ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

Geological CO_2 Storage Safety: A Comprehensive Review of Leakage Risks, Assessment Methods, and Mitigation Strategies

Hao Xue[1, a]

[1]University of Glasgow, Glasgow, G12 8QQ, UK

[a]Alex-X05@outlook.com

Abstract

Anthropogenic CO_2 emissions elevate atmospheric carbon levels and induce severe global climate risks, as documented in IPCC assessment reports. Carbon capture, utilization and storage (CCUS) is critical for climate mitigation, yet large-scale geological CO_2 storage deployment is hindered by incomplete understanding of time-dependent trapping behaviours and non-standardized risk-assessment protocols. This review aims to synthesize recent advances in geological carbon storage and deliver practical technical references for site-related engineering practice. Four typical geological storage formations are reviewed, and the temporal evolution of multiple CO_2 trapping mechanisms is analysed to interpret long-term sequestration stability. Representative risk-assessment workflows, including the SPR framework, probabilistic leakage evaluation, numerical simulation and uncertainty analysis are systematically compiled. Results reveal that structural, residual, solubility and mineral trapping exhibit time-varying relative contributions governing reservoir-specific storage security; coupled deterministic-probabilistic risk workflows mitigate decision biases during site screening. This review provides actionable guidance for site screening, monitoring deployment and regulatory compliance, supporting safe large-scale implementation of geological CO_2 storage within CCUS systems.

Keywords

Geological CO_2 storage; trapping mechanisms; risk assessment; Source–Pathway–Receptor framework.

1. INTRODUCTION

CO_2 emissions from human activity are at an all-time high, and geological carbon storage is now essential for reaching net-zero goals. Deep saline aquifers and depleted oil and gas fields alone can permanently store billions of tonnes of carbon [1,2,3]. Current commercial projects have demonstrated technical feasibility, yet long-term containment security remains insufficiently validated over centennial timescales, quantitative risk assessment methods lack standardized protocols, and the vast number of legacy wells presents a formidable unknown risk factor [2,3,4]. This study systematically synthesizes the complete risk assessment workflow—from geological storage carriers and time-dependent trapping mechanisms, through the Source–Pathway–Receptor framework and probabilistic leakage evaluation, to numerical simulation techniques and uncertainty-based risk ranking—providing an accessible technical reference for safe and sustainable CO_2 geological storage [4].

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

2. STORAGE MECHANISMS AND TRAPPING FORMS

Four major geologic unit types are viable for subsurface CO_2 geological storage. Deep saline aquifers hold the highest theoretical storage volume among these options, consisting of strata filled with high-salinity formation brine [4]. Depleted oil-gas reservoirs have well-documented geological characteristics, and existing wellbores can be repurposed for injection operations [5]. In unmineable coal seams, CO_2 is retained through adsorption processes on coal-matrix surfaces [6]. Basalt formations react rapidly with injected CO_2; carbon can be locked away as carbonate minerals via mineral carbonation within just a few years [7].

Supercritical CO_2 is introduced into underground strata to minimise subsurface storage volume [1]. Once emplaced underground, CO_2 becomes trapped through four time-dependent trapping processes [5]. Right after injection, caprock and adjacent stratigraphic layers act as physical barriers, stopping buoyant CO_2 from migrating upward. Over the following years, capillary forces hold discrete CO_2 phases within rock micropores and pore throats, limiting fluid movement. Across decades to centuries, CO_2 slowly dissolves in formation brine. Higher fluid density makes the CO_2-charged brine sink deeper into the subsurface. After hundreds to thousands of years, geochemical reactions take place between CO_2 and host rock minerals, forming stable solid carbonate minerals — this is mineral trapping. It provides the most secure, long-term sequestration outcome; carbon fixed in this form will not be remobilised even under seismic disturbance [7,8].

Leakage risk for subsurface-stored CO_2 changes with storage time and generally declines over the project lifecycle [8,9]. Shortly after injection commences, mobile CO_2 plumes are still present underground, creating the greatest chance of fugitive leakage. As CO_2 dissolves and transforms into mineral phases, leakage hazards fall to negligible levels. For this reason, stringent, high-frequency monitoring is critical in early-stage injection. Monitoring intensity may then be progressively reduced as storage time proceeds [8,9].

