Available online at
https://mijuitm.com.my

e-ISSN: 2735-0703

Mathematical Sciences and Informatics Journal 7(1) 2026, 146 – 156.

Mathematical Sciences and Informatics Journal

Efficient Big Data Compression for Distributed Acoustic Sensing via Gradient Quantization

Shoaib UL Hassan[1*], Noureen Talpur[1], Shuhaida Mohamed Shuhidan[1], Mohd Hilmi Hasan[1], Shahneela Pitafi[2], Kazim Raza Talpur[3]

[1]Department of Computing at Universiti Teknologi PETRONAS, Perak, Malaysia.
[2]Faculty of Data Science and Information Technology, INTI International University, Negeri Sembilan, Malaysia.
[3]Faculty of Artificial Intelligence, Universiti Teknologi Malaysia (UTM), Kuala Lumpur, Malaysia.

ARTICLE INFO

Article history:
Received 31 January 2026
Revised 3 April 2026
Accepted 14 April 2026
Online first
Published 30 April 2026

Keywords:
Big Data Compression
DAS Data Compression
Data Compression
Gradient Quantization

DOI:
10.24191/mij.v7i1.11942

ABSTRACT

Distributed Acoustic Sensing (DAS) data are vast, reaching terabytes, and can generate about 650 GB to 20 TB of data daily. To manage these large data volumes efficiently and ensure rapid network transfer and long-term storage, robust and reliable data processing techniques are essential. DAS enables real-time or near-real-time monitoring. Existing techniques are time-consuming, ineffective, and inefficient. In this study, we propose an enhanced version of the Gradient Quantization (GQ) model, focusing on improving the quantization steps during compression while preserving the real signals of DAS events. Additionally, we compared the results with six widely used data compression techniques. According to the results, our proposed method achieved higher compression and a higher space-saving ratio, outperforming other models while preserving the key features of DAS signals recorded by the Interrogator Unit (IU). In contrast, existing techniques often result in the loss of important event signals when higher compression is applied, thereby reducing data quality and usability, and limiting the precision of DAS in seismic data exploration. The proposed framework demonstrated superior performance, particularly in terms of data compression. Our approach reduced DAS data by a factor of 3.92 with a 74.46% space-saving ratio and achieved a Peak Signal-to-Noise Ratio (PSNR) of 65.81 dB, indicating high-quality signal preservation after compression.

[1*] Corresponding author. *E-mail address*: shoaib_24004745@utp.edu.my

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

1. INTRODUCTION

In recent years, DAS technology has become popular in many fields. It is based on phase-sensitive optical time-domain reflectometry (Φ-OTDR) technology. DAS utilizes optical fiber cables as sensing sources and records strain rates, also known as fingerprints (Otchere et al., 2024). It introduces a highly sensitive recording mechanism using a device called an IU. It records vibrational signal frequencies from various sources due to internal or external disturbances. For instance, these disturbances may occur when devices become faulty, because of anthropogenic activities, or because of poor installation, such as weak coupling conditions. Several studies reveal that DAS routinely reaches terabytes of data and generates ~650 GB to 20 TB of data per day (Arief et al., 2021; Shao et al., 2025). Optical fiber cables used in DAS technology act as continuous sensors along their entire length. These cables can be installed on the surface and in the subsurface. The reflected waves are measured using an IU. Any imperfection that occurs during transmission causes the waves to return to the device, regardless of the distance at which the imperfection occurred. This phenomenon is known as Rayleigh backscatter. Our investigation highlights the need for attention to the rapid generation of massive data for long-term monitoring. If DAS data are not handled appropriately, they later become costly and require more storage, efficiency, and computational power for analysis.

Thus, in this research, we propose a robust model for data compression (DC) of DAS. This technique can also be used for multiple purposes without losing actual data. Our model performed well and gave the highest confidence in the DC ratio.

