Taiwo, R., & Zayed, T. (2023). Analysis and ranking of corrosion causes for water pipelines: A critical review. *Npj Clean Water*, 6(1), 65. <https://doi.org/10.1038/s41545-023-00275-5>
- Inglada, J., & Mercier, G. (2007). A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5), 1432–1445.
- Ismail, A., & Elshamy, A. (2009). Engineering behaviour of soil materials on the corrosion of mild steel. *Applied Clay Science*, 42(3–4), 356–362. <https://doi.org/10.1016/j.clay.2008.03.003>

- Jol, H. M. (2009a). *Ground penetrating radar: Theory and applications* (1st ed). Elsevier Science.
- Jol, H. M. (2009b). *Ground penetrating radar: Theory and applications* (1st ed). Elsevier Science.
- Kang, M.-S., Kim, N., Im, S. B., Lee, J.-J., & An, Y.-K. (2019a). 3D GPR Image-based U-Net for Enhancing Underground Cavity Detectability. *Remote Sensing*, *11*(21), Article 21. <https://doi.org/10.3390/rs11212545>
- Kang, M.-S., Kim, N., Im, S. B., Lee, J.-J., & An, Y.-K. (2019b). 3D GPR Image-based U-Net for Enhancing Underground Cavity Detectability. *Remote Sensing*, *11*(21), Article 21. <https://doi.org/10.3390/rs11212545>
- Kang, M.-S., Kim, N., Im, S. B., Lee, J.-J., & An, Y.-K. (2019c). 3D GPR Image-based U-Net for Enhancing Underground Cavity Detectability. *Remote Sensing*, *11*(21), Article 21. <https://doi.org/10.3390/rs11212545>
- Kang, M.-S., Kim, N., Lee, J. J., & An, Y.-K. (2020). Deep learning-based automated underground cavity detection using three-dimensional ground penetrating radar. *Structural Health Monitoring*, *19*(1), 173–185. <https://doi.org/10.1177/1475921719838081>
- Kaoser, S., Barrington, S., Elektorowicz, M., & Ayadat, T. (2006). The influence of hydraulic gradient and rate of erosion on hydraulic conductivity of sand-bentonite mixtures. *Soil & Sediment Contamination*, *15*(5), 481–496.
- Khudoyarov, S., Kim, N., & Lee, J.-J. (2020). Three-dimensional convolutional neural network-based underground object classification using three-dimensional ground penetrating radar data. *Structural Health Monitoring*, *19*(6), 1884–1893. <https://doi.org/10.1177/1475921720902700>
- Kim, J., Baek, J., & Kim, E. (2015). A Novel On-Road Vehicle Detection Method Using π HOG. *IEEE Transactions on Intelligent Transportation Systems*, *16*(6), 3414–3429. IEEE Transactions on Intelligent Transportation Systems. <https://doi.org/10.1109/TITS.2015.2465296>
- Kim, N., Kim, S., An, Y.-K., & Lee, J.-J. (2019). Triplanar Imaging of 3-D GPR Data for Deep-Learning-Based Underground Object Detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(11), 4446–4456. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. <https://doi.org/10.1109/JSTARS.2019.2953505>
- King-Wah Lau, P., Wei-Yat Cheung, B., Wai-Lok Lai, W., & Fung-Chu Sham, J. (2021).