3. RISK ASSESSMENT METHODS FOR GEOLOGICAL STORAGE

Geological storage projects involve inherent uncertainties in reservoir properties, caprock integrity, and wellbore conditions. A robust risk assessment framework transforms these uncertainties into actionable information for site selection, injection design, monitoring optimization, and regulatory compliance [5]. Risk management for CO_2 leakage encompasses three sequential steps: first, assessing the likelihood and potential impact of leakage; second, implementing measures to mitigate and avoid unacceptable risks; third, developing contingency plans for risks that remain tolerable but cannot be fully eliminated [1]. Following this logic, this section systematically reviews the established Source–Pathway–Receptor (SPR) framework for risk identification, probabilistic methods for leakage estimation, consequence analysis and receptor vulnerability assessment, numerical simulation tools, and uncertainty-based risk ranking approaches that guide monitoring resource allocation.

3.1. Risk Assessment Framework: Source–Pathway–Receptor

The Source-Pathway-Receptor (SPR) framework offers a structured approach to recognise and quantify risks for geological CO_2 storage projects [9]. Under this analytical logic, one single leakage source may link to multiple migration pathways, and each pathway can affect several receptors. This forms a risk matrix that demands thorough evaluation. To carry out systematic risk identification, the subsurface storage system is split into three functional parts: Origin, Pathway and Endpoint [5,6].

Such system decomposition supports the orderly identification of Features, Events, and Processes (FEPs) that may trigger CO_2 leakage. Originated from performance-assessment

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

practice for nuclear waste repositories, the FEP approach has become a fundamental tool for CO_2 storage risk analysis [3,4]. Features stand for static geological and engineering attributes, including lithology, porosity, permeability, caprock thickness, faults and pre-existing wells. Events refer to abrupt safety perturbations: seismic activities, adjacent drilling operations, injection pressure surges and casing failure fall within this category [2,3].

Recent improvements to FEP methods centre on quantifying dynamic risks driven by inter-component connections. The Process Influence Diagram (PID) maps causal links among different FEPs and helps construct well-organised leakage scenarios [2]. Interaction matrices take FEPs as diagonal entries and mark cross-item interactions, which can be tailored for site-specific risk appraisal [2]. The Quintessa FEP database stores roughly 200 hierarchical FEPs with detailed descriptions and referenced sources [8]. It has been utilised in major global storage demonstrations such as Sleipner, Weyburn, In Salah and Decatur [2,3,4].

Survival-analysis-based surrogate models make long-term leakage forecasting feasible, removing the need for computationally expensive century-scale full-physics simulations [10,11]. Updated FEP-integrated systems take cement-CO_2 geochemical reactions, thermo-mechanical reservoir variations and corrosion of aged wellbores into account, so as to evaluate long-term integrity of abandoned wells [10,12]. These technical advances enable three-phase risk modelling: building the site conceptual model, mathematically defining key leakage scenarios, and quantitatively evaluating potential consequences [11]. A standardised performance-assessment protocol also delivers unified workflows combining FEP screening, scenario filtering and quantitative risk computation for geological carbon storage [4].

3.2. Leakage Probability Assessment

When field measurements are insufficient, qualitative risk assessment heavily depends on expert knowledge. Analysts refer to historical oil-gas leakage cases to assign probability values for different failure scenarios [5]. Fault Tree Analysis (FTA) and Event Tree Analysis (ETA) serve as core quantitative risk instruments [3,5,6]. FTA breaks down top-level leakage hazards into elementary component failures using Boolean logic rules [5]. ETA tracks how initial disturbances propagate, judging success or failure of each safety barrier along the evolution path [6]. Both techniques can couple with Bayesian Networks to achieve more refined risk computation [12-14].