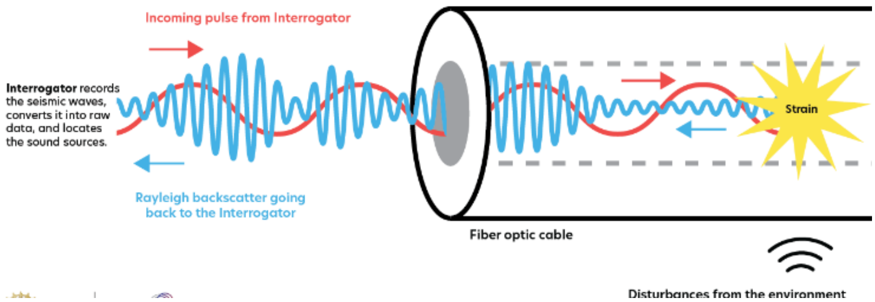

Fig. 1. DAS working model (EarthScope Consortium. 2025)

Fig. 1 illustrates the working principle of DAS based on the use of a fiber optic cable. DC is a primary problem in the continuous increase of data, in which our primary goal is to shrink the size of data for better storage and transmission. Massive data generation by DAS creates a major problem. To address this, it is critical to efficiently handle DAS data by introducing efficient storage techniques (Al-okaily & Tbakhi, 2024). The goal of DC is not only to facilitate data exploration but also to save storage space and time (Gao & Jiang, 2025). This requires a robust method to encode the information, often deleting redundant data. Algorithms always make trade-offs among DC ratio, speed, and efficiency (Chen & Sun, 2025). Efficiency often involves constraints on processing time (Carpentieri, 2025). DAS is considered big data primarily due to the enormous volume, high velocity, and diverse variety of the data it generates. These characteristics contribute to its complexity and highlight the need for robust processing and analysis techniques. To deal with massive data, advanced techniques, particularly Machine Learning (ML) and Deep Learning (DL), are increasingly integrated with DAS systems to handle the data. DAS fiber cables are typically laid over tens of kilometers, transforming them into an array of sensors. DAS provides continuous, real-time monitoring with high temporal resolution.

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

The DAS interrogator generates massive amounts of data over long distances, and the entire fiber optic cable is utilized as a sensing source. For instance, one experiment produced 128 TB of data by sampling 12,000 channels at 500 Hz over three months. (Arief et al., 2021). The enormous data size and complexity are due to diverse types of noise. The sheer volume of data creates challenges in transferring and storing data over long periods. When such voluminous data are collected, traditional methods such as Short-Term Average/Long-Term Average (STA/LTA) become less effective. Additionally, variable SNR and low-magnitude events often pose a problem due to the data's unprecedented size and velocity, coupled with challenges in data quality and SNR (Dong et al., 2020; Luo et al., 2023; Yu, F.-H., et al., 2024). Computational processing requires advanced data-management and analytical tools to extract meaningful information from the complex and noisy signals (Ashry et al., 2022). Thus, the above metrics are important to maintain and preserve key features of DAS signals.

This research aims to propose a technique for DAS data compression using machine learning techniques. However, since this technology generates several terabytes of data over long periods, robust and efficient compression algorithms are essential. In this research, we introduce an extended version of GQ, incorporating our proposed technique for compression and decompression of big data, such as DAS. The proposed technique is compared with six existing DC techniques including wavelet, down-sampling, autoencoder, Huffman, Lempel-Ziv-Welch (LZW), and Lempel-Ziv-Storer-Szymanski (LZSS). In the results and discussion section, the comparison outcomes are presented. As shown in Figure 7, two evaluation metrics, compression ratio and space-saving (%), were used. Our proposed GQ technique outperformed all six existing methods. Furthermore, Table 1 presents the performance of GQ across various quality metrics.

2. RELATED WORK

Researchers' investigations have revealed that data storage is a problem, as data routinely reaches terabytes. Additionally, it is costly to handle and store data for long periods. To address this problem, many researchers have tried to solve it and have proposed frameworks for data compression. A recent study by Isaenkov et al. (2021) implemented a model that reduces the ~1.5 TB of data to ~500- MB. Fig. 2 illustrates the compression technique followed by the authors in the study.

