energies

MDPI

Review

An Open-Source Wellbore-Stability Modelling Framework for Geothermal Wells: Linking Failure Mechanisms, Evidence, and Lifecycle Decisions

Iman Najafi [1] and Dariusz Knez [2,*]

[1] World Business Council, 02-001 Warsaw, Poland; iman.najafi86@gmail.com
[2] Department of Drilling and Geoengineering, Faculty of Drilling, Oil and Gas, AGH University of Krakow, 30-059 Krakow, Poland
* Correspondence: knez@agh.edu.pl

Abstract

The economic viability of geothermal energy depends on wells that can be constructed and operated reliably. Wellbore instability can delay drilling, require remediation, and interrupt operation. Stability modelling supports risk reduction when its mechanisms and input requirements match the formation, loading history, well geometry, and engineering decision. This critical review compares petroleum and geothermal evidence across drilling, completion, injection, operation, and intervention. We distinguish measured from simulated outcomes, calibration from independent testing, and rock stability from casing and barrier performance. Connected fractures can control circulation losses below an intact-rock fracture limit. Cold injection changes rock and barrier stresses, while long-term deformation is less well constrained at well scale. Strength criteria and bedding assumptions can also change pressure-window predictions. We develop a conditional model-selection framework and companion tool that connect well conditions to model families, required inputs, and observations for testing. Retrospective demonstrations and case assessments examine mechanism coverage and interpretation, including two sources excluded from rule development. Weak-material and fines-transport conditions are explicit, while casing diagnosis requires a separate justification for stress modelling. The outcome-informed assessments establish neither predictive accuracy nor operational benefit. Independent testing against drilling, injection, and intervention observations remains necessary.

Keywords: geothermal wells; wellbore stability; model selection; thermal loading; creep; independent validation

Check for updates

Academic Editor: Reza Rezaee

Received: 7 October 2026
Revised: 20 October 2026
Accepted: 26 October 2026
Published: 29 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Access to subsurface energy resources depends on wells that can be drilled, completed, and operated under difficult geological conditions. This is especially clear in geothermal development. A reservoir may contain sufficient heat and permeability for commercial production, yet that potential cannot be realized if the required wells are too difficult or too costly to construct. The International Energy Agency has emphasized that experience developed in oil and gas—particularly in subsurface characterization, drilling, completion, and high-temperature well construction—can help accelerate conventional and next-generation geothermal development [1]. That experience can inform geothermal design where the governing mechanism and operating conditions are comparable; geothermal-specific thermal, fracture, and lifecycle effects require separate assessment.

The U.S. Department of Energy identifies subsurface access as a major determinant of geothermal project viability and notes that casing and cement materials and services can account for as much as half of geothermal well cost in some designs [2]. Hard, abrasive rock slows drilling, while high temperature restricts the operating envelope of fluids, tools, and downhole measurements. Natural fractures create a more difficult trade-off. They may provide the permeability needed for heat extraction, yet the same structures can become severe loss zones during drilling. Vivas et al. [3] described this combination of temperature, hard rock, drilling-fluid limitations, and fractured formations as a central geothermal drilling challenge.

Lost circulation is therefore not a secondary inconvenience in geothermal wells. Saleh et al. [4] reported that it can represent about 10% of well cost in mature geothermal areas and more than 20% in exploratory or developing fields. A later four-field study by Winn et al. [5] found loss-related costs reaching up to 30% of drilling cost in some settings and showed that loss events were frequently associated with faults, fractures, and productive intervals. Fluid loss can lead to deterioration of hole cleaning and hydraulic support if not controlled on time. Pack-off, stuck pipe, washout, or an unplanned casing decision may follow. The same fracture system can therefore be both part of the resource and a threat to successful well construction.

Drilling decisions must balance competing stability mechanisms. Raising wellbore pressure may suppress compressive failure, but it can increase losses into a connected fracture or an under-pressured interval. Lowering pressure may reduce losses while removing support from the borehole wall. Temperature further changes this balance: cool drilling or injection fluid entering hot rock can alter near-wellbore stresses and shift collapse or fracture limits. The appropriate operating range therefore depends on the local fracture system, pressure history, thermal contrast, and strength assumptions.

Thermoporoelastic formulations describe how temperature and pore-pressure changes modify near-well stresses under specified boundary conditions [6–8]. The wellbore-stability problem also extends beyond the open hole. Poor hole geometry can complicate casing placement and cement displacement, while loss zones may prevent cement from remaining where it is required. Reviews of geothermal well integrity and cementing identify high temperature, corrosive fluids, lost circulation, casing design, and cement durability as recurring concerns [9,10]. During later operation, heating, cooling, injection, and shutdown impose repeated loads on the casing–cement–formation system. Deng et al. [11] modelled plastic deformation, interface debonding, and microannulus development under alternating thermal and pressure loading; t Field evidence from IDDP-1 shows severe corrosion and cracking of carbon-steel components after extreme geothermal exposure [12]. These outcomes require distinct engineering checks, although construction conditions connect them through the well lifecycle.

The relevant methods of wellbore-stability assurance differ in what they calculate and in how directly their outputs have been tested. Elastic wall-stress solutions give a transparent starting point [13,14]; comparison of strength criteria shows that the same stress state can produce different collapse limits [15–17]. Weak-plane formulations alter the predicted failure geometry in laminated rock [18,19]. Poroelastic, chemical, and thermal models address loading paths that an instantaneous elastic calculation omits [20–26]. Discontinuum models, uncertainty propagation, and data-driven classifiers address different questions again [27–30]. Treating these methods as a single ladder of increasing sophistication obscures the distinct evidence needed to judge each one.

The more consequential gap lies between a calculated limit and an independently checked forecast. Breakout images and stress tests may calibrate a pre-drill model, yet a model fitted to those observations has not predicted them independently. Formation-

specific loss histories, calliper enlargement, pressure history, and later cementing or integrity records offer subsequent tests, but few studies retain the original prediction and follow the resulting well through those stages. The distinction matters most in geothermal wells, where a fracture that causes drilling losses may also provide useful flow during production or injection.

This review asks three questions:

- Which physical mechanisms alter a wellbore-stability decision?
- What evidence supports the corresponding model in its tested setting?
- Which observation would reveal a failed transfer to a new well or lifecycle stage?

Earlier reviews have examined numerical and discrete-element approaches to wellbore stability [27], machine-learning prediction [30], and geothermal well integrity and cementing [9,10]. This review asks a different, decision-level question: which model mechanisms are warranted for a specified geothermal well condition, what evidence supports their transfer from petroleum or other settings, and what later observation could test the recommendation? The contribution is an auditable mechanism–decision–observation mapping across the geothermal well lifecycle: it connects a stated condition to a model family, its input burden, and an outcome reserved for testing. Section 4 implements that mapping as categorical rules. The selector is decision support, not a new numerical geomechanical model or a validated failure predictor. Rock failure, circulation loss, casing loading, and barrier leakage remain distinct outcomes; agreement with one does not validate prediction of the others.

Geothermal operation adds loading paths that a drilling pressure window does not describe. Injection changes temperature and pore pressure around an open or perforated interval; production and shut-in cycle the casing–cement–formation assembly; and corrosion, scale, or casing damage can govern later intervention [11,31–37]. In susceptible formations, deformation may also accumulate with time. The connection between stages is physical, but their measurements are not interchangeable: a casing pressure test assesses a barrier, whereas a loss event or borehole image bears more directly on rock failure. The review follows these outcomes across the well lifecycle while retaining separate validation targets.

Figure 1 summarizes the principal mechanisms across the geothermal well lifecycle, linking controlling conditions to engineering decisions and the observations needed to assess rock stability, fracture-controlled flow, and barrier performance.

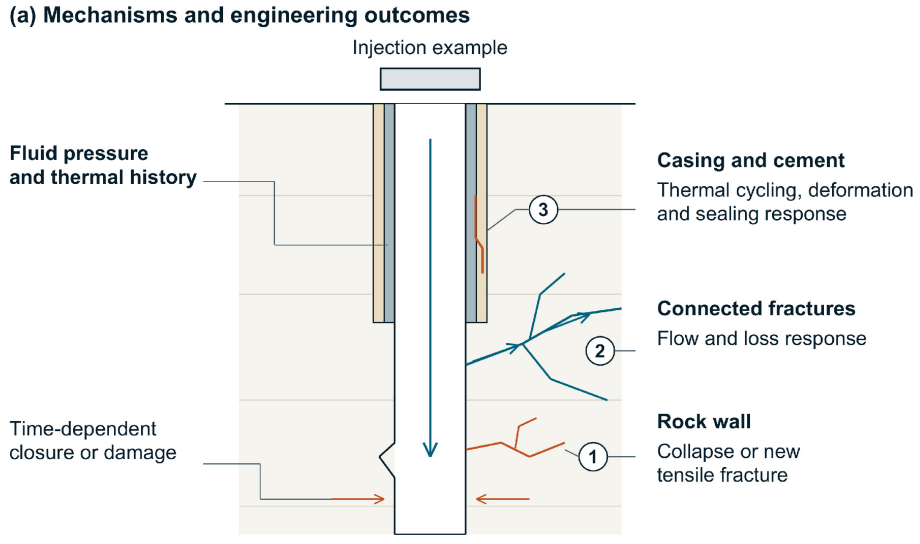


Figure 1. *Cont.*

(b) Lifecycle decisions and test observations

Lifecycle stage	Controlling conditions	Engineering decision	Test observations
Before drilling	Stress, fabric and strength	Trajectory and initial pressure window	Cores, image logs and pressure tests
Drilling	Pressure, exposure and connected fractures	Collapse and loss management	Cavings, breakouts and loss depth
Completion	Placement and barrier geometry	Cement placement and barrier assessment	Placement monitoring and pressure tests
Injection / production	Thermal and pressure history; time-dependent deformation	Rate, pressure and temperature limits; monitoring	Pressure-temperature history, injectivity and deformation
Maintenance	Existing damage and re-entry loading	Inspection, repair and re-entry	Caliper or camera; post-repair tests

Figure 1. Geothermal wellbore mechanisms, lifecycle decisions, and test observations. (**a**) An idealized injection-well section distinguishes rock-wall failure (1), flow through connected natural fractures (2), and casing–cement deformation or sealing response (3). Pressure and thermal histories affect these outcomes through different mechanisms. Time-dependent closure or damage requires a separate constitutive assessment. The drawing is schematic and not to scale. (**b**) Representative controlling conditions, engineering decisions, and test observations across the well lifecycle. Mechanisms may overlap lifecycle stages; rock stability and barrier performance require distinct assessment criteria and observations.