- Characterizing pipe leakage with a combination of GPR wave velocity algorithms. *Tunnelling and Underground Space Technology*, 109, 103740. <https://doi.org/10.1016/j.tust.2020.103740>
- Klotzsche, A., Jonard, F., Looms, M. c., van der Kruk, J., & Huisman, J. a. (2018). Measuring Soil Water Content with Ground Penetrating Radar: A Decade of Progress. *Vadose Zone Journal*, 17(1), 180052. <https://doi.org/10.2136/vzj2018.03.0052>
- Klysz, G., & Balayssac, J.-P. (2007). Determination of volumetric water content of concrete using ground-penetrating radar. *Cement and Concrete Research*, 37(8), 1164–1171. <https://doi.org/10.1016/j.cemconres.2007.04.010>
- Kofman, L., Ronen, A., & Frydman, S. (2006). Detection of model voids by identifying reverberation phenomena in GPR records. *Journal of Applied Geophysics*, 59(4), 284–299. <https://doi.org/10.1016/j.jappgeo.2005.09.005>
- Kowalsky, M. B., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., Ward, A., & Gee, G. (2005). Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data. *Water Resources Research*, 41(11).
- Kwan, R. L.-Y., & Lai, W. W.-L. (2024). Archaeological investigation of burials preluded by ground penetrating radar and geospatial technologies. *Journal of Archaeological Science*, 170, 106058. <https://doi.org/10.1016/j.jas.2024.106058>
- Lai, W. L., Kind, T., & Wiggerhauser, H. (2011). Using ground penetrating radar and time–frequency analysis to characterize construction materials. *NDT & E International*, 44(1), Article 1. <https://doi.org/10.1016/j.ndteint.2010.10.002>
- Lai, W. L., Kou, S. C., & Poon, C. S. (2012). Unsaturated zone characterization in soil through transient wetting and drying using GPR joint time–frequency analysis and grayscale images. *Journal of Hydrology*, 452–453, 1–13. <https://doi.org/10.1016/j.jhydrol.2012.03.044>
- Lai, W. W. L., Chang, R. K. W., & Sham, J. F. C. (2017). Detection and imaging of city's underground void by GPR. *2017 9th International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, 1–6. <https://doi.org/10.1109/IWAGPR.2017.7996055>
- Lai, W. W. L., Chang, R. K. W., Sham, J. F. C., & Pang, K. (2016a). Perturbation mapping of water leak in buried water pipes via laboratory validation experiments with high-frequency ground penetrating radar (GPR). *Tunnelling and Underground Space Technology*, 52, 157–167. <https://doi.org/10.1016/j.tust.2015.10.017>
- Lai, W. W. L., Chang, R. K. W., Sham, J. F. C., & Pang, K. (2016b). Perturbation mapping of water leak in buried water pipes via laboratory validation experiments with high-

- frequency ground penetrating radar (GPR). *Tunnelling and Underground Space Technology*, 52, 157–167. <https://doi.org/10.1016/j.tust.2015.10.017>
- Lai, W. W. L., Dérobert, X., & Annan, P. (2018). A review of Ground Penetrating Radar application in civil engineering: A 30-year journey from Locating and Testing to Imaging and Diagnosis. *NDT & E International*, 96, 58–78. <https://doi.org/10.1016/j.ndteint.2017.04.002>
- Lai, W. W. L., Ho, M. L.-Y., Chang, R. K. W., Sham, J. F. C., & Poon, C.-S. (2017). Tracing and imaging minor water seepage of concealed PVC pipe in a reinforced concrete wall by high-frequency ground penetrating radar. *Construction and Building Materials*, 151, 840–847. <https://doi.org/10.1016/j.conbuildmat.2017.06.148>
- Lai, W. W.-L., Kind, T., Kruschwitz, S., Wöstmann, J., & Wiggenhauser, H. (2014). Spectral absorption of spatial and temporal ground penetrating radar signals by water in construction materials. *NDT & E International*, 67, 55–63. <https://doi.org/10.1016/j.ndteint.2014.06.009>
- Lai, W. W.-L., Kind, T., Sham, J. F.-C., & Wiggenhauser, H. (2016). Correction of GPR wave velocity at different oblique angles between traverses and alignment of line objects in a common offset antenna setting. *NDT & E International*, 82, 36–43. <https://doi.org/10.1016/j.ndteint.2016.03.003>
- Lai, W.-L., Kind, T., Stoppel, M., & Wiggenhauser, H. (2013). Measurement of accelerated steel corrosion in concrete using ground-penetrating radar and a modified half-cell potential method. *Journal of Infrastructure Systems*, 19(2), 205–220.
- Lee, J. H., & Lee, J. S. (2022). Trajectory optimization using dynamic programming and Q-learning. *2022 IEEE 5th Student Conference on Electric Machines and Systems (SCEMS)*, 1–4. <https://doi.org/10.1109/SCEMS56272.2022.9990639>
- Léger, E., Saintenoy, A., & Coquet, Y. (2014). *Estimating saturated hydraulic conductivity from ground-based GPR monitoring Porchet infiltration in sandy soil*. 124–130.
- Leucci, G., Negri, S., & Carrozzo, M. T. (2003). Ground Penetrating Radar (GPR): An application for evaluating the state of maintenance of the building coating. *Annals of Geophysics*, 46(3).
- Li, H., Li, N., Wu, R., Wang, H., Gui, Z., & Song, D. (2021). GPR-RCNN: An Algorithm of Subsurface Defect Detection for Airport Runway Based on GPR. *IEEE Robotics and Automation Letters*, 6(2), 3001–3008. <https://doi.org/10.1109/LRA.2021.3062599>
- Li, N., Wu, R., Li, H., Wang, H., Gui, Z., & Song, D. (2022). MV-GPRNet: Multi-View Subsurface Defect Detection Network for Airport Runway Inspection Based on GPR.