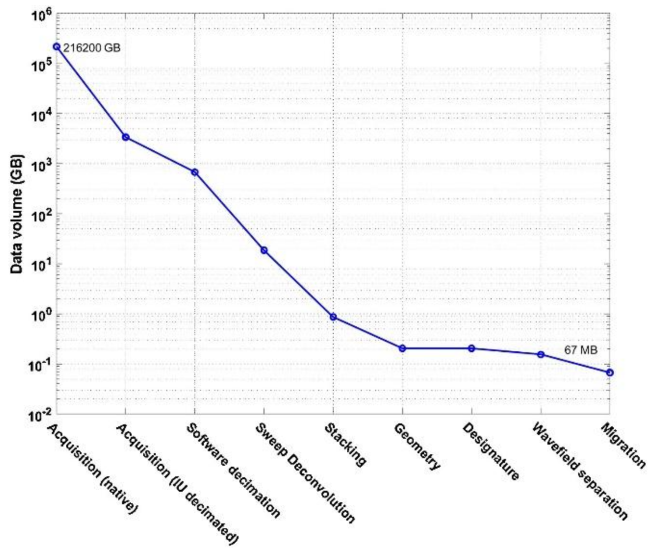

Fig. 2. Data reduction of DAS. Given volume reduced from 216 TB to 67 MB (Isaenkov et al., 2021)

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

They acquired Vertical Seismic Profile (VSP) data from the Carbon Capture and Storage (CCS) project, where a DAS monitoring system was installed at the CO_2CRC Otway site in Australia in early 2020. In another study, Yu, P.-L., et al. (2024) proposed a DC solution for DAS data. The method not only outperformed existing techniques but also offered two key advantages: demodulation of compressed data and phase information, enabling concurrent acquisition of amplitudes while preserving a high density of channels (sensors). This provides an affordable way to store acoustic signals efficiently. The model in the study outperformed existing approaches and achieved a 19.53 compression ratio and a 94.88% space-saving ratio. However, in DAS data, higher degradation of signals is considered a failure of the compression methodology.

Yu et al. (2022) proposed a DC framework to resolve the data storage problem. They achieved a 98.75 percent space-saving ratio by converting 128 MB of data to 1.6 MB. Their results also revealed that vibrational signal frequencies could be accurately reconstructed from under-sampled 1-bit data, but the SNR was not satisfactory. These parameters are critical quality metrics for DAS data. However, our proposed GQ outperformed the existing approaches in terms of compression rate and space saving while preserving signal quality. Our proposed model yielded better performance in the results and discussion section.

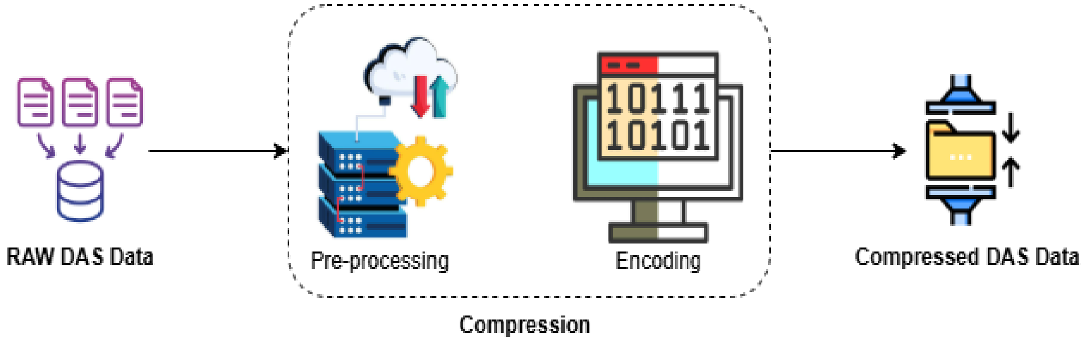

Fig. 3. Compression strategy followed by Dong et al. (2022)

Another study by Dong et al. (2022) on DAS data was conducted to compress the data in real-time or near-real-time for monitoring applications as shown in Fig. 3. They used two techniques, DASSA and H5TurboPF for DC, and managed to reduce 40% of the data. Wang et al. (2025) proposed a scheme for DC and reduced the data size from 9.31 GB to 238.42 MB, achieving a ≈ 39.95 compression ratio. They acquired data from a 6-km-long sensing optical cable and achieved a 40-fold reduction. Further analysis proved that the strategy could successfully reconstruct the data. Shen et al. (2022) proposed a novel classification technique and compressed DAS data. Their novel scheme compressed raw data and classified different acoustic signals. In the experiment, their model achieved 90% accuracy. It also handled four-class classification using 30% of the actual data. The authors also observed that the data compressed using their scheme demonstrated better performance, achieving a higher SNR and measurement ratio.