2. Review Method and Evidence Base

2.1. Source Identification and Eligibility

This structured critical review synthesizes 210 source records addressing wellbore-stability mechanisms, model requirements, diagnostics and geothermal lifecycle decisions. Selection is purposive: sources represent foundational formulations, mechanisms capable of changing an engineering decision, and geothermal observations across construction and operation. Source identification combines topical searching and citation tracing across rock-mechanics, petroleum-engineering and geothermal journals and technical proceedings. OpenAlex provides discovery and citation information; Crossref supports bibliographic verification; publisher platforms, including ScienceDirect and Taylor & Francis Online, and conference archives provide article records and accessible texts. Citation counts inform reading priority. Reliability is assessed from the specific methods, observations and transfer conditions supporting a claim. S1 distinguishes sources used directly in manuscript claims or rule rationale from additional mechanism/context records and records their review extent under stable identifiers.

Eligibility requires a connection to rock failure, fracture-controlled loss, thermal or chemical loading, time-dependent deformation, or a casing or barrier decision. Petroleum and material-scale evidence is retained where the mechanism and transfer limits are relevant to geothermal applications. Titles and abstracts establish scope; original methods and results resolve ambiguous relevance. Each exclusion has a reason and evidence locator. Resource assessments, regional deformation, equipment durability and monitoring methods without a specified wellbore decision are excluded. Inaccessible full text alone does not justify exclusion. Source type and publication status are checked separately from publisher identity. Duplicate checks compare DOI, title, authors, year and version; repeated copies count once. Related reports are linked without assuming identical evidence. Editorials provide context.

Indexed abstracts are distinguished from original proceedings papers and coded only from the named entry. Their support is qualitative and cannot establish unreported sample sizes, uncertainty or predictive validity.

2.2. Coding and Claim-Specific Appraisal

Each source has one primary lifecycle stage, assigned from its principal engineering decision or observed outcome; secondary tags identify additional stages. Figure 2 is generated from S1. Current model-family coding describes the appraised method; a mention of a technique in an article's references does not establish that the article uses it. Application-domain coding uses supporting source passages where available, and mixed or indeterminate domains remain explicit. Context records do not supply a sample denominator, comparative prediction-error estimate or validation claim. Historical ordinal labels are retained separately from current appraisal and are not quality findings. Screening and appraisal decisions are recorded in a consolidated matrix; independent duplicate screening and inter-assessor agreement statistics are not claimed. A discrepant interpretation is checked against the source and criterion, with the alternatives and reason for resolution recorded. An unresolved field remains not established and does not support a quantitative claim.



Figure 2. Review workflow linking source identification, eligibility and identity checks, claim-specific appraisal, mechanism–decision synthesis, and conditional model selection. The diagram represents the analytical workflow rather than a database retrieval-count flow.

Mechanism-based comparative analysis groups sources addressing the same decision and failure process. For each directly used source, appraisal records the model family, study unit, predicted or measured outcome, calibration–test relationship, uncertainty and transfer limit. Review extent identifies original-paper passages, publisher full text, preprints or abstract-only support. A source that supports a mechanism description may not support an estimate of predictive reliability. Sample sizes are recorded only where the source establishes the relevant population; cited background examples and synthetic model sets are not counted as independent field wells. Quantitative statements are checked against their tables or results, and conflicts are retained without an unsupported correction. A laboratory comparison supports its tested material and load path; a diagnostic field association does not establish forecast accuracy. No aggregate quality score or pooled model-accuracy ranking is inferred. S1A links each categorical rule to sources, required observations, engineering judgement and limitations.

2.3. Synthesis and Evaluation Boundaries

This critical synthesis does not estimate field-wide prevalence or claim exhaustive selection. Nine OpenAlex query families provide a separate discovery inventory covering drilling, completion, injection, production, thermal cycling, creep and maintenance/intervention. OpenAI was used for automating the process of research and finding the most relevant papers and ranking them based on number of citations, using OpenAlex API. Exact requests, retrieval totals and recorded dispositions are supplied in S1A. The broad inventory is a coverage and citation-discovery resource; it is not represented as a completely screened systematic-review pool. Only sources with recorded inclusion and a stated claim/context role enter the synthesis. Unassessed inventory records are neither included nor treated

as exclusions. Repeated identifiers are reconciled and related publication versions are linked. Source records are not independent experiments, and overlapping sample counts are not pooled. Heterogeneous materials, load paths, scales and outcomes preclude pooled prediction-error ranking. Table 1 summarizes findings and inference limits. Table 2 gives conditional recommendations.

2.4. Framework Evaluation

Framework evaluation distinguishes demonstrations of rule application from coverage checks using sources excluded from rule development. Our online model selector, Wellbore Studio Version 0.2.0, is evaluated with published geothermal cases spanning drilling, injection, stimulation and maintenance. Each assessment records source-supported conditions, analyst-assigned categories, selected model layers, a conceptual comparator, reported observations, and mechanisms outside the selected layers. Blank inputs represent unknown conditions. Complete case encodings and source locators are supplied in S1A and Supplementary Material S2.

The two sources used for the separate coverage evaluation were selected after the engine, rules and comparison criteria were frozen. Published outcomes remained visible during encoding. The evaluation therefore assesses mechanism coverage and recommendation sensitivity; it does not estimate blinded prediction error or operational benefit (Figure 2). Rule checks vary thermal, fracture, uncertainty and data-evidence inputs. These categorical perturbations are reported separately from the physical stress–strength–temperature calculations used to illustrate decision sensitivity.

3. Evidence Map and Lifecycle Findings

The primary-stage evidence map is concentrated on design and drilling decisions, with fewer records assigned to operation or crosscutting questions (Figure 3). Each source contributes once to the primary-stage total. Secondary tags retain construction, injection or production, thermal cycling, and intervention coverage where recorded. This distribution describes the assembled synthesis and does not estimate the publication balance across the field. The smaller operation group limits the support available for choosing creep laws or forecasting decades of service.

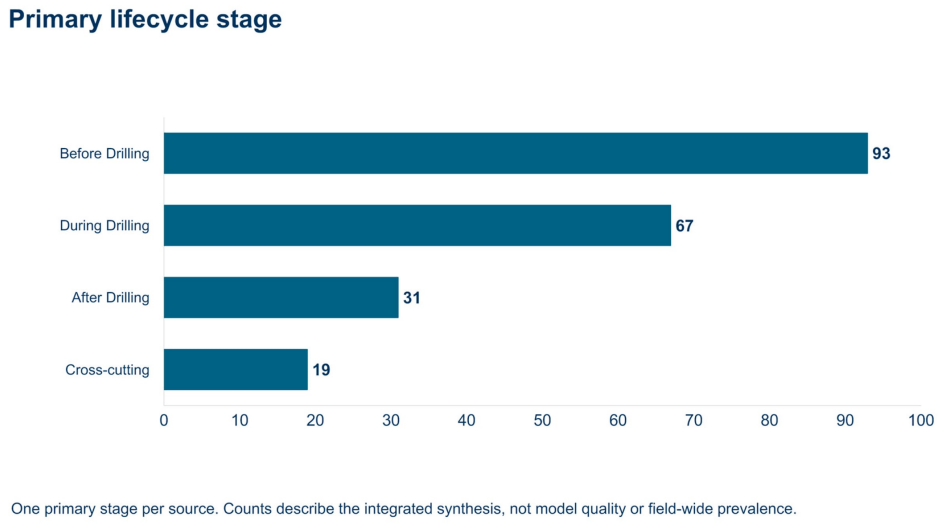


Figure 3. Mutually exclusive primary lifecycle classifications generated from Supplementary Material S1. A source is counted once according to its principal coded decision or outcome. Counts describe this synthesis, not model quality or field-wide research prevalence; the record-level screening status is retained in S1.

3.1. Before-Drill Wellbore-Stability Modelling Forecasts

Before drilling, stability models combine stress, pore pressure, wellbore pressure, and rock strength to estimate provisional collapse and fracture limits. These estimates guide mud-weight and trajectory decisions, subject to the quality of the input data. Mechanical Earth Models add stratigraphic and offset-well constraints, but greater detail does not remove uncertainty in stress magnitude or strength [38–43]. McLean and Addis showed that changing the compressive failure criterion can materially change the calculated minimum mud weight for the same input state [15]. Ma et al. obtained a lower minimum mud weight with Mogi–Coulomb than with Mohr–Coulomb in their breakout-width analysis [16]. Neither comparison establishes which criterion will predict a new well more accurately; it identifies model-form sensitivity that must be checked against observed failure.

This uncertainty is not confined to the lower, collapse-related boundary of the stability window. The calculated upper boundary also depends on formation properties, well geometry, and the simplifying assumptions of the model. A tensile-failure calculation estimates the pressure required to initiate a new fracture in intact rock, but drilling fluid may enter or reopen an existing natural fracture at a lower wellbore pressure. The operational upper limit may therefore be controlled by fracture connectivity and activation rather than by intact-rock tensile strength. Case studies from sandstone and offshore wells show why calculated fracture and collapse limits should be checked against pressure tests, losses, and other drilling observations [44–48]. This distinction is particularly important in naturally fractured geothermal formations. Winn et al. [5], for example, associated many loss events across four geothermal fields with faults, fracture intersections, and productive intervals. These observations demonstrate where losses occurred and how they were related to geological structure; they do not define a universal fracture-pressure threshold. For wellbore-stability analysis, the implication is important: a calculated mud-weight window may appear mechanically safe while the actual well begins losing circulation at a lower pressure through connected fractures. The upper operational limit should therefore reflect both intact-rock tensile failure and the pressure at which existing fractures can become hydraulically active.

Several studies demonstrate substantial model-form uncertainty. For example, Gholami et al. reported a 1.42–1.48 g/cm^3 safe mud-weight window with Mohr–Coulomb and a wider 1.37–1.80 g/cm^3 range under the other criteria considered [49]. This result demonstrates decision sensitivity to criterion choice in that well, not a ranking transferable to every lithology. Other uncertainty analyses propagate uncertain inputs through collapse and fracture calculations [50,51], while anisotropic comparisons show that bedding orientation and weak-plane strength can change the predicted failure location and trajectory preference [18,19,52]. Image-log and borehole-fracture studies constrain stress orientation and failure geometry [53–56]; they become an independent test only when the relevant prediction was specified before those observations were used for fitting.

3.2. During Drilling Wellbore-Stability Modelling

During drilling, the pre-drill stability model can be tested against actual well behaviour. The pressure at which losses occur can indicate whether natural fractures have opened or become hydraulically connected, while the depth and character of cavings provide evidence of borehole-wall failure. Comparing these observations with the predicted stability window helps determine whether the failure mechanism assumed before drilling is consistent with the response of the well.

Geothermal field evidence shows why this distinction matters. Winn et al. [5] examined McGinness Hills, Don A. Campbell, Steamboat Hills, and Puna and found that circulation losses were commonly associated with faults and fracture intersections, with more loss events occurring in production intervals. This suggests that the upper operational

limit of the well may be controlled by connected natural fractures rather than by a single intact-rock fracture gradient. The observations do not imply that mud pressure should always be increased or decreased when losses occur. Instead, the stability model should be updated to reflect the geological structure and hydraulic behaviour of the interval. Well trajectory also matters because the likelihood of intersecting or activating fractures depends on their orientation relative to the borehole and the stress field, while high temperatures can further change the allowable operating range [57]. The appropriate response is therefore to reconsider the assumed failure mechanism and the resulting stability window, rather than simply adjusting a pressure threshold until it matches the observed losses.