- Remote Sensing*, 14(18), Article 18. <https://doi.org/10.3390/rs14184472>
- Liu, H., Birken, R., & Wang, M. L. (2016). Automatic pavement layer identification with multichannel ground penetrating radar at traffic speed. *Journal of Applied Remote Sensing*, 10(4), 046023. <https://doi.org/10.1117/1.JRS.10.046023>
- Liu, H., Deng, X., Dai, D., Chen, Z., He, Q., & Meng, X. (2023). A Study on the Method for Detecting and Locating Leakage in Underground Pipeline. *Modern Radar*, 45(12), 7–14.
- Liu, H., Huang, Z., Yue, Y., Cui, J., & Hu, Q. (2022). Characteristics Analysis of Ground Penetrating Radar Signals for Groundwater Pipe Leakage Environment. *Journal of Electronics & Information Technology*, 44(211213), 1257.
- Liu, H., Lin, C., Cui, J., Fan, L., Xie, X., & Spencer, B. F. (2020). Detection and localization of rebar in concrete by deep learning using ground penetrating radar. *Automation in Construction*, 118, 103279. <https://doi.org/10.1016/j.autcon.2020.103279>
- Liu, S., Chen, Y., Luo, C., Jiang, H., Li, H., Li, H., & Lu, Q. (2022). Particle Swarm Optimization-Based Variational Mode Decomposition for Ground Penetrating Radar Data Denoising. *Remote Sensing*, 14(13), Article 13. <https://doi.org/10.3390/rs14132973>
- Lowekamp, B. C., Chen, D. T., Ibáñez, L., & Blezek, D. (2013). The design of SimpleITK. *Frontiers in Neuroinformatics*, 7, 45.
- Lu, Q., Guo, S., Zeng, Z., Liu, S., Liu, K., & Li, X. (2023). Application of Variational Mode decomposition in GPR Detection of Oil Contaminated Soil. *Journal of Physics: Conference Series*, 2651(1), 012035. <https://doi.org/10.1088/1742-6596/2651/1/012035>
- Luo, T. X. H., & Lai, W. W. L. (2020). GPR pattern recognition of shallow subsurface air voids. *Tunnelling and Underground Space Technology*, 99, 103355. <https://doi.org/10.1016/j.tust.2020.103355>
- Luo, T. X. H., Lai, W. W. L., Chang, R. K. W., & Goodman, D. (2019a). GPR imaging criteria. *Journal of Applied Geophysics*, 165, 37–48. <https://doi.org/10.1016/j.jappgeo.2019.04.008>
- Luo, T. X. H., Lai, W. W. L., Chang, R. K. W., & Goodman, D. (2019b). GPR imaging criteria. *Journal of Applied Geophysics*, 165, 37–48. <https://doi.org/10.1016/j.jappgeo.2019.04.008>
- Luo, T. X. H., Lai, W. W. L., & Giannopoulos, A. (2020a). Forward modelling on GPR responses of subsurface air voids. *Tunnelling and Underground Space Technology*, 103,

103521. <https://doi.org/10.1016/j.tust.2020.103521>

- Luo, T. X. H., Lai, W. W. L., & Giannopoulos, A. (2020b). Forward modelling on GPR responses of subsurface air voids. *Tunnelling and Underground Space Technology*, 103, 103521. <https://doi.org/10.1016/j.tust.2020.103521>
- Luo, T. X. H., Lai, W. W. L., & Lei, Z. (2023). Intensity Normalisation of GPR C-Scans. *Remote Sensing*, 15(5), Article 5. <https://doi.org/10.3390/rs15051309>
- Luo, T. X.-H., & Lai, W. W. L. (2020a). Subsurface Diagnosis With Time-Lapse GPR Slices and Change Detection Algorithms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 935–940. <https://doi.org/10.1109/JSTARS.2020.2975659>
- Luo, T. X.-H., & Lai, W. W. L. (2020b). Subsurface diagnosis with time-lapse GPR slices and change detection algorithms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 935–940.
- Lv, N., Chen, C., Qiu, T., & Sangaiah, A. K. (2018). Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images. *IEEE Transactions on Industrial Informatics*, 14(12), 5530–5538.
- Malagodi, S., Orlando, L., Piro, S., & Rosso, F. (1996). Location of archaeological structures using GPR method: Three-dimensional data acquisition and radar signal processing. *Archaeological Prospection*, 3(1), 13–23.
- Mehdizadeh, A., Disfani, M. M., Evans, R., Arulrajah, A., & Ong, D. (2017). Mechanical consequences of suffusion on undrained behaviour of a gap-graded cohesionless soil-an experimental approach. *Geotechnical Testing Journal*, 40(6), 1026–1042.
- Merkle, D., Frey, C., & Reiterer, A. (2021). Fusion of ground penetrating radar and laser scanning for infrastructure mapping. *Journal of Applied Geodesy*, 15(1), 31–45. <https://doi.org/10.1515/jag-2020-0004>
- Metwaly, M. (2015). Application of GPR technique for subsurface utility mapping: A case study from urban area of Holy Mecca, Saudi Arabia. *Measurement*, 60, 139–145.
- Moffat, R., Fannin, R. J., & Garner, S. J. (2011). Spatial and temporal progression of internal erosion in cohesionless soil. *Canadian Geotechnical Journal*, 48(3), 399–412.
- Momeni, A., Piratla, K. R., Anderson, A., Chalil Madathil, K., & Li, D. (2023). Stochastic model-based leakage prediction in water mains considering pipe condition uncertainties. *Tunnelling and Underground Space Technology*, 137, 105130. <https://doi.org/10.1016/j.tust.2023.105130>
- Muggleton, J. M., Brennan, M. J., & Gao, Y. (2011a). Determining the location of buried plastic