The literature related to DC shows that robust strategies and frameworks are essential, particularly for large-scale DAS monitoring projects. In such systems, optical fiber cables serve as sensing sources, generating enormous volumes of data that can reach terabytes per day. This study attempts to address this problem in DAS. In recent years related work has focused mainly on other types of data rather than DAS data. Additionally, this study offers a better understanding of DAS compression and enhances compression at a higher rate thereby standing out among prior frameworks.

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

3. DAS DATA COMPRESSION

In the literature review, numerous studies have been conducted on DAS data compression. Our investigation found that only a few studies have not only proposed efficient frameworks but also provided solutions that offered additional advantages for DAS data after compression. They reduced the size and outperformed existing methods. Related studies have shown how important it is to use effective DC techniques when managing the enormous amounts of DAS data. Many applications, such as pipeline leakage detection, road and marine traffic, enhanced oil and gas exploration, structural monitoring, and seismic detection, require efficient DAS data handling. Classification features have even been incorporated into the compression pipelines of some methods. Despite these developments, most current solutions have primarily addressed data reduction and compression, with little direct relevance to long-term DAS monitoring systems. As a result, this study closes a significant gap by addressing compression issues unique to DAS, suggesting a more efficient solution suited to the special characteristics of DAS data, and making a significant contribution to scalable and affordable DAS data management.

4. PROPOSED MODEL

In this study, the proposed data compression model utilized Gradient Quantization (GQ) for DAS data. However, the compression techniques followed in this study are extended and given a high compression rate. In this study, we achieved 92.56% data compression with different quantization steps. The model achieved high compression and successfully reconstructed DAS files in their original form with no loss of data during decompression actions. Fig. 4 shows the GQ-based compression and decompression framework. The framework begins with loading SEGY file data, followed by three main compression steps: gradient compression, quantization, and entropy coding. For decompression, the reverse operations are entropy decoding, dequantization, and gradient decompression. In the last steps, the original DAS signal is reconstructed. This model enables better storage while preserving essential signal characteristics without losing desired signals.

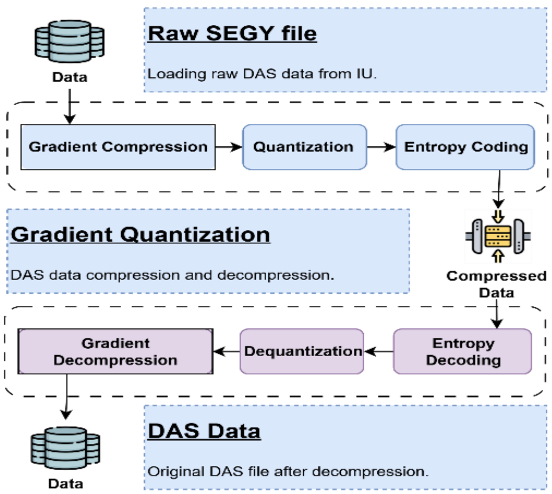

Fig. 4. Gradient Quantization for DAS

In this study, we proposed an efficient, lossy compression framework tailored for high-dimensional DAS data stored in SEGY file format, followed by industry applications. Based on Fig. 4, the proposed

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

framework integrates signal transformation, a quantization technique, and entropy coding to achieve significant space savings while preserving essential waveform characteristics for downstream interpretation. The complete framework flow is based on the following steps:

(i) DAS Data: The frameworks begin by loading DAS files containing traces directly from SEGY files using ObsPy, a Python library for seismological data. It is a publicly available library and is widely used. Each trace is extracted and converted into a 2D NumPy array. Fig. 6 shows the single DAS trace, which contains the strain-rate recorded by the Interrogator Unit (IU).

(ii) Gradient-Based Compression: To exploit temporal redundancy inherent in DAS signals, we apply first-order differencing along the time axis. This operation emphasizes dynamic changes in the signal, making it more amenable to compression by reducing low-frequency redundancy. The initial column of the original signal is stored separately to enable full reconstruction.

(iii) Scalar Quantization: The difference signals are then quantized using a uniform scalar quantizer with a fixed step size (q_step). This reduces the numerical precision of the data while enabling effective bitstream compression. The choice of q_step allows for a tunable trade-off between fidelity and compression rate.