Time-dependent shale instability provides another way to evaluate the pre-drill stability model. Bol et al. [21] showed that pressure penetration and shale hydration can contribute to instability after the formation has been exposed to drilling fluid. Coupled models developed by Chen et al. [22] and Yu et al. [23] further show that osmotic effects, membrane efficiency, and solute diffusion can alter near-wellbore pore pressure and, consequently, the effective stresses controlling borehole failure. These studies support the need to consider exposure time and shale–drilling-fluid interaction when instability develops progressively rather than immediately after drilling.

However, the parameters governing these processes are formation- and fluid-dependent and should not be transferred directly to another shale or drilling-fluid system without appropriate laboratory calibration [24–26]. In the field, increasing cavings with open-hole exposure time can therefore indicate time-dependent shale instability, particularly when accompanied by changes in drilling-fluid properties or other evidence of fluid–rock interaction. Cavings alone are not sufficient to identify the mechanism, since poor hole cleaning and mechanically induced borehole failure can produce similar drilling symptoms.

Recent data-driven studies in geothermal drilling have focused on real-time fracture and drilling-event detection rather than direct prediction of wellbore collapse or the safe operating pressure window. Mellal et al. combined low-cost drilling measurements, continuous wavelet transforms, and machine learning to identify fractures in Utah FORGE data and evaluated the method on a separate well [58]. Aspiras et al. investigated real-time warning of drilling incidents from geothermal drilling time-series data, including operations conducted under total loss of returns. Such approaches can improve early recognition of fractures and abnormal drilling conditions, particularly when their performance is evaluated on wells that were not used for model development. Their outputs, however, remain diagnostic indicators rather than geomechanical stability limits. A fracture indication or drilling-event warning does not by itself determine whether mud weight or ECD should be changed, whether a lost-circulation treatment is required, or whether another operational response is appropriate. Demonstrating operational value therefore requires the detected event to be linked to the actual well condition, the action taken, and the resulting drilling response.

3.3. Post-Drill Assessment and Model Corrections

After drilling, borehole observations provide an opportunity to check whether the assumptions used in the wellbore-stability models were consistent with the behaviour of the drilled well. Image logs, calliper data, drilling-induced tensile fractures (DITFs), and breakouts can be used to reassess the in situ stress model and the predicted failure mechanism. In the Soultz and KTB wells, Brudy and Zoback [53] compared observed DITFs with independent stress information and examined the effects of mud pressure and borehole cooling on fracture initiation. Wiprut et al. [54] combined image-log observations with other stress constraints to estimate stress orientation and magnitude and assess wellbore

stability for different well trajectories. Other field and modelling studies have similarly used breakout and fracture geometry to improve stress interpretation [55,56,59].

These observations are valuable for post-drill model calibration, but calibration should not be confused with independent validation. If an image log is used to adjust the stress model, the same image cannot then be used as independent evidence that the revised model is correct. The calibrated model should instead be tested against observations that were not used in the calibration, such as failure in another interval or a subsequent well in the same formation.

A useful post-drill review should therefore compare the original stability prediction with what actually occurred while drilling. The original mud-weight window, predicted unstable intervals, input ranges, and assumed failure mechanism should be retained together with the actual mud weight/ECD and pressure-temperature history. These predictions can then be compared with the observed depths of breakouts, hole enlargement, DITFs, losses, or other instability events.

The purpose of this comparison is not simply to obtain a better match after the well has been drilled. It is to identify why the original model differed from the observed behaviour. The discrepancy may result from an incorrect stress estimate, uncertain rock strength, an inappropriate failure criterion, unrepresented anisotropy or fractures, or thermal and hydraulic effects that were neglected. Numerical breakout studies also show that similar borehole geometries can be produced under different constitutive assumptions [60]. A match to breakout width or shape alone therefore does not prove that the model has represented the correct failure mechanism.

Figenschou et al. [61] describe post-drill diagnosis in the Ringhorne field using ultrasonic calliper measurements, high-resolution images and time-based drilling parameters to investigate borehole damage and shear failure. The accessible publisher abstract supports this diagnostic use. It does not establish a reserved prediction test or demonstrate that the analysis improved a later drilling decision. S1 identifies this abstract-supported contextual citation separately from the synthesis records.

The main value of post-drill analysis is therefore the recalibration of the stability model for the next interval or well. The revised model should record which assumptions were changed, why they were changed, and whether those changes improve prediction when applied to new data. This closes the lifecycle loop between pre-drill prediction, observations acquired while drilling, and post-drill validation.

3.4. Numerical Models and Uncertainty

Numerical modelling is useful when fractures, discontinuities, or progressive rock damage cannot be represented adequately by analytical models. Santarelli et al. [29] examined instability in heavily fractured rock, while continuum-damage and crystalline-rock studies showed how damage and fracture networks influence breakout development [62,63]. These studies support numerical modelling of complex failure, but laboratory agreement does not demonstrate that the same model can predict field-scale losses through connected fractures. Greater numerical resolution cannot compensate for uncertain fracture geometry or poorly constrained properties.

Uncertainty analysis addresses a different problem. Gholami et al. [49] showed that uncertain inputs and the selected failure criterion can significantly change the predicted mud-weight window, with maximum horizontal stress being particularly influential. Probabilistic methods propagate uncertainty in stress, strength, and pressure into failure estimates [28,50,51], but they cannot correct for a missing failure mechanism. Bayesian and machine-learning approaches can further update predictions [64], although their re-

liability should ultimately be tested against wells or drilling observations not used for model calibration.

3.5. Geothermal Injection, Operation, and Maintenance

Cold-water injection creates a coupled thermal and hydraulic load not represented by the drilling mud-weight window. Huan et al. simulated a casing–cement–formation assembly and found a stress concentration at the cased-to-open-hole junction; in a scenario with a 160 °C injection–formation temperature difference, computed hoop tensile stress in granite reached 43.7 MPa and radial tensile stress at the casing–cement interface reached 31.9 MPa [31]. In preliminary workshop scenarios, Hals and Berre found predicted tensile damage sensitive to well orientation and injected-water temperature; injection-rate dependence was weaker within their tested model conditions [32]. These are mechanistic predictions under specified boundary conditions. They identify what to monitor and test, but neither supplies a field-validated operating threshold for another reservoir.

Well-integrity studies address a different outcome from open-hole stability and show the importance of the well's thermal history. Deng et al. [11] modelled cement-sheath damage under cyclic thermal and pressure loading. In contrast, Corina and Moghadam [37] tested a 1.5 m casing–cement assembly through 39 cycles between 20 and 80 °C and found no increase in cement permeability, although a microannulus was present after curing. These results are not contradictory because the materials, geometry, loading conditions, and measured responses differ. Marbun et al. [33] further examined thermally induced liner loads in a high-temperature injection well. Together, these studies show that formation stability, casing loading, and barrier leakage should be evaluated separately, using the actual pressure and temperature history of the well.

Maintenance evidence is predominantly diagnostic rather than predictive. In two long-abandoned Dieng wells, Marbun et al. examined casing condition and workover feasibility but lacked original design, cement, and operational records needed for definitive failure attribution [35]. Their related design study considered changing casing loads through production and possible conversion to injection [34]. In Poland, Bujakowski et al. reported borehole reconstructions guided by inspection and intervention records, including repair of a leaking injection well with a cemented liner before return to service [36]. These cases show that an intervention can restore a barrier without testing the original open-hole collapse calculation. They also show why missing well-history data limit any attempt to infer which earlier design assumption failed.

The available geothermal sources therefore support different levels of inference at different stages. Multi-field drilling records identify where losses occurred [5]; coupled injection studies predict stresses rather than measure failure thresholds [31,32]; a low-enthalpy laboratory assembly tests sealing under a defined cycle [37]; and intervention reports document defects and repair outcomes [35,36]. Field fibre-optic monitoring adds direct evidence of cement placement and hydration during construction [65]. No single source among these follows one frozen rock-stability forecast through drilling, injection, barrier degradation, and workover. That missing longitudinal comparison limits claims of lifecycle validation more than the number of studies in any stage category.

4. Evidence-Constrained Model Selection

The comparisons above support a conditional choice of model family, not a universal ranking. A method is relevant when it represents the mechanism capable of changing the engineering decision, and its parameters can be constrained in the setting where that decision is made. The evidence needed to support the choice differs by output: image-log geometry can test a stress or breakout interpretation, a loss event challenges the assumed

upper pressure mechanism, and a pressure or leak test assesses a completed barrier. Table 2 states these links and the limits that prevent direct transfer. A mechanism is treated as decision-critical when a plausible range of its effect could move a predicted limit across a proposed pressure, trajectory, injection, or barrier decision. If that effect is unmeasured or smaller than the decision margin, a sensitivity calculation is appropriate before promoting a coupled model to the primary design model.

4.1. What the Comparative Evidence Supports

Analytical wall-stress and failure calculations remain valuable because their assumptions are visible. The criterion comparisons of McLean and Addis, Ma et al., and Gholami et al. show that changing the failure rule can alter mud-weight recommendations even without adding more coupled physics [15,16,49]. Lee et al. show a different kind of model inadequacy: in laminated rock, the predicted failure location and lower pressure bound depend on bedding orientation and weak-plane strength [18]. These findings support testing criterion and fabric sensitivity before increasing numerical detail. They do not justify selecting one criterion for every stress regime or assigning an operational limit without local strength and observation data.

Time-dependent loading introduces additional model requirements. Shale experiments and chemical–mechanical models connect exposure, water activity, and pore-pressure evolution to delayed failure [21–23]; geothermal injection simulations connect cooling and pressure diffusion to stresses at both open-hole rock and cased interfaces [31,32]. Adding those mechanisms is warranted when their time scale and estimated magnitude overlap the proposed operation. The transfer limit remains material-specific: membrane efficiency and transport parameters require mud–shale tests, whereas thermal and barrier calculations require temperature, material, and pressure histories. Numerical detail without matching material and loading constraints can make a prediction appear precise while leaving its controlling parameters unidentified.

Table 1. Descriptive source comparisons: reported model result, evidence type, and transfer limit.