- water pipes from measurements of ground surface vibration. *Journal of Applied Geophysics*, 75(1), 54–61. <https://doi.org/10.1016/j.jappgeo.2011.06.030>
- Muggleton, J. M., Brennan, M. J., & Gao, Y. (2011b). Determining the location of buried plastic water pipes from measurements of ground surface vibration. *Journal of Applied Geophysics*, 75(1), 54–61. <https://doi.org/10.1016/j.jappgeo.2011.06.030>
- Muggleton, J. M., & Papandreou, B. (2014). A shear wave ground surface vibration technique for the detection of buried pipes. *Journal of Applied Geophysics*, 106, 164–172. <https://doi.org/10.1016/j.jappgeo.2014.04.021>
- Nie, X. H., Li, X. G., Du, C. W., & Cheng, Y. F. (2009). Temperature dependence of the electrochemical corrosion characteristics of carbon steel in a salty soil. *Journal of Applied Electrochemistry*, 39(2), 277–282. <https://doi.org/10.1007/s10800-008-9669-1>
- Noor, E. A., & Al-Moubaraki, A. H. (2014). Influence of Soil Moisture Content on the Corrosion Behavior of X60 Steel in Different Soils. *Arabian Journal for Science and Engineering*, 39(7), 5421–5435. <https://doi.org/10.1007/s13369-014-1135-2>
- Ocaña-Levario, S. J., Carreño-Alvarado, E. P., Ayala-Cabrera, D., & Izquierdo, J. (2018). GPR image analysis to locate water leaks from buried pipes by applying variance filters. *Journal of Applied Geophysics*, 152, 236–247. <https://doi.org/10.1016/j.jappgeo.2018.03.025>
- O'Day, D. K. (1989). External corrosion in distribution systems. *Journal-American Water Works Association*, 81(10), 45–52.
- Orlando, L. (2013). GPR to constrain ERT data inversion in cavity searching: Theoretical and practical applications in archeology. *Journal of Applied Geophysics*, 89, 35–47. <https://doi.org/10.1016/j.jappgeo.2012.11.006>
- Ort, T., Gilitschenski, I., & Rus, D. (2020). Autonomous Navigation in Inclement Weather Based on a Localizing Ground Penetrating Radar. *IEEE Robotics and Automation Letters*, 5(2), 3267–3274. <https://doi.org/10.1109/LRA.2020.2976310>
- Pachideh, V., & Mir Mohammad Hosseini, S. M. (2019). A new physical model for studying flow direction and other influencing parameters on the internal erosion of soils. *Geotechnical Testing Journal*, 42(6), 1431–1456.
- Paglieroni, D. W., Pechard, C. T., & Beer, N. R. (2015). Change Detection in Constellations of Buried Objects Extracted From Ground-Penetrating Radar Data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5), 2426–2439. <https://doi.org/10.1109/TGRS.2014.2360097>