(iv) Entropy Coding: In this step of the compression framework, we serialized the data and subjected the quantized gradient data to entropy coding using zlib, which is a lossless compression technique and a dictionary-based compression scheme. This step further reduces the data volumes. This reduction is beneficial because DAS systems generate enormous data volumes by eliminating statistical redundancies in the byte representation of the quantized array.

(v) Decompressed File: Three components are stored during compression: (i) the entropy-encoded data stream, (ii) the shape of the quantized data, and (iii) the initial signal column required for reconstruction. These files are used to recreate the original signal during decompression.

(vi) Decompression: The decompression pipeline reverses each step in sequence. First, the compressed binary is decoded using zlib. The resulting quantized gradients are then dequantized to approximate the original gradient values. Using the stored initial column, the original signal is reconstructed by cumulative summation.

Fig. 5 illustrates the workflow of the proposed GQ technique along with other reconstructible data compression methods. The process begins with data input and preprocessing, followed by the application of different compression techniques, including gradient quantization, wavelet transform, autoencoder, and downsampling. Each of these methods aims to efficiently reduce data size while maintaining the integrity of DAS signals. However, during implementation, we observed that increasing the quantization steps significantly improves the compression rate compared to other techniques but simultaneously reduces signal clarity.

The compressed data are then subjected to evaluation and analysis using various compression metrics to assess performance in terms of compression efficiency, signal preservation, and reconstruction quality. This workflow provides a clear overview of the proposed approach and its comparative assessment against other conventional compression methods.

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

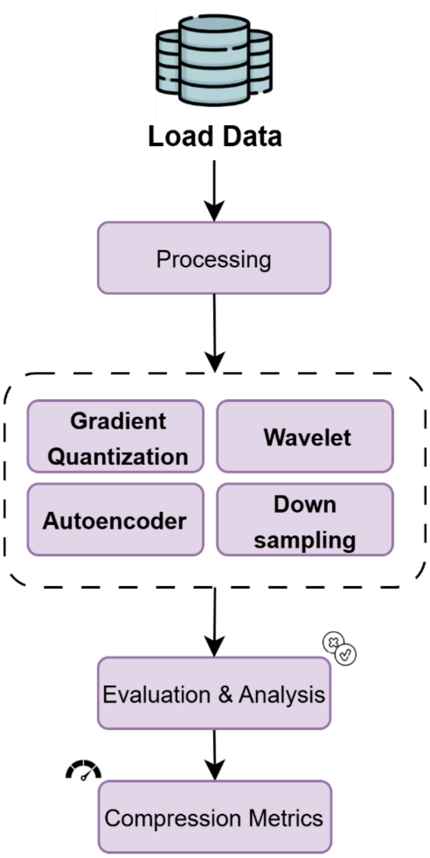

Fig. 5. Gradient Quantization model flow with reconstructible methods

5. RESULTS AND DISCUSSION

The experimental results illustrated in Fig. 6 show the comparative file sizes before and after applying GQ to a DAS SEGY dataset. The left bar represents the size of the original uncompressed data, which is approximately 15.41 MB. The right bar represents the total size of the compressed output, including the entropy-coded file, shape metadata, and the initial signal column required for reconstruction. The compressed size is approximately 3.94 MB, yielding a substantial space reduction. This corresponds to a compression ratio of approximately 3.92× and an overall space saving of approximately 74.46% (as shown in Fig. 7), demonstrating the effectiveness of the gradient-based quantization and entropy encoding approach in minimizing storage without retaining the entire uncompressed dataset. Such a reduction is particularly valuable in large-scale DAS deployments where data volume can quickly become a limiting factor for storage and transmission. During the data compression experiments, we compared and evaluated six techniques: wavelet, downsampling, autoencoder, Huffman, LZW, and LZSS and autoencoder, against the proposed GQ method. The experimental results showed that GQ achieved a higher compression rate of over 90%, demonstrating significant improvement over previous work.