Comparison	Evidence and Setting	Reported Result	Boundary of Inference
Failure criterion [15,16,49]	Analytical comparisons; well risk study	In [49], Mohr–Coulomb gave 1.42–1.48 g cm^{-3}; the other criteria considered gave 1.37–1.80 g cm^{-3}	Criterion sensitivity in that well, not a universal accuracy ranking
Bedding anisotropy [18,19,52]	Weak-plane models and case analyses	Bedding orientation changes predicted failure position and lower pressure bound	Requires local fabric orientation and plane strength
Shale chemistry [21–23]	Laboratory observations and coupled transport models	Fluid activity and ion transport alter pore pressure and possible failure location	Mud–shale properties and exposure time are material-specific
Fractured rock [29,62,63]	Fracture analysis and laboratory damage comparison	Discontinuities and damage laws change predicted breakout or loss response	Breakout agreement does not validate fracture flow
Uncertainty [28,49–51]	Single-well risk study and parameter propagation	Input and criterion uncertainty widen possible pressure windows	Probabilities depend on input distributions and model form

4.2. Conditional Model Choice

Table 1 summarizes what selected studies report. Table 2 is the authors' prescriptive model-selection guide derived from that evidence. For each condition, Table 2 identifies a

starting model, the decision it informs, an observation that could test it, and the boundary beyond which transfer is unsupported. It does not rank software or imply that every method has been validated in geothermal wells. Figure 4 links conditional model selection to site-specific calibration, sensitivity assessment, and independent testing. Figure 5 presents the corresponding categorical model-selection matrix. In the Wellbore Studio version 0.2.0 rule set, Required means that the selected condition makes omission of that modelling layer inconsistent with the stated decision; Recommended means the layer should normally be considered but its decision effect still requires local checking; Conditional means it is promoted only when data and sensitivity justify it; and Support Only denotes a diagnostic or auxiliary layer that does not itself define a geomechanical limit. These are authors' engineering judgments informed by the cited mechanism studies, not categories estimated from comparative prediction-error data. When several conditions match, the tool retains the highest category for each layer (Required above Recommended above Conditional above Support Only). S1A's Rule Trace and the archived source list the exact condition-to-layer assignments. Weak or poorly consolidated material has a distinct condition and does not automatically activate the competent-rock rule. An explicit fines/clogging mechanism requires a transport and injectivity assessment. Post-drill diagnostic evidence makes MEM integration Recommended; a separate rock/stress decision is needed to establish its necessity for casing inspection.

The level of evidence required should reflect the consequence of the drilling decision. Sensitivity analysis may be sufficient for preliminary screening, whereas mud-weight, trajectory, or casing decisions require locally constrained stress, pore pressure, rock strength, and fracture or loss data. More complex coupled or data-driven models are justified only when the added mechanism can change the decision and the prediction can be checked against independent observations. The framework should therefore be used as an evidence-based model-selection guide, not as a field-validated design standard. The thermal rule is activated by the user's "thermally disturbed" condition, not by geothermal use alone: the user must first judge whether plausible cooling, heating, and operating duration could change the decision. No numerical temperature-difference cutoff is encoded. High decision-changing uncertainty promotes probabilistic treatment; low data availability or absent validation adds a warning in the evaluated baseline. Consequently, an output with missing stress, fracture, or thermal parameters is a screening recommendation and cannot establish a safe operating limit. Stress magnitude, strength and bedding tests, fracture/loss pressure and connectivity, and measured temperature histories are the observations most likely to resolve the choice.

Table 2. Prescriptive guide linking well conditions to model choices, decisions, and independent checks.

Condition	Evidence-Based Starting Model	Decision or Output	Independent Check and Avoid-Use Boundary
Competent rock; intact-wall failure or new tensile-fracture initiation [13–17]	Elastic wall stress and locally calibrated compressive/tensile failure criterion	Initial collapse, trajectory, and intact-fracture initiation screen	Compare breakouts and pressure tests; this is not a connected-fracture loss limit
Bedded rock [18,19,52,66,67]	Weak-plane or anisotropic strength	Collapse and trajectory sensitivity	Image bedding and failure geometry; avoid isotropic transfer
Reactive shale [21–26]	Time-dependent poroelastic and chemical transport	Fluid chemistry and exposure decision	Test mud–shale interaction and caving timing; avoid borrowed membrane parameters

Table 2. *Cont.*

Condition	Evidence-Based Starting Model	Decision or Output	Independent Check and Avoid-Use Boundary
Connected natural fractures or losses [5,29,63]	Fracture-aware hydraulic and geomechanical interpretation; explicit discontinuities when decision-relevant	Loss onset and reservoir-compatible pressure management	Map loss depth, dynamic pressure, connectivity, and treatment response; do not use intact-rock tensile pressure as the loss threshold
Cold injection [31,32]	Coupled thermal–hydraulic–mechanical sensitivity	Injection pressure and temperature envelope	Record field cycles and fracture response; avoid transferring scenario stress thresholds
Cased barrier cycling [11,37]	Casing–cement structural and sealing assessment	Integrity and inspection plan	Leak or pressure test; avoid equating rock stability with sealability
Narrow uncertain window [28,49–51]	Mechanism model with uncertainty propagation	Casing and pressure contingency	Constrain input distributions; avoid precise probabilities from sparse data
Drilling incident warning [58,68–71]	Time-ordered detection model plus physical baseline	Warning or diagnostic aid	Hold out another well; avoid treating classification as a failure pressure
Cement placement [65]	Placement and hydration monitoring	Identify incomplete construction	Check later barrier integrity separately
Workover [35,36]	Inspection-led barrier and rock re-entry assessment	Repair or re-entry decision	Use camera/calliper and post-repair pressure test; avoid inferring original rock-model validity

4.3. Case-Based Boundaries of Transfer

Field evidence also defines the limits of model transfer. Winn et al. [5] found that geothermal circulation losses frequently occurred near faults and within productive intervals. In such settings, connected natural fractures may take fluid below the tensile-fracture pressure predicted for intact rock. The appropriate upper operating limit therefore depends on the actual loss depth, pressure history, and role of the fracture in reservoir performance rather than on a universal correction to fracture pressure.

Borehole images provide useful evidence for stress and stability interpretation. Studies from Soultz, KTB, and the Norwegian margin used drilling-induced fractures together with other measurements to constrain stress conditions and assess borehole response [53,54]. These observations can calibrate a stability model, but calibration is not the same as prediction. Stronger validation requires the calibrated model to predict breakouts, tensile fractures, or losses in another interval or well using observations that were not used to build the model.

The same caution applies to injection and well-integrity studies. The 43.7 MPa thermally induced hoop stress reported by Huan et al. [31] is a modelled result for a specified temperature difference, not a universal geothermal failure limit. Likewise, the absence of increased permeability after 39 thermal cycles in the experiment of Corina and Moghadam [37] should not be interpreted as a general safe-cycle limit. Field assessment requires the actual pressure and temperature history together with local formation and barrier properties. Workover

evidence has a similar limitation: successful liner repair demonstrates the effectiveness of that intervention [36], but does not validate the original stability model. Where failure histories are incomplete, as reported for the Dieng cases [35], the cause of damage may remain uncertain. Closing the lifecycle loop therefore requires pre-drill predictions to be retained and compared with subsequent drilling, temperature, pressure, inspection, and integrity observations.

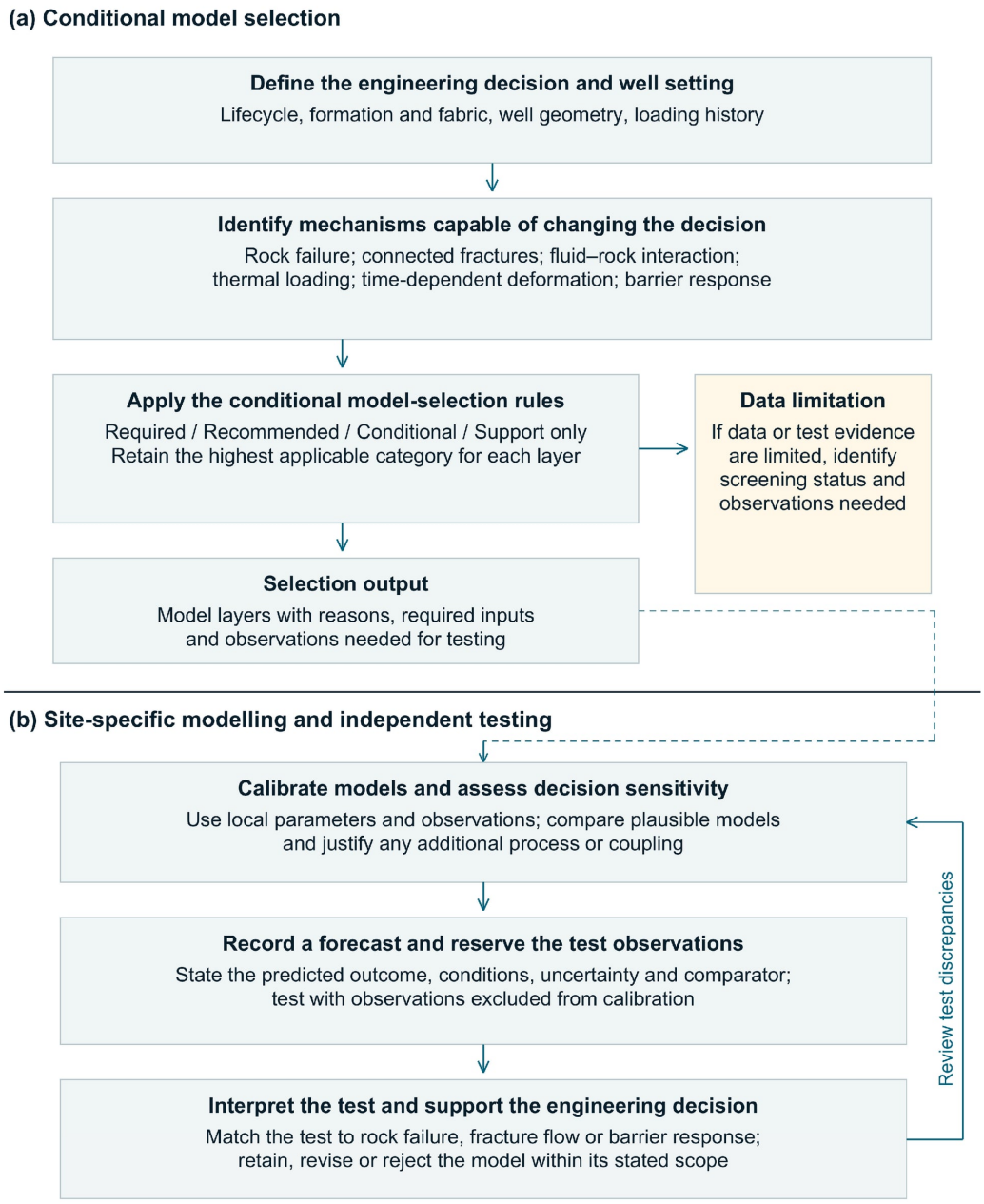


Figure 4. Conditional model selection and subsequent modelling and testing. (**a**) The engineering decision and well conditions identify relevant mechanisms and activate categorical model-layer recommendations. The highest applicable category is retained for each layer. Limited data or missing validation evidence require screening guidance and identification of the observations needed for further assessment. The dashed connection links selection to site-specific application. (**b**) Model calibration and sensitivity assessment precede a recorded forecast and testing against observations excluded from calibration. Tests address the specified rock-failure, fracture-flow, or barrier-response outcome; discrepancies inform model revision. This panel describes the application procedure rather than completed independent predictive validation. Numerical operating limits and time-dependent creep require appropriate site-specific analyses.