The SHARP research project combined event-tree logic and Monte-Carlo simulation to evaluate fault- and seismic-related leakage risks for North Sea storage sites [12]. Jing et al. (2024) constructed a fault-tree Bayesian network fed with fuzzy expert probabilities. For a representative storage site, the aggregated leakage probability was calculated as 29 %; fault fracture leakage contributed 32 %, caprock breakthrough 28 %, and well-related leakage 19 %, placing the site within medium-low risk rank [13]. As open-source software, NRAP-Open-IAM separates reservoir simulation and leakage calculation modules and supports modelling of time-changing leakage hazards [14]. Moopen et al. (2024) adopted this tool to compute multi-segment well-leakage probabilities, complying with California's criteria for permanent carbon-storage certification [14,15].

3.3. Leakage Consequence Assessment and Receptor Vulnerability

Consequence assessment quantifies key indicators including total CO_2 leakage mass, peak contaminant concentration, arrival time of migrating plumes, and spatial extent of impacted zones [5]. Groundwater vulnerability is heavily conditioned by aquifer mineral composition. Carbonate-rich aquifers possess strong pH-buffering capacity, whereas sandstone formations

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

are prone to acidification. Even small amounts of intruding CO_2 may trigger mobilisation of heavy-metal elements within sandstone aquifers. Risks to local ecosystems and human populations are tied to biological species richness and how close residents live relative to potential leakage zones [4].

The two-dimensional Risk Assessment Matrix (RAM) offers visualisation for both leakage probability and the magnitude of associated harm [1]. Quantitative RAM scoring can draw on either mean values or worst-case scenario outputs [8]. Modified RAM workflows now incorporate FEP outputs and Bayesian event-tree analysis. This combination yields complete sets of mutually-exclusive leakage scenarios alongside tiered risk acceptance thresholds [9,15]. At Canada's Quest CCS site, bow-tie risk analysis first distinguished four primary hazard groups. Researchers then ran NRAP-Open-IAM simulations to confirm low groundwater contamination potential. Based on these results, the formal monitoring area was scaled down from 460 km^2 to 100 km^2 [15]. For sites hosting large numbers of legacy wells, such integrated RAM-probabilistic approaches can substantially lower overall monitoring burdens [15].

3.4. Numerical Simulation in Risk Assessment

Numerical simulation converts measured geological attributes into quantifiable forecasts for CO_2 migration, geochemical alteration and geomechanical responses [5,6]. The international Sim-SEQ model inter-comparison exercise deployed TOUGH2, STOMP, Eclipse and VESA for the S-3 test case. Discrepancies rooted in different conceptual model set-ups produced considerable output divergence: predicted breakthrough time spanned 8–36 days, while post-injection gas saturation ranged from 79 % to 95 % [15,16,17]. All simulated results therefore require calibration against real-world field monitoring data [16,17].

Multiphase flow simulators are routinely adopted for regulatory submissions to model long-range CO_2 plume movement [6]. Codes including TOUGH2, STOMP and Eclipse can represent complex fault structures and heterogeneous grid domains [18,19]. The STOMP-WCSE injection model reproduced observed injection rates and bottom-hole pressure iteration behaviour at the S-3 site [19]. High-resolution multiphase modelling work for the Cranfield pilot further demonstrated that sedimentary facies and permeability exert dominant control over supercritical CO_2 saturation patterns [16,19].

Reactive-transport modelling couples multiphase flow computation with water-rock geochemical reactions [18]. PFLOTRAN simulations for the Cranfield site showed dissolved CO_2 spread across a 3.5 km^2 area surrounding injection wells; pH values dropped to 3.3–5.5 near plume fringes [18]. Fully coupled modelling also identified river-channel-related geochemical hotspots as critical controls governing CO_2 transport behaviour [16,18]. Over a simulated 20-year period, geochemical reactions halved the areal extent of the free-phase CO_2 plume and enhanced dissolved-trapping performance [18].