- Park, B., Kim, J., Lee, J., Kang, M.-S., & An, Y.-K. (2018). Underground Object Classification for Urban Roads Using Instantaneous Phase Analysis of Ground-Penetrating Radar (GPR) Data. *Remote Sensing*, 10(9), Article 9. <https://doi.org/10.3390/rs10091417>
- Pipan, M., Baradello, L., Forte, E., Prizzon, A., & Finetti, I. (1999). 2-D and 3-D processing and interpretation of multi-fold ground penetrating radar data: A case history from an archaeological site. *Journal of Applied Geophysics*, 41(2), 271–292. [https://doi.org/10.1016/S0926-9851\(98\)00047-0](https://doi.org/10.1016/S0926-9851(98)00047-0)
- Rahnema, H., & Mirasi, S. (2012). Seismic and Geotechnical Study of Land Subsidence and Vulnerability of Rural Buildings. *International Journal of Geosciences*, 03(04), 878–884. <https://doi.org/10.4236/ijg.2012.324089>
- Rajani, B., & Kleiner, Y. (2003). Protecting ductile-iron Water Mains: What protection method works best for what soil condition? *Journal AWWA*, 95(11), 110–125. <https://doi.org/10.1002/j.1551-8833.2003.tb10497.x>
- Rasol, M., Pais, J. C., Pérez-Gracia, V., Solla, M., Fernandes, F. M., Fontul, S., Ayala-Cabrera, D., Schmidt, F., & Assadollahi, H. (2022). GPR monitoring for road transport infrastructure: A systematic review and machine learning insights. *Construction and Building Materials*, 324, 126686. <https://doi.org/10.1016/j.conbuildmat.2022.126686>
- Reichman, D., Collins, L. M., & Malof, J. M. (2017a). Improvements to the Histogram of Oriented Gradient (HOG) prescreener for buried threat detection in ground penetrating radar data. *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXII*, 10182, 166–174. <https://doi.org/10.1117/12.2263107>
- Reichman, D., Collins, L. M., & Malof, J. M. (2017b). Learning improved pooling regions for the Histogram of Oriented Gradient (HOG) feature for buried threat detection in ground penetrating radar. *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXII*, 10182, 175–187. <https://doi.org/10.1117/12.2263108>
- Reza, A. M. (2004). Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, 38, 35–44.
- Ristic, A. V., Petrovacki, D., & Govedarica, M. (2009). A new method to simultaneously estimate the radius of a cylindrical object and the wave propagation velocity from GPR data. *Computers & Geosciences*, 35(8), 1620–1630.
- Rokni, K., Ahmad, A., Solaimani, K., & Hazini, S. (2015). A new approach for surface water change detection: Integration of pixel level image fusion and image classification techniques. *International Journal of Applied Earth Observation and Geoinformation*, 34, 226–234.

- Rowell, C., Pidlisecky, A., Irving, J., & Ferguson, R. (2010). Characterization of lava tubes using ground penetrating radar at Craters of the Moon National Monument. *Idaho, USA: CREWES Research Report*, 22.
- Salazar, S. E., Barnes, A., & Coffman, R. A. (2015). Development of an internal camera-based volume determination system for triaxial testing. *Geotechnical Testing Journal*, 38(4), 549–555.
- Sandirasegaram, N., & English, R. (2005). Comparative analysis of feature extraction (2D FFT and wavelet) and classification (Lp metric distances, MLP NN, and HNeT) algorithms for SAR imagery. *Algorithms for Synthetic Aperture Radar Imagery XII*, 5808, 314–325. <https://doi.org/10.1117/12.597305>
- Šarlah, N., Podobnikar, T., Ambrožič, T., & Mušič, B. (2020). Application of Kinematic GPR-TPS Model with High 3D Georeference Accuracy for Underground Utility Infrastructure Mapping: A Case Study from Urban Sites in Celje, Slovenia. *Remote Sensing*, 12(8), 1228. <https://doi.org/10.3390/rs12081228>
- Seo, J., Park, W., & Kim, T. (2022). Feature-based approach to change detection of small objects from high-resolution satellite images. *Remote Sensing*, 14(3), 462.
- Sham, J. F. C., & Lai, W. W. L. (2016). Development of a new algorithm for accurate estimation of GPR's wave propagation velocity by common-offset survey method. *NDT & E International*, 83, 104–113. <https://doi.org/10.1016/j.ndteint.2016.05.002>
- Shang, J. Q., & Umana, J. A. (1999a). Dielectric Constant and Relaxation Time of Asphalt Pavement Materials. *Journal of Infrastructure Systems*, 5(4), 135–142. [https://doi.org/10.1061/\(ASCE\)1076-0342\(1999\)5:4\(135\)](https://doi.org/10.1061/(ASCE)1076-0342(1999)5:4(135))
- Shang, J. Q., & Umana, J. A. (1999b). Dielectric Constant and Relaxation Time of Asphalt Pavement Materials. *Journal of Infrastructure Systems*, 5(4), Article 4. [https://doi.org/10.1061/\(ASCE\)1076-0342\(1999\)5:4\(135\)](https://doi.org/10.1061/(ASCE)1076-0342(1999)5:4(135))
- Simanjuntak, R. M., Kuffer, M., & Reckien, D. (2019). Object-based image analysis to map local climate zones: The case of Bandung, Indonesia. *Applied Geography*, 106, 108–121.
- Sterpi, D. (2003). Effects of the erosion and transport of fine particles due to seepage flow. *International Journal of Geomechanics*, 3(1), 111–122.
- Stone, K., Keller, J. M., & Shaw, D. (2015). Multi-scale HOG prescreening algorithm for detection of buried explosive hazards in FL-IR and FL-GPR data. *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XX*, 9454, 262–276. <https://doi.org/10.1117/12.2177663>