Categorical model-selection rules v0.2.0

Condition	Elastic	Failure	MEM	Bed.	Poro	Chem	THM	Num.	Prob.	AI/ML	Post	Barrier	Fines
Early screening	R	R	C	–	–	–	–	–	Rec	–	–	–	–
Competent rock	R	R	Rec	–	–	–	–	C	C	–	–	C	–
Deviated well	R	R	Rec	C	–	–	–	C	Rec	–	–	C	–
Bedded / anisotropic	Rec	R	Rec	R	C	C	–	C	Rec	–	–	C	–
Reactive shale	Rec	R	Rec	C	R	R	C	C	Rec	S	–	Rec	–
HPHT / deepwater	Rec	R	R	–	Rec	–	R	C	Rec	–	–	Rec	–
Thermal effects	–	Rec	R	–	Rec	–	R	C	Rec	–	–	Rec	–
Hydrate-bearing	–	C	R	–	Rec	–	R	Rec	C	–	C	Rec	–
Fractured / faulted	C	Rec	R	C	–	–	C	R	Rec	–	Rec	R	–
Uncertain window	R	R	R	C	C	–	C	C	R	–	–	Rec	–
Real-time symptoms	C	Rec	R	C	C	C	C	C	Rec	S	C	Rec	–
Post-drill check	C	Rec	Rec	C	C	C	C	C	C	S	R	R	–
Injection / production	–	C	Rec	–	C	–	C	C	Rec	–	Rec	R	–
Maintenance	–	C	Rec	–	–	–	C	C	Rec	–	R	R	–
Multiwell labels	–	–	Rec	–	–	–	–	–	Rec	Rec	Rec	C	–
Weak / unconsolidated	C	C	C	–	Rec	–	–	C	Rec	–	–	C	C
Fines / clogging	–	–	–	–	Rec	–	–	C	Rec	–	–	C	R

R Required Rec Recommended C Conditional S Support only – Not primary

Figure 5. Categorical model-selection rules (selector v0.2.0). R, Required; Rec, Recommended; C, Conditional; S, Support Only; dash, not primary. Elastic, Kirsch stress; Failure, strength criterion; MEM, Mechanical Earth Model; Bed., bedding anisotropy; Poro, poroelasticity; Chemo, shale–fluid interaction; THM, thermal–hydraulic–mechanical coupling; Num., numerical analysis; Prob., uncertainty; Post, back-analysis; Barrier, integrity linkage; Fines, particle-transport and injectivity assessment. Categories express author judgement, not measured accuracy. Full labels and rules are supplied in S1A.

4.4. Validation and Predictive Transfer

Gholami et al. [49] provide a concrete model-form sensitivity check for one petroleum well. Their reported greater-than-90%-certainty safe density window is 1.42–1.48 g cm^{-3} with Mohr–Coulomb and 1.37–1.80 g cm^{-3} with Hoek–Brown, modified Lade, or Mogi–Coulomb. A hypothetical 1.50 g cm^{-3} density lies above the former window but inside the latter three; the classification reverses with criterion choice. Their sensitivity analysis identifies maximum horizontal stress as the most influential uncertain input. Breakout or drilling-induced-fracture observations, leak-off or formation-integrity tests, and local strength tests would constrain the stress and strength assumptions before transferring any window. S1A reproduces this threshold comparison without assigning probabilities or treating the petroleum result as a geothermal operating limit. S1A also supplies an explicitly illustrative intact-wall sensitivity calculation. With a vertical circular borehole, principal horizontal stresses aligned with the cross-section, and a simplified elastic thermal correction, uncertainty in minimum stress, maximum stress, tensile strength, or cooling can reverse the comparison between a proposed pressure and the calculated initiation pressure. The inputs are analyst-selected examples rather than geothermal field estimates. Fracture-connected loss requires a separate pressure and connectivity observation; it cannot be inferred from this intact-wall calculation.

Data-driven studies require a different transfer test. Classifiers and machine-learning-assisted finite-element models can identify events or estimate parameters from labelled drilling data [30,69–71]. The geothermal examples by Mellal et al. and Aspiras et al. test real drilling streams [58,68], but fracture or incident detection is not equivalent to a calibrated collapse limit. A held-out well, time-ordered test, clear event definition, and comparison

with a physical or operational baseline are needed before a warning is interpreted as evidence for a pressure or casing decision.

Table 3 distinguishes the observed outcome from the larger claim often attached to it. The rows show why a multi-field loss record, a numerical injection scenario, a laboratory barrier test, and a workover case cannot be pooled into one validation-strength score. The observation needed for a new well follows the question being asked: fracture location for a loss mechanism, breakout geometry for a stress or criterion interpretation, and leakage or pressure response for a barrier.

Table 3. Geothermal evidence by observed or calculated outcome and inference limit.

Study and Stage	Direct Evidence or Calculation	Supported Inference	What Remains Untested
Winn et al. [5]; drilling	Records from four geothermal fields	Losses frequently associate with faults and productive intervals	A universal loss-pressure correction
Huan et al. [31]; injection	Coupled numerical scenario; 160 °C temperature difference	High computed stress at rock and cased/open-hole junction	Field-validated failure threshold
Hals and Berre [32]; operation	Thermal-fracture scenarios	Orientation and injected-water temperature affect predicted damage	Prospective well performance
Corina and Moghadam [37]; barrier	1.5 m assembly, 39 cycles between 80 and 20 °C	No measured permeability increase under tested conditions	High-temperature field sealability
Hart et al. [65]; construction	Field fibre-optic monitoring with model and laboratory comparison	Cement placement and hydration can be observed	Long-term barrier integrity
Marbun et al. [35]; workover	Two abandoned Dieng wells and incomplete histories	Casing condition can guide intervention	Definitive original failure attribution
Bujakowski et al. [36]; maintenance	Inspection and reconstruction cases in Poland	Defects and repair outcome can be documented	Validation of a pre-drill collapse forecast

Independent validation remains uncommon in wellbore-stability studies. Laboratory tests can support a failure mechanism, while post-drill back-analysis can calibrate a model to an individual well. Neither demonstrates that the calibrated model will predict another interval or well. Stronger validation requires the original prediction to be retained and compared with later observations that were not used for calibration. Both successful and unsuccessful predictions should be reported so that model performance can be assessed against simpler approaches.

Model confidence should be stated for the conditions under which it was tested, including formation, failure mechanism, pressure–temperature conditions, input uncertainty and validation data. For this review, complete lifecycle validation requires a model and its pre-event inputs fixed before drilling, a rock-failure or loss forecast tested against drilling outcomes excluded from calibration, and injection, production or barrier predictions tested against subsequent observations in the same well or a declared transfer well. The reviewed evidence does not establish this complete sequence. This finding applies to the included sources and does not establish its absence from the wider literature.

Table 4 reports the principal category changes in the Soultz reference scenario. Thermal disturbance, connected fractures, and injection activate coupled and fracture-aware layers. Removing one fracture-related input may leave the same mechanism active through another input. Reduced data or missing test evidence produces a warning even when the layer categories do not change. These are changes in the classification of modelling requirements, not calculated pressure or temperature limits.

Table 4. Sensitivity of categorical model requirements to changes in the Soultz reference inputs.

Input Change	Effect on Selected Layers	Interpretation and Limit
Thermal flag removed	THM changes from Required to Conditional.	The flag represents a judgement about thermal significance; the change does not show that cooling is negligible.
Both fracture-trigger inputs cleared	Numerical analysis changes from Required to Conditional; elastic stress and failure screening become Required.	Changing only formation or only the discontinuity flag leaves the fracture rule active.
Uncertainty raised from medium to high	Probabilistic treatment changes from Recommended to Required.	The category requests uncertainty analysis; it is not a calculated failure probability.
Data availability reduced or test evidence absent	Layer categories are retained; data-acquisition or screening warnings are added.	Missing observations prevent an operating-limit claim despite an unchanged category vector.

THM, thermal–hydraulic–mechanical modelling. The reference scenario and complete vectors are provided in S1A and Supplementary Material S2.

4.5. Selector Demonstrations and External Retrospective Evaluation

The evaluation comprises three geothermal demonstrations: GRT-1-to-GRT-2 drilling transfer [72], Soultz fractured-reservoir injection [73], and Polish injection-well reconstruction [36]. Szentes injectivity decline and Leyte casing inspection provide two development-linked assessments [74,75]. Material tests, numerical scenarios, reservoir-deformation observations and salt-closure analyses provide distinct evidence for time-dependent deformation [76–81]; their transfer limits are examined in Section 5.5. Two further sources, Groß Schönebeck stimulation and Dezhou reinjection, are excluded from rule formulation and evaluated against the frozen Wellbore Studio version 0.2.0 specification [82,83].

Table 5 consolidates the controlling conditions, selected model layers, published observations and limits of inference. Source-supported material and operating conditions are distinguished from analyst choices for uncertainty, data availability and consequence. The comparators are intact-wall analysis, uncoupled modelling or inspection guidance, as appropriate to the decision; they are conceptual baselines rather than calibrated competing field forecasts. S1A supplies the complete input/output vectors and sensitivity checks. Supplementary Material S2 records the expanded assessments and source-specific qualifications.

The assessments distinguish rock-wall failure, connected-fracture flow, particle transport and barrier diagnosis. Szentes requires a fines and injectivity assessment rather than automatic competent-rock screening. Leyte supports inspection and integrity assessment, while the need for stress modelling requires a separate rock-mechanics question. In the two sources excluded from rule development, the selector identifies relevant fracture or particle layers but does not select proppant-contact, reactive-scaling or screen-design laws. Mechanism coverage is therefore partial in both cases.

These comparisons demonstrate how the rules are interpreted and where additional models or observations are needed. They do not establish that the selector predicted damage, improved a historical field decision or outperformed a calibrated alternative. Independent

predictive testing requires fixed decision-time inputs, a recorded forecast, a predefined comparator and observations excluded from calibration and case encoding.

Table 5. Geothermal demonstrations and retrospective coverage assessments of the model-selection framework.

Case and Controlling Conditions	Principal Selected Layers	Reported Observation and Comparator	Limit of Inference
GRT-2 drilling transfer [72] Deviated well; limited local data.	Elastic stress and failure criterion Required; MEM and uncertainty Recommended.	Reported shear-failure threshold below 1.3 g cm^{-3}; instability at 1.07–1.12 g cm^{-3}.	Retrospective comparison; swelling is also plausible. No retained selector forecast.
Soultz injection [73] Fractured reservoir; thermal loading.	MEM, THM and fracture-aware analysis Required.	Coupled-model comparison with GPK-2 injectivity; uncoupled modelling is the conceptual comparator.	A history match constrains flow; it does not test a reserved rock-failure or barrier forecast.
Polish injection-well reconstruction [36] Maintenance; barrier decision.	Back-analysis and integrity linkage Required; MEM Recommended.	Inspection-led repair of Biały Dunajec PAN-1; post-repair capacity 375 m^3 h^{-1} at 6.3 MPa.	The post-repair check tests the intervention, not an earlier open-hole collapse model.
Szentes injection [74] Weak sandstone; fines/clogging.	Particle-transport/injectivity, uncertainty and integrity assessment Required.	Particle migration and clogging reported; injectivity improved after pump lift and temporarily after skin fracturing.	The output identifies a mechanism to assess; it does not establish treatment attribution or sustained benefit.
Leyte casing inspection [75] Maintenance; diagnostic inputs.	Back-analysis, uncertainty and integrity linkage Required; MEM Recommended.	Callipers in four selected wells among fourteen candidates showed thinning, deformation and a collar opening.	Inspection agreement does not establish stress-model necessity, external corrosion or remaining life.
Groß Schönebeck stimulation [82] Discontinuities; proppant/fines.	Fracture-aware analysis, MEM and fines assessment Required.	Laboratory flow tests and GtGrSk4/05 production/flowmeter observations; intact-wall analysis omits the reported mechanisms.	Partial coverage: contact, embedding and conductivity laws require additional local modelling.
Dezhou reinjection [83] Semi-consolidated sandstone; particles.	Injectivity/particle assessment Required; competent-rock failure screening Conditional.	PYR1–PY2 pumping, reinjection and tracer observations; intact-wall screening omits particle and screen blockage.	Partial coverage: screen/gravel design, scaling, oxygen and gas processes require additional analysis.