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

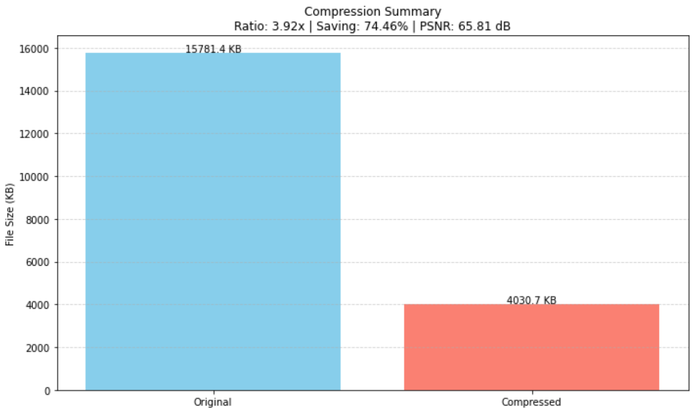

Fig. 6. Compression result comparison with original DAS data file

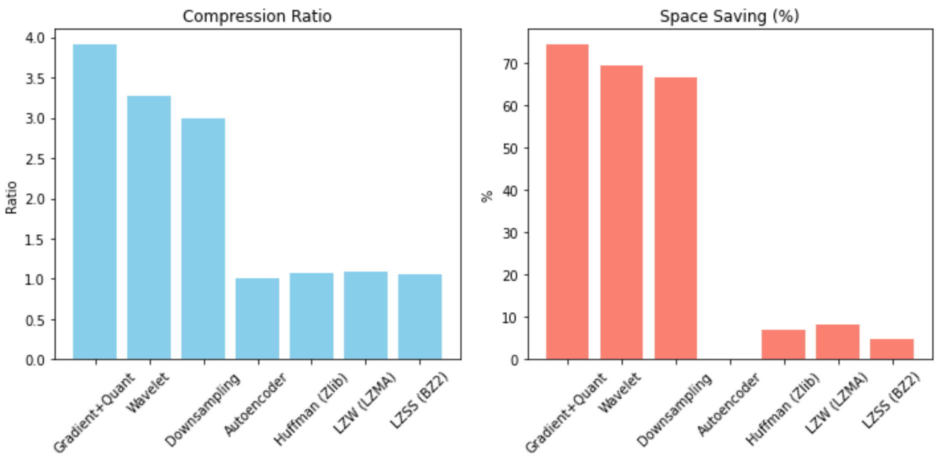

Fig. 7. Comparison of GQ with six compression techniques based on two metrics: compression ratio, space saving

Table 1. Data Quality Metrics for DAS with Existing DC Techniques and GQ Robustness

Metric	GQ	Wavelet	Down-sampling
SNR (dB)	26.49	24.85	4.47
PSNR (dB)	59.55	57.91	37.54
Spectral Entropy	3.3449	3.3458	2.8608
Kurtosis	384.98	398.82	371.74
Skewness	0.0174	-0.0082	0.1262
Zero-Crossing Rate	0.1813	0.1866	0.1030

Table 1 highlights the robustness and effectiveness of the GQ method for DAS data, particularly in achieving superior compression and space-saving performance. While applying compression techniques, we ensured that the data quality remained consistent.

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

Table 2. Signal Quality Metrics for Original and Reconstructible Compression Methods

Method	SNR (dB)	PSNR (dB)	Spectral Entropy	Kurtosis	Skewness	Zero-Crossing Rate
Original (Uncompressed)	∞	∞	3.3447	386.7182	-0.0073	0.1889
GradientQuant	26.4854	59.5477	3.3449	384.9844	0.0174	0.1813
Wavelet	24.8501	57.9123	3.3458	398.8152	-0.0082	0.1866
Downsampling	4.4749	37.5371	2.8608	371.7360	0.1262	0.1030

Furthermore, Table 2 compares GQ with other reconstructible compression techniques, where GQ exhibited the most robust performance, achieving higher SNR (26.49 dB) and PSNR (59.55 dB) values—indicating strong signal preservation and minimal quality loss, which is crucial for DAS data. The GQ method also maintained spectral entropy and zero-crossing rate values closely aligned with the original uncompressed data, confirming its ability to preserve both temporal and frequency characteristics without degradation. Additionally, when compared with the wavelet and downsampling techniques, GQ achieved superior compression and space-saving ratios while maintaining signal fidelity, making it highly suitable for DAS signal preservation.