- Stow, D. (1999). Reducing the effects of misregistration on pixel-level change detection. *International Journal of Remote Sensing*, 20(12), 2477–2483.
- Su, Y., Wang, J., Li, D., Wang, X., Hu, L., Yao, Y., & Kang, Y. (2023). End-to-end deep learning model for underground utilities localization using GPR. *Automation in Construction*, 149, 104776. <https://doi.org/10.1016/j.autcon.2023.104776>
- Su, Y., Wang, J., Wang, X., Yao, Y., & Shou, W. (2023). 3D reconstruction in underground utilities. *Automation in Construction*, 156, 105100. <https://doi.org/10.1016/j.autcon.2023.105100>
- Suban, M., Cvelbar, R., & Bundara, B. (n.d.). *THE IMPACT OF STAGNANT WATER ON THE CORROSION PROCESSES IN A PIPELINE*.
- Sugak, V. G., & Sugak, A. V. (2010). Phase Spectrum of Signals in Ground-Penetrating Radar Applications. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4), 1760–1767. *IEEE Transactions on Geoscience and Remote Sensing*. <https://doi.org/10.1109/TGRS.2009.2036163>
- Sukhanov, D. Ya., Ponomarev, O. G., Zavyalova, K. V., Khmelev, V. L., & Roslyakov, S. N. (2017). Radar with a local positioning video-system. *2017 Progress In Electromagnetics Research Symposium - Spring (PIERS)*, 3723–3728. <https://doi.org/10.1109/PIERS.2017.8262405>
- Tan, F., Tan, W., Yan, F., Qi, X., Li, Q., & Hong, Z. (2022). Model Test Analysis of Subsurface Cavity and Ground Collapse Due to Broken Pipe Leakage. *Applied Sciences*, 12(24), Article 24. <https://doi.org/10.3390/app122413017>
- Tanoli, W. A., Sharafat, A., Park, J., & Seo, J. W. (2019). Damage Prevention for underground utilities using machine guidance. *Automation in Construction*, 107, 102893. <https://doi.org/10.1016/j.autcon.2019.102893>
- Torrione, P. A., Morton, K. D., Sakaguchi, R., & Collins, L. M. (2014). Histograms of Oriented Gradients for Landmine Detection in Ground-Penetrating Radar Data. *IEEE Transactions on Geoscience and Remote Sensing*, 52(3), 1539–1550. *IEEE Transactions on Geoscience and Remote Sensing*. <https://doi.org/10.1109/TGRS.2013.2252016>
- Tzanis, A. (2015). The Curvelet Transform in the analysis of 2-D GPR data: Signal enhancement and extraction of orientation-and-scale-dependent information. *Journal of Applied Geophysics*, 115, 145–170. <https://doi.org/10.1016/j.jappgeo.2015.02.015>
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., & contributors, the scikit-image. (2014). scikit-image: Image

- processing in Python. *PeerJ*, 2, e453. <https://doi.org/10.7717/peerj.453>
- Walter, V. (2004). Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3–4), 225–238.
- Wan, C. F., & Fell, R. (2004). Laboratory tests on the rate of piping erosion of soils in embankment dams. *Geotechnical Testing Journal*, 27(3), 295–303.
- Wang, M., & Yin, X. (2022). Construction and maintenance of urban underground infrastructure with digital technologies. *Automation in Construction*, 141, 104464. <https://doi.org/10.1016/j.autcon.2022.104464>
- Wang, W., Li, C.-Q., & Shi, W. (2019). Degradation of mechanical property of corroded water pipes after long service. *Urban Water Journal*, 16(7), 494–504. <https://doi.org/10.1080/1573062X.2019.1687744>
- Wang, X.-W., & Xu, Y.-S. (2022). Investigation on the phenomena and influence factors of urban ground collapse in China. *Natural Hazards*, 113(1), 1–33. <https://doi.org/10.1007/s11069-022-05304-z>
- Wang, Y., Qin, H., Tang, Y., Zhang, D., Yang, D., Qu, C., & Geng, T. (2022). RCE-GAN: A Rebar Clutter Elimination Network to Improve Tunnel Lining Void Detection from GPR Images. *Remote Sensing*, 14(2), Article 2. <https://doi.org/10.3390/rs14020251>
- Warren, C., Giannopoulos, A., & Giannakis, I. (2016). gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar. *Computer Physics Communications*, 209, 163–170. <https://doi.org/10.1016/j.cpc.2016.08.020>
- Warren, C., Giannopoulos, A., Gray, A., Giannakis, I., Patterson, A., Wetter, L., & Hamrah, A. (2019). A CUDA-based GPU engine for gprMax: Open source FDTD electromagnetic simulation software. *Computer Physics Communications*, 237, 208–218. <https://doi.org/10.1016/j.cpc.2018.11.007>
- Wasim, M., Shoaib, S., Mubarak, N. M., Inamuddin, & Asiri, A. M. (2018). Factors influencing corrosion of metal pipes in soils. *Environmental Chemistry Letters*, 16(3), 861–879. <https://doi.org/10.1007/s10311-018-0731-x>
- Water Supplies Department. (2025). *Number of Water Main Bursts*. <https://www.wsd.gov.hk/en/core-businesses/major-infrastructure-projects/risk-based-improvement-programme-of-water-mains/number-of-cases-includes-damage-by-other-parties-/index.html>
- Wei, Y., Avci, C., Liu, J., Belezamo, B., Aydın, N., Li, P., & Zhou, X. (2017). Dynamic programming-based multi-vehicle longitudinal trajectory optimization with simplified car following models. *Transportation Research Part B: Methodological*, 106, 102–129.