MEM, Mechanical Earth Model; THM, thermal–hydraulic–mechanical modelling. Only principal layers are shown. Source observations and analyst-assigned categories are separated in S1A. The three demonstrations and the Szentes/Leyte assessments inform interpretation of the specification; the two final sources are excluded from rule development. All outcomes were visible during encoding.

5. Discussion

The browser-based Wellbore Stability Model Selector implements the conditional framework. Users specify lifecycle stage, formation, geometry, pressure–temperature setting, expected mechanisms, data availability, uncertainty, validation evidence and integrity consequence. The tool returns Required, Recommended, Conditional and Support-Only model layers, with supporting reasons, observations needed for testing and limits of application. It links recommendations to relevant literature and exports the selection as a

machine-readable record. The output supports the choice of modelling approach; it does not calculate an operating pressure or service-life limit. Site-specific geomechanical analysis remains necessary for design decisions.

5.1. Transfer from Petroleum to Geothermal Wells

Petroleum studies provide useful stress, failure, and shale-transport methods, but geothermal transfer depends on the mechanism and loading path. Criterion comparisons established in sedimentary wells can guide a competent geothermal interval when local stress and strength are constrained [15,16,49]. They do not explain losses into a connected productive fracture [5] or thermal stress during long-term cold-water injection [31,32]. The model family may remain the same while the boundary conditions, failure surface, and observation used to test it change. This is why geothermal relevance is demonstrated by a matching process, not by relabelling a petroleum calibration.

The distinction is equally important for completed wells. Huan et al.'s coupled calculation locates potentially severe stresses at a cased-to-open-hole transition [31]; Corina and Moghadam measured stable sealing over a limited low-enthalpy cycling test [37]; field intervention reports identify damaged casing and specific repairs [35,36]. These studies examine different systems and cannot be compared as tests of the same prediction. Their joint implication is that open-hole fracture response, liner loading, and cement leakage require separate outputs and tests. A single "wellbore stability" label conceals those differences.

5.2. Evidence Gaps and Conflicting Results

The source comparisons reveal a reporting limitation: explanatory fits and calibrated simulations are often presented without a retained forecast and later observations reserved for testing. Pressure, temperature, damage, and intervention histories are needed to test a model across the well lifecycle. Calibration to a later borehole image or casing defect does not establish predictive performance. The evidence considered here is insufficient to rank long-term creep laws or certify a barrier life through the selector.

The comparisons support a conceptual warning against ranking models by physical detail alone; they do not demonstrate comparative predictive accuracy. The choice of compressive criterion changed calculated mud-weight windows in one field case [49]; a damage model reproduced laboratory breakout geometry under specified stress conditions [62]; geothermal THM simulations predicted high local stress under severe cooling [31,32]. None of these results establishes the same measure of accuracy. Future comparisons should report a common predicted quantity, its uncertainty, the subsequent observed failure or non-failure, and the simpler model used as a baseline. Table 6 identifies the measurements that would discriminate among the currently plausible explanations. The 43.7 and 31.9 MPa values in Huan et al. [31] and the 39-cycle result of Corina and Moghadam [37] remain specific to their simulated or tested conditions and are not operating or safe-cycle limits.

A particularly useful negative result would show that adding a mechanism changes calculated stress but not the selected operation, or that it fails an independent well test. The existing literature tends to report successful fits and illustrative scenarios; without failed predictions, the decision benefit of additional physics or machine learning remains uncertain [28,49,50,64,69–71]. Preserving unsuccessful comparisons would help separate genuine transfer from parameter adjustment after an event.

Calibration and validation must therefore be reported separately. Stress magnitude inferred from a breakout, for example, can be used to explain that breakout; testing the derived state against a different depth or well is a stronger validation [53,54]. A drilling classifier trained on nearby samples can similarly appear accurate because the samples share the same event history; time-ordered and well-level holdouts are more relevant to

deployment [58,68]. The matrix records source descriptions of validation, but it cannot reconstruct an independent test where a publication does not report one.

Table 6. Evidence gaps and observations that would distinguish competing explanations.

Unresolved Comparison	Evidence Prompting It	Discriminating Observation
Intact failure versus fracture loss	Criterion studies [15,49] and field losses [5]	Loss interval, dynamic pressure, fracture image, and treatment response
Stress versus damage-law error	Borehole images [53,54] and damage model [62]	Frozen breakout forecast across another interval or well
Cooling-induced rock fracture	THM and orientation scenarios [31,32]	Field injection temperature, pressure, fracture response, and non-failure cases
Thermal barrier damage	Sealing model [11] and low-enthalpy test [37]	Comparable material, cycle and pressure-controlled leak tests
Incident detection versus decision benefit	Geothermal classifiers [58,68]	Held-out well, time-ordered alert, action taken, and outcome
Construction defect versus later leakage	Fibre monitoring [65] and workover cases [35,36]	Linked placement, inspection, pressure-test, and repair history
Calibration versus transfer	Field back-analyses [49,53,54]	Retained pre-event prediction tested on a separate interval or well

5.3. Implications for Model Use

The framework's practical use is to make a model claim auditable. A stability recommendation should identify the failure mode and pressure or trajectory decision it informs, the data used to fit the model, and the observation reserved to test it. For geothermal injection, rock-fracture response and casing–cement sealing require different observations [31,37]. This format exposes when a result is only a sensitivity calculation and when field evidence supports changing an operation. It does not prescribe a universal workflow or replace the operator's site-specific design assessment.

A model revision is scientifically informative only if the pre-event estimate remains available. Where a well loses fluid below the predicted upper bound, geological fracture connectivity, equivalent circulating density, and the assumed failure criterion are alternative explanations [5,15,49]. The observed loss pressure can help choose among them when its depth, dynamic pressure, and treatment history are recorded. In a later workover, inspection and pressure testing can establish the state of a barrier [35,36]; they are not retrospective measurements of the original open-hole collapse limit. Maintaining that distinction prevents the lifecycle narrative from conflating rock mechanics with well integrity.

5.4. Limits of the Review

This structured critical review supports interpretation within a purposively selected evidence base. The broad discovery inventory is not fully eligibility-screened, and neither exhaustive coverage, field-wide prevalence nor inter-assessor agreement is estimated. Source roles and review extent distinguish claim-specific original-paper appraisal from context, preprint and abstract-only support; this is not a complete risk-of-bias assessment. Case outcomes were available during input encoding. Two cases inform the specification; two additional sources are excluded from rule development and assessed after a recorded rule freeze. Their retrospective mechanism agreement does not establish predictive accuracy. The selector supplies no calibrated failure probabilities, numerical operating limits or treatment-success forecasts. Application requires observations appropriate to the stated

stress, material, thermal, fracture-flow or barrier decision. Long-term creep remains outside the categorical constitutive selection.

5.5. Long-Term Deformation and Scale of Inference

The timescale of a geothermal decision changes the required constitutive description. Instantaneous strength and elastic thermal-stress calculations address different outcomes from accumulated creep, thermally induced damage, and repeated barrier loading. A model suitable for a short drilling exposure may therefore require additional state variables and observations when used to assess prolonged injection or production. The need for that extension depends on material response, temperature and pressure history, and the magnitude of deformation relative to the operating margin.

Jiang et al. measured temperature-dependent limestone creep [76], and Zhang and Zhao represented creep and thermal damage in long-duration geothermal simulations [77]. Jain and Singh compared a temperature-sensitive elastoplastic wellbore solution with finite-element results [78]. The Heber study constrained reservoir deformation using a multi-decadal field history [79], while the Soultz study compared a coupled fractured-reservoir model with injectivity observations [73]. These sources address distinct outcomes and spatial scales.

The limestone tests establish material-scale creep behaviour [76], and the 20-year scenarios illustrate a coupled model rather than a field forecast [77]. Heber observations constrain reservoir-scale deformation, not borehole creep or casing failure [79]. The elasto-plastic comparison addresses model form rather than creep [78]; the Soultz history match constrains flow, not long-term borehole integrity [73]. Long-term creep is relevant in susceptible rocks, but calibration and well-scale predictive evidence remain separate questions.

Salt closure and fractured-rock creep require different constitutive descriptions. Weijermars et al. identify effective viscosity, pressure imbalance and lateral far-field motion as controls on uncased salt-well closure [80]. Zhang et al. report six creep tests on pre-fractured granite from one block under three confining-pressure levels and acid or alkaline flow [81]. The measured strain and permeability responses support coupling between sustained stress, fracture deformation and fluid–rock interaction. These analytical and specimen-scale results do not establish decades of geothermal borehole performance. The granite study lacks replicated test cells and contains an unresolved specimen–solution assignment; specimen-specific failure attribution is not used as a quantitative selection threshold. Detailed source checks are retained in S1 and Supplementary Material S2.

Where expected service-life deformation could change a rock-stability or barrier decision, a time-dependent constitutive analysis should be considered alongside thermal and pore-pressure histories. Its parameters require material-specific creep tests or defensible field calibration; repeated calliper, image, pressure, temperature, and barrier observations should be reserved to test the forecast. The selector flags creep as outside its categorical rules and does not choose a constitutive law or predict a service lifetime. When those inputs are absent, sensitivity bounds and an explicit monitoring trigger are more defensible than a precise lifetime prediction. S1 and Supplementary Material S2 document the material-specific evidence and its transfer limits.