Geomechanical simulation evaluates reservoir stress evolution, ground surface uplift, and fault re-activation risk [19,20]. TOUGH-FLAC provides coupled computation for subsurface flow and rock mechanical deformation [20]. Five-year CO_2 injection operations at In Salah (Algeria) elevated reservoir pore pressure from 18 MPa up to 30 MPa and induced roughly 2 cm of surface uplift [19]. Thirty-year modelling for the Jurong depleted oil reservoir indicated single-well injection of 5.43 million tonnes CO_2 could drive formation-pressure increases as high as 9.5 MPa [19,21]. For Brazil's Parnaíba Basin, a 3-D Mohr-Coulomb geomechanical model was applied to define safe upper limits for injection pressure, so that storage resource utilisation could be maximised [19,20].

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

3.5. Uncertainty Analysis, Sensitivity Assessment, and Risk Ranking

Reservoir-scale heterogeneity introduces uncertainties for key input parameters such as porosity, permeability, fault friction coefficients and geochemical reaction constants [5,6]. Complete statistical distributions for these parameters are rarely available from field-scale real-world storage projects [21,22]. Uncertainty-quantification (UQ) workflows adopt Monte-Carlo, polynomial-chaos and Bayesian approaches to propagate input-parameter errors and derive probability distributions for model outputs [22,23]. The arbitrary polynomial-chaos (aPC) algorithm performs well under data-sparse conditions and does not require fully defined parameter distribution functions [22,23,25]. Köppel carried out benchmark comparisons across four UQ algorithms and verified that aPC can yield stable outputs with substantially reduced simulation counts [25]. Machine-learning-based surrogate models can speed up uncertainty estimation for saline-aquifer storage capacity, with Latin-Hypercube sampling and Bayesian optimisation as common supporting techniques [24]. Polynomial response-surface methods serve as an alternative to computationally heavy full-physics simulations and enable fast probabilistic assessment [26].

Global-sensitivity analysis approaches including Sobol indices and Morris screening quantify how much each individual parameter contributes to total output variance [24,26]. One TOUGH2-based modelling exercise for California's San Joaquin Basin singled out permeability as the governing factor: higher permeability mitigates pressure build-up yet broadens the spatial extent of the CO_2 plume [26,27]. Simulations for the Farnsworth Unit demonstrated that net-to-gross ratio and permeability multipliers exert primary control over reservoir connectivity and pressure dissipation behaviour [25,27,28]. Sobol-based sensitivity work targeting oil-water bearing formations in Northern Jiangsu, China further indicated permeability dominates fluid production responses, whereas porosity mostly dictates achievable storage capacity [28,29].

Risk-ranking schemes integrate leakage probability and potential damage severity into standardised tier classifications (high / medium / low). Such tiering supports rational allocation of monitoring resources across candidate sites [8,9]. Random-forest machine-learning models (F1-score = 0.986) trained on 300 stochastic simulation realisations have been applied to delineate high-risk geographical zones. Researchers also distinguished alternative development plans by interpreting P10/P50/P90 confidence interval bounds from model outputs [27,28].

4. CONCLUSION

This paper systematically sorts out the basic storage principles and complete risk assessment technical system of geological CO_2 sequestration. Four reservoir types and time-evolving trapping mechanisms determine long-term storage safety. SPR identification, probabilistic calculation, numerical modeling and risk grading jointly quantify leakage hazards and receptor vulnerability. Geological parameter uncertainty and differing parameter sensitivities greatly affect prediction accuracy. Comprehensive integrated risk workflows optimize site selection, injection schemes and monitoring layouts to minimize underground CO_2 leakage risks. However, these assessment techniques have distinct application limits. SPR supports rapid basin-scale screening but ignores local reservoir heterogeneity; probabilistic methods require abundant field data; multi-physics numerical modeling suffers from high computational cost; risk grading lacks globally accepted unified criteria.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

Future research should address key gaps: advancing digital-twin platforms for real-time reservoir monitoring, developing AI-driven risk-prediction tools, and optimizing multi-scale THMC coupled models to capture time-dependent trapping and leakage behaviours. Sufficient pilot-project datasets are also needed to constrain uncertain geological parameters and calibrate model outputs.