<https://doi.org/10.1016/j.trb.2017.10.012>

- Wei, Y., Tian, Q., Guo, J., Huang, W., & Cao, J. (2019). Multi-vehicle detection algorithm through combining Harr and HOG features. *Mathematics and Computers in Simulation*, 155, 130–145. <https://doi.org/10.1016/j.matcom.2017.12.011>
- Weil, G. J. (1995). *Remote infrared thermal sensing of sewer voids*. 2454, 229–237.
- Wen-Zhong Shi, An-Shu Zhang, & Ho, O.-K. (2013). Spatial analysis of water mains failure clusters and factors: A Hong Kong case study. *Annals of GIS*, 19(2), 89–97. <https://doi.org/10.1080/19475683.2013.782509>
- World Bank. (2024). *Urban Development Overview*. <https://www.worldbank.org/en/topic/urbandevelopment/overview>
- Xiao, X., Ihamouten, A., Villain, G., & Dérobert, X. (2017). Use of electromagnetic two-layer wave-guided propagation in the GPR frequency range to characterize water transfer in concrete. *NDT & E International*, 86, 164–174. <https://doi.org/10.1016/j.ndteint.2016.08.001>
- Xie, F., Wu, C. G.-W., Lai, W. W.-L., & Sham, J. F.-C. (2018). Correction of multi-frequency GPR wave velocity with distorted hyperbolic reflections from GPR surveys of underground utilities. *Tunnelling and Underground Space Technology*, 76, 76–91. <https://doi.org/10.1016/j.tust.2018.02.005>
- Xu, Q., Gao, H., Pang, Z., & Lei, W. (2024). GPR Bscan Change Detection Network for Structural Defect Evolution. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–15. <https://doi.org/10.1109/TGRS.2024.3480122>
- Xu, X., Peng, S., Xia, Y., & Ji, W. (2014). The development of a multi-channel GPR system for roadbed damage detection. *Microelectronics Journal*, 45(11), 1542–1555. <https://doi.org/10.1016/j.mejo.2014.09.004>
- Xu, Y., & Chen, W. (2018). Performance Analysis of GPS/BDS Dual/Triple-Frequency Network RTK in Urban Areas: A Case Study in Hong Kong. *Sensors*, 18(8), Article 8. <https://doi.org/10.3390/s18082437>
- Yamaguchi, T., & Mizutani, T. (2021). Detection and localization of manhole and joint covers in radar images by support vector machine and Hough transform. *Automation in Construction*, 126, 103651. <https://doi.org/10.1016/j.autcon.2021.103651>
- Yamaguchi, T., Mizutani, T., Meguro, K., & Hirano, T. (2022a). Detecting Subsurface Voids From GPR Images by 3-D Convolutional Neural Network Using 2-D Finite Difference Time Domain Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3061–3073. <https://doi.org/10.1109/JSTARS.2022.3165660>