6. Conclusions

This review considered how to choose wellbore-stability models for geothermal decisions from construction through injection, operation, and intervention. It distinguished model capability from the evidence needed to transfer a prediction. The main findings are as follows:

(1) The reviewed studies do not establish that more physical detail improves prediction; this is a conceptual inference, not a pooled error comparison. Failure criteria, bedding, shale–fluid interactions, connected fractures, thermal loading, and time-dependent deformation matter when they can change the decision under local conditions.
(2) Published comparisons support matching a model to the controlling mechanism and matching its test to the stated output. The proposed framework turns that evidence into conditional model-family recommendations; those recommendations remain the authors' judgments. A loss event, borehole image, injection pressure response, and barrier leak test assess different outputs; none is interchangeable with a calibration datum.
(3) Independent longitudinal testing remains a central gap in the reviewed evidence. Retained pre-event predictions should be tested against later intervals or wells, reporting correct and failed forecasts and preserving operating and intervention histories. Weak-material injectivity requires explicit assessment of particle mobility, while maintenance recommendations must distinguish inspection from stress modelling and remaining-life prediction. The case assessments illustrate these distinctions, but their outcome-informed interpretation does not establish predictive accuracy. No universal thermal-stress or safe-cycle threshold is supported; independent testing with fixed inputs and reserved outcomes remains necessary.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/en19215145/s1, S1 provides the evidence matrix, study characteristics, eligibility and source-identity records, claim-specific appraisal, source locators and editable assessment rubric. S1A provides the model-selection rules, source-to-rule links, case inputs and outputs, sensitivity analyses, claim checks and reference crosswalk. S2 provides expanded demonstrations, retrospective case assessments, categorical sensitivity details and material-specific creep source checks. Supporting search records and reproducible analysis files are available at https://doi.org/10.5281/zenodo.23496344.

Author Contributions: Conceptualization, I.N. and D.K.; methodology, I.N.; investigation, I.N.; data curation, I.N.; formal analysis, I.N. and D.K.; writing—original draft preparation, I.N. and D.K.; writing—review and editing, I.N. and D.K.; visualization, I.N.; supervision, D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the AGH University of Krakow under the Subsidy and Excellence Initiative–Research University (IDUB) programme.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The evidence matrix and claim appraisal are provided in Supplementary Material S1. Model-selection rules, case inputs and outputs, sensitivity analyses, and source-to-claim records are provided in S1A; expanded assessment details are provided in S2. Supporting search records, appraisal audit files, the selector implementation and verification scripts are archived at Zenodo: https://doi.org/10.5281/zenodo.23496344. The companion interface, Wellbore Studio version 0.2.0, reports screening status and required observations at https://najafiib.github.io/wellbore-stability-model-selector/ (accessed on 20.10.2026).

Acknowledgments: OpenAI Codex assisted with literature triage, source-text extraction, bibliographic checking, claim appraisal and retrospective case evaluation, drafting and language editing, preparation of figures and spreadsheets, and software implementation and testing. AI-assisted outputs support the documented analysis and do not constitute independent human assessment or predictive validation. The authors are responsible for source verification, scientific interpretation and the manuscript content.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. International Energy Agency (IEA). *The Future of Geothermal Energy*; IEA: Paris, France, 2024. Available online: https://www.iea.org/reports/the-future-of-geothermal-energy (accessed on 7 October 2026).
2. U.S. Department of Energy; Office of Geothermal. *Subsurface Accessibility*; U.S. Department of Energy, Office of Geothermal: Washington, DC, USA, 2026. Available online: https://www.energy.gov/hgeo/geothermal/subsurface-accessibility (accessed on 29 September 2026).
3. Vivas, C.; Salehi, S.; Tuttle, J.D.; Rickard, B. Challenges and Opportunities of Geothermal Drilling for Renewable Energy Generation. *GRC Trans*. **2020**, *44*, 904–918.
4. Saleh, F.; Teodoriu, C.; Salehi, S.; Ezeakacha, C.P. Geothermal Drilling: A Review of Drilling Challenges with Mud Design and Lost Circulation Problem. In Proceedings of the 45th Workshop on Geothermal Reservoir Engineering, Stanford, CA, USA, 10–12 February 2020. Available online: https://pangea.stanford.edu/ERE/db/IGAstandard/record_detail.php?id=29309 (accessed on 7 October 2026).
5. Winn, C.; Dobson, P.; Ulrich, C.; Kneafsey, T.; Lowry, T.S.; Akerley, J.; Delwiche, B.; Samuel, A.; Bauer, S. Context and mitigation of lost circulation during geothermal drilling in diverse geologic settings. *Geothermics* **2023**, *108*, 102630. [CrossRef]
6. Chen, G.; Ewy, R.T. Thermoporoelastic Effect on Wellbore Stability. *SPE J*. **2005**, *10*, 121–129. [CrossRef]
7. Wang, Y.; Dusseault, M.B. A coupled conductive-convective thermo-poroelastic solution and implications for wellbore stability. *J. Pet. Sci. Eng*. **2003**, *38*, 187–198. [CrossRef]
8. Chen, X.; Gao, D.; Yang, J.; Luo, M.; Feng, Y.; Li, X. A Comprehensive Wellbore Stability Model Considering Poroelastic and Thermal Effects for Inclined Wellbores in Deepwater Drilling. *J. Energy Resour. Technol*. **2018**, *140*, 092903. [CrossRef]
9. Allahvirdizadeh, P. A review on geothermal wells: Well integrity issues. *J. Clean. Prod*. **2020**, *275*, 124009. [CrossRef]
10. Arbad, N.; Emadi, H.; Watson, M. A comprehensive review of geothermal cementing from well integrity perspective. *J. Pet. Sci. Eng*. **2022**, *217*, 110869. [CrossRef]
11. Deng, K.; Zhou, N.; Lin, Y.; Peng, Y.; Yan, K.; Qin, H.; Xie, P.; Li, Z. Failure mechanism and control method of cement sheath sealing integrity under alternating thermal-stress coupling in geothermal wells. *Renew. Energy* **2024**, *226*, 120448. [CrossRef]
12. Thorbjornsson, I.O.; Kaldal, G.S.; Krogh, B.C.; Palsson, B.; Markusson, S.H.; Sigurdsson, P.; Einarsson, A.; Gunnarsson, B.S.; Jonsson, S.S. Materials investigation of the high temperature IDDP-1 wellhead. *Geothermics* **2020**, *87*, 101866. [CrossRef]
13. Guenot, A. Borehole breakouts and stress fields. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1989**, *26*, 185–195. [CrossRef]
14. Haimson, B. Micromechanisms of borehole instability leading to breakouts in rocks. *Int. J. Rock Mech. Min. Sci*. **2007**, *44*, 157–173. [CrossRef]
15. McLean, M.R.; Addis, M.A. Wellbore Stability: The Effect of Strength Criteria on Mud Weight Recommendations. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 23–26 September 1990. [CrossRef]
16. Ma, T.; Chen, P.; Yang, C.; Zhao, J. Wellbore stability analysis and well path optimization based on the breakout width model and Mogi-Coulomb criterion. *J. Pet. Sci. Eng*. **2015**, *135*, 678–701. [CrossRef]
17. Singh, A.; Rao, K.S.; Ayothiraman, R. An analytical solution to wellbore stability using Mogi–Coulomb failure criterion. *J. Rock Mech. Geotech. Eng*. **2019**, *11*, 1211–1230. [CrossRef]
18. Lee, H.; Ong, S.H.; Azeemuddin, M.; Goodman, H. A wellbore stability model for formations with anisotropic rock strengths. *J. Pet. Sci. Eng*. **2012**, *96–97*, 109–119. [CrossRef]
19. Ong, S.H.; Roegiers, J.C. Influence of anisotropies in borehole stability. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1993**, *30*, 1069–1075. [CrossRef]
20. Zhang, J.; Bai, M.; Roegiers, J.C. Dual-porosity poroelastic analyses of wellbore stability. *Int. J. Rock Mech. Min. Sci*. **2003**, *40*, 473–483. [CrossRef]
21. Bol, G.M.; Wong, S.W.; Davidson, C.J.; Woodland, D.C. Borehole Stability in Shales. *SPE Drill. Complet*. **1994**, *9*, 87–94. [CrossRef]
22. Chen, G.; Chenevert, M.E.; Sharma, M.M.; Yu, M. A study of wellbore stability in shales including poroelastic, chemical, and thermal effects. *J. Pet. Sci. Eng*. **2003**, *38*, 167–176. [CrossRef]
23. Yu, M.; Chenevert, M.E.; Sharma, M.M. Chemical-mechanical wellbore instability model for shales: Accounting for solute diffusion. *J. Pet. Sci. Eng*. **2003**, *38*, 131–143. [CrossRef]
24. Zeynali, M.E. Mechanical and physico-chemical aspects of wellbore stability during drilling operations. *J. Pet. Sci. Eng*. **2012**, *82–83*, 120–124. [CrossRef]
25. Ma, T.; Chen, P. A wellbore stability analysis model with chemical-mechanical coupling for shale gas reservoirs. *J. Nat. Gas Sci. Eng*. **2015**, *26*, 72–98. [CrossRef]
26. Rafieepour, S.; Zamiran, S.; Ostadhassan, M. A cost-effective chemo-thermo-poroelastic wellbore stability model for mud weight design during drilling through shale formations. *J. Rock Mech. Geotech. Eng*. **2020**, *12*, 768–779. [CrossRef]
27. Kang, Y.; Yu, M.; Miska, S.; Takach, N.E. Wellbore Stability: A Critical Review and Introduction to DEM. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 4–7 October 2009. [CrossRef]