For engineering and regulatory practice, staged evaluation workflows are suggested: apply simplified screening for regional surveys, and implement high-fidelity simulation paired with long-term monitoring for priority sites. Injection and monitoring strategies shall be dynamically updated following the evolution of trapping mechanisms, instead of static one-time design. Regulators are advised to establish whole-life-cycle technical specifications and adopt digital-twin-enabled early-warning tools, facilitating safe large-scale deployment of geological CO_2 sequestration within CCUS chains.

REFERENCES

[1] Benson, S. M., & Surles, T. (2006). Carbon dioxide capture and storage: Status and pathways. Proceedings of the National Academy of Sciences, 103(46), 17293–17300. https://doi.org/10.1073/pnas.0607185103

[2] IPCC. (2005). Carbon dioxide capture and storage. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/03/srccs_wholereport-1.pdf

[3] CGS Europe. (2012). Knowledge report: Site selection and characterization for CO_2 storage. http://repository.cgseurope.net/eng/cgseurope/knowledge-repository/key-reports/evaluation/6/2/3.aspx

[4] CSLF Risk Assessment Task Force. (2009). Phase I final report: Risk assessment for geological storage of CO_2. https://fossil.energy.gov/archives/cslf/sites/default/files/documents/tg_PhaseIFinalReportRiskAssessmentTaskForce1009.pdf

[5] Al-Khdheeawi, E. A., & Mahdi, H. S. (2023). Deep saline aquifers for CO_2 geological storage: A review of capacity estimation and site selection. Energy Reports, 9, 2056–2072. https://doi.org/10.1016/j.egyr.2023.08.072

[6] Ringrose, P. S., & Meckel, T. A. (2023). Depleted oil and gas fields as CO_2 storage sites: Leveraging legacy infrastructure. International Journal of Greenhouse Gas Control, 128, 103920. https://doi.org/10.1016/j.ijggc.2023.103920

[7] Hassanpouryouzband, A., Joonaki, E., & Haszeldine, R. S. (2020). CO_2 storage in unmineable coal seams: Adsorption mechanisms and site screening criteria. Fuel, 279, 118478. https://doi.org/10.1016/j.fuel.2020.118478

[8] Snæbjörnsdóttir, S. Ó., Sigfússon, B., & Oelkers, E. H. (2024). Rapid mineral carbonation in basalt formations: Field evidence and reaction kinetics. Nature Geoscience, 17(2), 128–136. https://doi.org/10.1038/s41561-023-01121-9

[9] Quintessa Limited. (2004). Online database for CO_2 storage features, events & processes. https://www.quintessa.org/latest-news/online-database-for-co2-storage-features-events-and-processes

[10] Gurwicz, A., Chen, J., Gutman, D. H., & Gildin, E. (2025). Assessing risk in long-term CO_2 storage under uncertainty via survival analysis-based surrogates. SPE Journal, 1–18. https://doi.org/10.2118/222449-PA

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

[11] SHARP Project. (2024). SHARP project – an integrated approach for assessing CO_2 storage containment risks. AMiner. https://www.aminer.cn/pub/65ebd79c13fb2c6cf635c18a

[12] Jing, M., Li, Q., Liu, G., & Xue, Q. (2024). Quantitative assessment of CO_2 leakage risk in geologic carbon storage management. Greenhouse Gases: Science and Technology, 14(6), 1068–1091. https://doi.org/10.1002/ghg.2309

[13] Vasylkivska, V., Lackey, G., & Dilmore, R. (2021). NRAP-Open-IAM: A flexible open-source integrated-assessment-model for geologic carbon storage risk assessment. Environmental Modelling & Software, 143, 105114. https://doi.org/10.1016/j.envsoft.2021.105114

[14] Moopen, M., Vasylkivska, V., Dilmore, R., & Oldenburg, C. (2024). Probabilistic wellbore leakage assessment for California's geologic CO_2 storage Certificate of Permanence using NRAP-Open-IAM. International Journal of Greenhouse Gas Control, 139, 104172. https://www.sciencedirect.com/science/article/pii/S1750583624000825