6. CONCLUSION AND FUTURE WORK

DAS produces massive data volumes (TB/day). An efficient DAS DC framework can significantly reduce the cost of storage, management, and data transfer. Our proposed model outperforms others and shows a high compression ratio and fast compression. This framework compresses the raw DAS data acquired by the IU. The proposed framework can be more impactful for data compression if the quantization steps are carefully treated. Furthermore, the focus should not only be on compression and the space-saving ratio, but also on preserving key signal features of DAS data during compression. These steps play a greater role in the effectiveness of techniques such as GQ. In big data applications such as DAS, robust data compression techniques are needed to address problems such as data storage. For example, this technology produces a high volume of data, which requires more space for long-term storage. The evaluation report in the above results and discussion section shows the performance of various techniques. Here, we considered four metrics for reconstructible techniques: SNR, PSNR, spectral entropy, and zero-crossing rate, and confirmed the robustness and reconstructability of the GQ approach.

7. ACKNOWLEDGEMENTS/FUNDING

The authors would like to acknowledge the support of the Center for Research in Data Science, Department of Computing, Faculty of Science, Management and Computing, and Universiti Teknologi PETRONAS for providing the facilities and financial support for this research under grant number 015BC3-001.

8. CONFLICT OF INTEREST STATEMENT

The authors agree that this research was conducted in the absence of any self-benefits, commercial or financial conflicts and declare the absence of conflicting interests with the funders.

9. AUTHORS' CONTRIBUTIONS

Shoaib UL Hassan and Noureen Talpur designed the study, developed the proposed methodology, conducted the experiments, and prepared the original draft, including review and editing. Shuhaida Mohamed Shuhidan, and Mohd Hilmi Hasan contributed to validation, methodology refinement, and formal analysis. Shahneela Pitafi and Kazim Raza Talpur were involved in validation and formal analysis.

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

10. DECLARATION OF GENERATIVE AI IN THE WRITING PROCESS

The authors declare that generative artificial intelligence (AI) tools may have been used solely to support language refinement during the preparation of this manuscript. Such tools were used only to improve grammar, sentence structure, clarity, and overall readability. All aspects of the research, including the study design, data collection, data extraction, analysis, interpretation of findings, and final academic decisions, were conducted, reviewed, and approved by the authors. The authors take full responsibility for the accuracy, integrity, originality, and scholarly content of this manuscript.

11. DATA AVAILABILITY/SUPPLEMENTARY MATERIALS

The data and/or supplementary materials supporting the findings of this study are available from the corresponding author upon reasonable request, where applicable.

12. ETHICS STATEMENT

The authors declare that this study did not involve human participants or animal subjects. Where applicable, all procedures were conducted in accordance with relevant institutional guidelines, safety requirements, and ethical standards.

REFERENCES

Al-okaily, A., & Tbakhi, A. (2025). A novel lossless encoding algorithm for data compression–genomics data as an exemplar. *Frontiers in Bioinformatics, 4*, Article 1489704. https://doi.org/10.3389/fbinf.2024.1489704

Arief, H. A., Wiktorski, T., & Thomas, P. J. (2021). A survey on distributed fibre optic sensor data modelling techniques and machine learning algorithms for multiphase fluid flow estimation. *Sensors, 21*(8), Article 2801. https://doi.org/10.3390/s21082801

Ashry, I., Mao, Y., Wang, B., Hveding, F., Bukhamsin, A., Ng, T. K., & Ooi, B. S. (2022). A review of distributed fiber-optic sensing in the oil and gas industry. *Journal of Lightwave Technology, 40*(5), 1407–1431. https://doi.org/10.1109/JLT.2021.3135653

Carpentieri, B. (2025). Data compression with a time limit. *Algorithms, 18*(3), Article 135. https://doi.org/10.3390/a18030135

Chen, S., & Sun, T. (2025). Research on data compression and efficient transmission technology in the framework of big data processing. *Applied Mathematics and Nonlinear Sciences, 10*(1), 1–16. https://doi.org/10.2478/amns-2025-0339

Dong, B., Popescu, A., Tribaldos, V. R., Byna, S., Ajo-Franklin, J., & Wu, K. (2022). Real-time and post-hoc compression for data from distributed acoustic sensing. *Computers & Geosciences, 166*, Article 105181. https://doi.org/10.1016/j.cageo.2022.105181