- Yamaguchi, T., Mizutani, T., Meguro, K., & Hirano, T. (2022b). Detecting Subsurface Voids From GPR Images by 3-D Convolutional Neural Network Using 2-D Finite Difference Time Domain Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3061–3073. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. <https://doi.org/10.1109/JSTARS.2022.3165660>
- Yamaguchi, T., Mizutani, T., Meguro, K., & Hirano, T. (2022c). Detecting Subsurface Voids From GPR Images by 3-D Convolutional Neural Network Using 2-D Finite Difference Time Domain Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3061–3073. <https://doi.org/10.1109/JSTARS.2022.3165660>
- Yamaguchi, T., Mizutani, T., & Nagayama, T. (2021). Mapping Subsurface Utility Pipes by 3-D Convolutional Neural Network and Kirchhoff Migration Using GPR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(8), 6525–6536. *IEEE Transactions on Geoscience and Remote Sensing*. <https://doi.org/10.1109/TGRS.2020.3030079>
- Yao, S., Pan, S., Wang, T., Zheng, C., Shen, W., & Chong, Y. (2015). A new pedestrian detection method based on combined HOG and LSS features. *Neurocomputing*, 151, 1006–1014. <https://doi.org/10.1016/j.neucom.2014.08.080>
- Yelf, R. (2004). Where is true time zero? *Proceedings of the Tenth International Conference on Grounds Penetrating Radar, 2004. GPR 2004.*, 1, 279–282.
- Yu, Q., Li, Y., Luo, T., Zhang, J., Tao, L., Zhu, X., Zhang, Y., Luo, L., & Xu, X. (2023). Cement pavement void detection algorithm based on GPR signal and continuous wavelet transform method. *Scientific Reports*, 13(1), 19710. <https://doi.org/10.1038/s41598-023-46752-2>
- Yu, Y., Safari, A., Niu, X., Drinkwater, B., & Horoshenkov, K. V. (2021). Acoustic and ultrasonic techniques for defect detection and condition monitoring in water and sewerage pipes: A review. *Applied Acoustics*, 183, 108282. <https://doi.org/10.1016/j.apacoust.2021.108282>
- Żakowski, K., & Sokółski, W. (1999). 24-hour characteristic of interaction on pipelines of stray currents leaking from tram tractions. *Corrosion Science*, 41(11), 2099–2111. [https://doi.org/10.1016/S0010-938X\(99\)00034-7](https://doi.org/10.1016/S0010-938X(99)00034-7)
- Zhang, J., Hu, Q., Zhou, Y., Zhao, P., & Duan, X. (2024). A Multi-Level Robust Positioning Method for Three-Dimensional Ground Penetrating Radar (3D GPR) Road Underground Imaging in Dense Urban Areas. *Remote Sensing*, 16(9), 1559. <https://doi.org/10.3390/rs16091559>
- Zhang, J., Lu, Y., Yang, Z., Zhu, X., Zheng, T., Liu, X., Tian, Y., & Li, W. (2022). Recognition

- of void defects in airport runways using ground-penetrating radar and shallow CNN. *Automation in Construction*, 138, 104260. <https://doi.org/10.1016/j.autcon.2022.104260>
- Zhang, M., & Shi, W. (2020). A feature difference convolutional neural network-based change detection method. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10), 7232–7246.
- Zhang, X., Feng, X., Nilot, E., & Zhang, M. (2018). Noise suppression of GPR data using Variational Mode Decomposition. *2018 17th International Conference on Ground Penetrating Radar (GPR)*, 1–4. <https://doi.org/10.1109/ICGPR.2018.8441627>
- Zhang, X., Nilot, E., Feng, X., Ren, Q., & Zhang, Z. (2018). IMF-Slices for GPR Data Processing Using Variational Mode Decomposition Method. *Remote Sensing*, 10(3), Article 3. <https://doi.org/10.3390/rs10030476>
- Zhang, Y., Candra, P., Wang, G., & Xia, T. (2015). 2-D Entropy and Short-Time Fourier Transform to Leverage GPR Data Analysis Efficiency. *IEEE Transactions on Instrumentation and Measurement*, 64(1), 103–111. *IEEE Transactions on Instrumentation and Measurement*. <https://doi.org/10.1109/TIM.2014.2331429>
- Zhou, Y., & Lai, W. W. L. (2024). Characterization of leakage signatures in buried water pipes by ground penetrating radar(GPR) and instantaneous frequency analysis. *Tunnelling and Underground Space Technology*, 153, 105984. <https://doi.org/10.1016/j.tust.2024.105984>
- Zhou, Y., Lai, W. W. L., & Zhu, X. (2023). Superpixel-based change detection for GPR time-lapse slices using fuzzy c-means and the Markov random field method. *Tunnelling and Underground Space Technology*, 141, 105369. <https://doi.org/10.1016/j.tust.2023.105369>
- Zhu, H., Xiao, F., Zhou, Y., Lai, W. W. L., & Zhang, Q. (2024). A framework for GPR-based water leakage detection by integrating hydromechanical modelling into electromagnetic modelling. *Near Surface Geophysics*, 22(2), 175–187. <https://doi.org/10.1002/nsg.12281>
- Zhu, X., & Liu, D. (2013). MAP-MRF approach to Landsat ETM+ SLC-Off image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2), 1131–1141.
- Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152–171.