[15] Brown, C. F., Lackey, G., Mitchell, N., et al. (2023). Integrating risk assessment methods for carbon storage: A case study for the Quest Carbon Capture and Storage Facility. International Journal of Greenhouse Gas Control (Revised manuscript). https://www.osti.gov/pages/servlets/purl/2274889

[16] Soltanian, M. R., et al. (2019). Multicomponent reactive transport of carbon dioxide in fluvial heterogeneous aquifers. Journal of Natural Gas Science and Engineering, 65, 212–223. https://doi.org/10.1016/j.jngse.2019.03.017

[17] Mukhopadhyay, S., Doughty, C., Bacon, D. H., et al. (2013). Preliminary model-comparison results from the Sim-SEQ project using TOUGH2, STOMP, Eclipse, and VESA approach. OSTI.GOV, PNNL-SA-89092. https://osti.gov/biblio/1110468

[18] Soltanian, M. R., et al. (2016). Simulating in-zone chemistry changes from injection time to longer periods of CO_2 storage. Environmental Earth Sciences, 75(20), 1389. https://doi.org/10.1007/s12665-016-6153-9

[19] Hosseinzadeh, S., Abdollahi, R., Salimzadeh, S., & Haghighi, M. (2024). Three-dimensional coupled temporal geomechanical model for fault-reactivation and surface-deformation evaluation during reservoir depletion and CO_2 sequestration. Sustainability, 16(19), 8482. https://doi.org/10.3390/su16198482

[20] Hawez, H. K., & Asim, T. (2024). Impact of regional pressure dissipation on carbon capture and storage projects: A comprehensive review. Energies, 17(8), 1889. https://doi.org/10.3390/en17081889

[21] Oladyshkin, S., & Nowak, W. (2012). Data-driven uncertainty quantification using arbitrary polynomial chaos. Water Resources Research, 48(11), W11509. https://doi.org/10.1029/2011WR011473

[22] Köppel, M., Franzelin, F., Kröker, I., et al. (2018). Comparison of data-driven uncertainty quantification methods for a carbon dioxide storage benchmark. CoRR, abs/1802.03064. https://arxiv.org/abs/1802.03064

[23] Oladyshkin, S., Class, H., Helmig, R., & Nowak, W. (2011). Polynomial response surfaces for probabilistic risk assessment and risk control via robust design. https://cdn.intechopen.com/pdfs/35515.pdf

[24] SPE. (2022). CO_2 storage capacity prediction in deep saline aquifers: Uncertainty and global sensitivity analysis. SPE Journal. https://search.spe.org/i2kweb/SPE/doc/onepetro:B2796CE4

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0021

[25] Ashraf, M., Oladyshkin, S., Class, H., & Nowak, W. (2013). Geological storage of CO_2: Global sensitivity analysis and risk assessment using arbitrary polynomial chaos expansion. International Journal of Greenhouse Gas Control, 19, 704–719. https://doi.org/10.1016/j.ijggc.2013.07.014

[26] Chen, B., et al. (2012). Modeling the performance of large-scale CO_2 storage systems: A comparison of different sensitivity analysis methods. National Risk Assessment Partnership, NETL Technical Report NRAP-TRS-III-002-2012. https://www.netl.doe.gov/projects/files/ModelingthePerformanceof%20LargeScaleCO2StorageSystemsAComparisonofDifferentSensitivity_102412.pdf

[27] Wang, H., et al. (2025). Failure analysis–informed risk assessment framework for geological carbon storage using numerical simulation and machine learning. SPE Journal. https://doi.org/10.2118/231438-PA

[28] Zhang, Y. J., Yang, Y., Gong, X. L., et al. (2025). Numerical simulation and parameter global sensitivity analysis of CO_2 geological storage and utilization in oil-water mixed layers. China Environmental Science, (8). https://cnki.net

[29] Moopen, M., Vasylkivska, V., Dilmore, R., & Oldenburg, C. (2024). Probabilistic wellbore leakage assessment for California's geologic CO_2 storage Certificate of Permanence using NRAP-Open-IAM. International Journal of Greenhouse Gas Control, 139, 104172. https://www.sciencedirect.com/science/article/pii/S1750583624000825