Dong, B., Tribaldos, V. R., Xing, X., Byna, S., Ajo-Franklin, J., & Wu, K. (2020). DASSA: Parallel DAS data storage and analysis for subsurface event detection. In *2020 IEEE 34th International Parallel and Distributed Processing Symposium (IPDPS)* (pp. 254–263). IEEE. https://doi.org/10.1109/IPDPS47924.2020.00035

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026

EarthScope Consortium. (2025). *Distributed acoustic sensing*. https://www.earthscope.org/what-is/das/

Gao, J., & Jiang, Y. (2025). *Efficient lossless compression strategy using data distribution further*. Research Square. https://doi.org/10.21203/rs.3.rs-6343043/v1

Isaenkov, R., Pevzner, R., Glubokovskikh, S., Yavuz, S., Yurikov, A., Tertyshnikov, K., Gurevich, B., Correa, J., Wood, T., Freifeld, B., Mondanos, M., Nikolov, S., & Barraclough, P. (2021). An automated system for continuous monitoring of CO_2 geosequestration using multi-well offset VSP with permanent seismic sources and receivers: Stage 3 of the CO2CRC Otway Project. *International Journal of Greenhouse Gas Control, 108*, Article 103317. https://doi.org/10.1016/j.ijggc.2021.103317

Luo, L., Wang, W., Yu, H., Chen, X., & Bao, S. (2023). Abnormal event monitoring of underground pipelines using a distributed fiber-optic vibration sensing system. *Measurement, 221*, Article 113488. https://doi.org/10.1016/j.measurement.2023.113488

Otchere, D. A., Latiff, A. H. A., & Tackie-Otoo, B. N. (2024). Distributed acoustic sensing in subsurface applications: Review and potential integration with artificial intelligence for an intelligent CO_2 storage monitoring system. *Geoenergy Science and Engineering, 237*, Article 212818. https://doi.org/10.1016/j.geoen.2024.212818

Shao, L., Zhang, J., Chen, X., Xu, D., Gu, H., Mu, Q., Yu, F., Liu, S., Shi, X., Sun, J., Huang, Z., Yang, X., Zhang, H., Ma, Y., Lu, H., Liu, C., & Yu, C. (2025). Artificial intelligence-driven distributed acoustic sensing technology and engineering application. *PhotoniX, 6*(1), Article 4. https://doi.org/10.1186/s43074-025-00160-z

Shen, X., Li, J., Wu, Z., Dang, H., Chen, J., Shao, L., Liu, H., Shum, P. P., Wu, H., Zhu, K., Li, Y., Zheng, H., & Lu, C. (2022). Compressed-domain data classification for distributed acoustic sensing system. In *2022 Asia Communications and Photonics Conference (ACP) and International Conference on Information Photonics and Optical Communications (IPOC)* (pp. 108–110). Optica Publishing Group. https://doi.org/10.1109/ACP55869.2022.10088818

Wang, Y., Jiang, J., Liu, K., Zhang, M., Wang, S., Xu, T., Zhang, X., Ding, Z., & Liu, T. (2025). Ultra compression-ratio distributed acoustic sensing system via combination of envelope extraction and compressed sensing. *Optics Express, 33*(5), 9854. https://doi.org/10.1364/OE.554324

Yu, F., Liu, S., Xu, D., Gu, H., Chen, X., Zhang, J., Yang, X., Xiao, D., Zhang, H., & Shao, L. (2024). Data compression solution for φ-OTDR: Encoding the extracted spectrum. *IEEE Sensors Journal, 24*(14), 22429–22438. https://doi.org/10.1109/JSEN.2024.3407752

Yu, F., Shao, L., Liu, S., Xu, W., Xiao, D., Liu, H., & Shum, P. P. (2022). Data reduction in phase-sensitive OTDR with ultra-low sampling resolution and undersampling techniques. *Sensors, 22*(17), Article 6386. https://doi.org/10.3390/s22176386

Yu, P., Zhu, T., Marone, C., Elsworth, D., & Yu, M. (2024). DASEventNet: AI-based microseismic detection on distributed acoustic sensing data from the Utah FORGE well 16A (78)-32 hydraulic stimulation. *Journal of Geophysical Research: Solid Earth, 129*(9), Article e2024JB029102. https://doi.org/10.1029/2024JB029102

© 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.24191/mij.v7i1.11942

©Authors, 2026