28. Moos, D.; Peska, P.; Finkbeiner, T.; Zoback, M. Comprehensive wellbore stability analysis utilizing Quantitative Risk Assessment. *J. Pet. Sci. Eng.* **2003**, *38*, 97–109. [CrossRef]
29. Santarelli, F.J.; Dahen, D.; Baroudi, H.; Sliman, K.B. Mechanisms of borehole instability in heavily fractured rock media. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1992**, *29*, 457–467. [CrossRef]
30. Xu, K.; Liu, Z.; Chen, Q.; Zhang, Q.; Ling, X.; Cai, X.; He, Q.; Yang, M. Application of machine learning in wellbore stability prediction: A review. *Geoenergy Sci. Eng.* **2024**, *232*, 212409. [CrossRef]
31. Huan, X.; Xu, G.; Zhang, Y.; Sun, F.; Xue, S. Study on Thermo-Hydro-Mechanical Coupling and the Stability of a Geothermal Wellbore Structure. *Energies* **2021**, *14*, 649. [CrossRef]
32. Hals, K.M.D.; Berre, I. Thermal Fracturing of Geothermal Wells and Effects of Borehole Orientation. In Proceedings of the 38th Workshop on Geothermal Reservoir Engineering, Stanford, CA, USA, 11–13 February 2013. Available online: https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2013/Hals.pdf (accessed on 7 October 2026).
33. Marbun, B.T.H.; Ridwan, R.H.; Sinaga, S.Z.; Purbantanu, B.A. Design of injection liner of high-temperature vapor-dominated geothermal reservoir. *Renew. Energy* **2020**, *150*, 649–655. [CrossRef]
34. Marbun, B.T.H.; Ridwan, R.H.; Nugraha, H.S.; Sinaga, S.Z.; Purbantanu, B.A. Casing setting depth and design of production well in water-dominated geothermal system with 330 °C reservoir temperature. *Energy Rep.* **2020**, *6*, 582–593. [CrossRef]
35. Marbun, B.T.H.; Ridwan, R.H.; Sinaga, S.Z.; Pande, B.; Purbantanu, B.A. Casing failure identification of long-abandoned geothermal wells in Field Dieng, Indonesia. *Geotherm. Energy* **2019**, *7*, 31. [CrossRef]
36. Bujakowski, W.; Bielec, B.; Miecznik, M.; Pająk, L. Reconstruction of geothermal boreholes in Poland. *Geotherm. Energy* **2020**, *8*, 10. [CrossRef]
37. Corina, A.N.; Moghadam, A. The Sealing Performance of Cement Sheaths under Thermal Cycles for Low-Enthalpy Geothermal Wells. *Energies* **2024**, *17*, 239. [CrossRef]
38. Mahetaji, M.; Brahma, J. Prediction of Minimum Mud Weight for Prevention of Breakout Using New 3D Failure Criterion to Maintain Wellbore Stability. *Rock Mech. Rock Eng.* **2024**, *57*, 2231–2252. [CrossRef]
39. Afsari, M.; Ghafoori, M.R.; Roostaeian, M.; Haghshenas, A.; Ataei, A.; Masoudi, R. Mechanical Earth Model (MEM): An Effective Tool for Borehole Stability Analysis and Managed Pressure Drilling (Case Study). In Proceedings of the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 15–18 March 2009. [CrossRef]
40. Afsari, M.; Amani, M.; Razmgir, S.M.; Karimi, H.; Yousefi, S. Using Drilling and Logging Data for Developing 1D Mechanical Earth Model for a Mature Oil Field to Predict and Mitigate Wellbore Stability Challenges. In Proceedings of the International Oil and Gas Conference and Exhibition in China, Beijing, China, 8–10 June 2010. [CrossRef]
41. Lee, D.; Cassanelli, J.P.; Frydman, M.; Palacio, J.; Delgado, R.; Collins, B. Using a Dynamic Mechanical Earth Model and Integrated Drilling Team to Reduce Well Costs and Drilling Risks in San Martin Field. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 5–8 October 2003. [CrossRef]
42. Moazzeni, A.; Nabaei, M.; Shahbazi, K.; Shadravan, A. Mechanical Earth Modeling Improves Drilling Efficiency and Reduces Non-Productive Time (NPT). In Proceedings of the SPE Deep Gas Conference and Exhibition, Manama, Bahrain, 24–26 January 2010. [CrossRef]
43. Motahari, M.; Hashemi, A.; Molaghab, A. Successful mechanical earth model construction and wellbore stability analysis using elastic and plasticity solutions, a case study. *Geomech. Energy Environ.* **2022**, *32*, 100357. [CrossRef]
44. Darvishpour, A.; Cheraghi Seifabad, M.; Wood, D.A.; Ghorbani, H. Wellbore stability analysis to determine the safe mud weight window for sandstone layers. *Pet. Explor. Dev.* **2019**, *46*, 1031–1038. [CrossRef]
45. Sen, S.; Kundan, A.; Kumar, M. Modeling Pore Pressure, Fracture Pressure and Collapse Pressure Gradients in Offshore Panna, Western India: Implications for Drilling and Wellbore Stability. *Nat. Resour. Res.* **2020**, *29*, 2717–2734. [CrossRef]
46. Das, B.; Chatterjee, R. Wellbore stability analysis and prediction of minimum mud weight for few wells in Krishna-Godavari Basin, India. *Int. J. Rock Mech. Min. Sci.* **2017**, *93*, 30–37. [CrossRef]
47. Radwan, A.E. Drilling in Complex Pore Pressure Regimes: Analysis of Wellbore Stability Applying the Depth of Failure Approach. *Energies* **2022**, *15*, 7872. [CrossRef]
48. Allawi, R.H.; Al-Jawad, M.S. Wellbore stability analysis and selecting safe mud weight window for Mishrif reservoir in Southern Iraq. *Arab. J. Geosci.* **2023**, *16*, 345. [CrossRef]
49. Gholami, R.; Rabiei, M.; Rasouli, V.; Aadnoy, B.; Fakhari, N. Application of quantitative risk assessment in wellbore stability analysis. *J. Pet. Sci. Eng.* **2015**, *135*, 185–200. [CrossRef]
50. Udegbunam, J.E.; Aadnoy, B.S.; Fjelde, K.K.R. Uncertainty evaluation of wellbore stability model predictions. *J. Pet. Sci. Eng.* **2014**, *124*, 254–263. [CrossRef]
51. de Fontoura, S.A.B.; Holzberg, B.B.; Teixeira, D.C.; Frydman, M. Probabilistic Analysis of Wellbore Stability During Drilling. In Proceedings of the SPE/ISRM Rock Mechanics Conference, Irving, TX, USA, 20–23 October 2002. [CrossRef]
52. Asaka, M.; Holt, R.M. Anisotropic Wellbore Stability Analysis: Impact on Failure Prediction. *Rock Mech. Rock Eng.* **2021**, *54*, 583–605. [CrossRef]

53. Brudy, M.; Zoback, M.D. Drilling-induced tensile wall-fractures: Implications for determination of in-situ stress orientation and magnitude. *Int. J. Rock Mech. Min. Sci.* **1999**, *36*, 191–215. [CrossRef]
54. Wiprut, D.; Zoback, M.; Hanssen, T.H.; Peska, P. Constraining the full stress tensor from observations of drilling-induced tensile fractures and leak-off tests: Application to borehole stability and sand production on the Norwegian margin. *Int. J. Rock Mech. Min. Sci.* **1997**, *34*, e365.e1–e365.e12. [CrossRef]
55. Nelson, E.J.; Meyer, J.J.; Hillis, R.R.; Mildren, S.D. Transverse drilling-induced tensile fractures in the West Tuna area, Gippsland Basin, Australia: Implications for the in situ stress regime. *Int. J. Rock Mech. Min. Sci.* **2005**, *42*, 361–371. [CrossRef]
56. Zhang, H.; Yin, S.; Aadnoy, B.S. Finite-Element Modeling of Borehole Breakouts for In Situ Stress Determination. *Int. J. Geomech.* **2018**, *18*, 04018174. [CrossRef]
57. Marbun, B.T.H.; Ridwan, R.H.; Nugraha, H.S.; Sinaga, S.Z.; Purbantanu, B.A. Review of directional drilling design and operation of geothermal wells in Indonesia. *Renew. Energy* **2021**, *176*, 135–152. [CrossRef]
58. Mellal, I.; Gabry, M.A.; Soliman, M.Y.; Khaled, M.S. Real-Time Automated Fracture Detection in Geothermal Wells Using Low-Cost Drilling Data. In Proceedings of the SPE Annual Technical Conference and Exhibition, Houston, TX, USA, 20–22 October 2025. [CrossRef]
59. Chatterjee, S.; Mukherjee, S. Review on drilling-induced fractures in drill cores. *Mar. Pet. Geol.* **2023**, *151*, 106089. [CrossRef]
60. Akbarpour, M.; Abdideh, M. Wellbore stability analysis based on geomechanical modeling using finite element method. *Model. Earth Syst. Environ.* **2020**, *6*, 617–626. [CrossRef]
61. Figenschou, J.H.; Shaver, M.A.; Caldas, P.F.B.; Criollo, D. Wellbore Damage Identification Combining High-Resolution Ultrasonic Images and Time-Based Drilling Parameters. In Proceedings of the 59th U.S. Rock Mechanics/Geomechanics Symposium, Santa Fe, NM, USA, 8–11 June 2025. [CrossRef]
62. Sahara, D.P.; Schoenball, M.; Gerolymatou, E.; Kohl, T. Analysis of borehole breakout development using continuum damage mechanics. *Int. J. Rock Mech. Min. Sci.* **2017**, *97*, 134–143. [CrossRef]
63. Sahara, D.P.; Schoenball, M.; Kohl, T.; Muller, B.I.R. Impact of fracture networks on borehole breakout heterogeneities in crystalline rock. *Int. J. Rock Mech. Min. Sci.* **2014**, *71*, 301–309. [CrossRef]
64. Ray, P.L.; Tipton, J.R.; Cheng, S.; Francom, D. Wellbore Stability Uncertainty Quantification Using a Bayesian Machine Learning Framework. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 25 September 2024. [CrossRef]
65. Hart, J.; Polat, B.; Wollin, C.; Lipus, M.; Schölderle, F.; Ledig, T.; Jousset, P.; Reinsch, T.; Krawczyk, C.M. A fiber optic approach for cement placement and hydration assessment of deep geothermal boreholes. *Sci. Rep.* **2025**, *15*, 11365. [CrossRef] [PubMed]
66. Kanfar, M.F.; Chen, Z.; Rahman, S.S. Effect of material anisotropy on time-dependent wellbore stability. *Int. J. Rock Mech. Min. Sci.* **2015**, *78*, 36–45. [CrossRef]
67. Kanfar, M.F.; Chen, Z.; Rahman, S.S. Risk-controlled wellbore stability analysis in anisotropic formations. *J. Pet. Sci. Eng.* **2015**, *134*, 214–222. [CrossRef]
68. Aspiras, A.H.; Zarrouk, S.J.; Winmill, R.; Kempa-Liehr, A.W. Real-time incident detection in geothermal drilling through machine learning. *Renew. Energy* **2025**, *250*, 123260. [CrossRef]
69. Okpo, E.E.; Dosunmu, A.; Odagme, B.S. Artificial Neural Network Model for Predicting Wellbore Instability. In Proceedings of the SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 2–4 August 2016. [CrossRef]
70. Liu, P.; Li, J.; Chen, B.; Yan, G.; Lei, Q.; Liang, L.; Huang, Y.; Zhao, H.; Wang, G.; Sun, M. Digital Wellbore Stability Prediction with Machine Learning. In Proceedings of the International Petroleum Technology Conference, Dhahran, Saudi Arabia, 12–14 February 2024. [CrossRef]
71. AlBahrani, H.; Morita, N. Risk-Controlled Wellbore Stability Criterion Based on Machine Learning Assisted Finite Element Model. In Proceedings of the SPE/IADC International Drilling Conference and Exhibition, Virtual, 8–12 March 2021. [CrossRef]
72. Sosio, G.; Boivineau, A.-S.; Burachok, O.; Ould Braham, R.; Zordan, E.; Mandiuc, A.; Spyrou, C.; Belouahchia, M.; Baujard, C.; Dalmais, E.; et al. Integrated Geological, Fluid Flow and Geomechanical Model of a Geothermal Field. In Proceedings of the European Geothermal Congress, Strasbourg, France, 19–24 September 2016. Available online: https://www.researchgate.net/publication/334560279_Integrated_Geological_Fluid_Flow_and_Geomechanical_Model_of_a_Geothermal_Field (accessed on 7 October 2026).
73. Abdel Azim, R.; Alatefi, S.; Aljehani, A. A fully coupled thermo-poroelastic model for energy extraction in naturally fractured geothermal reservoirs: Sensitivity analysis and flow simulation. *Geotherm. Energy* **2024**, *12*, 26. [CrossRef]
74. Farkas, M.P.; Magyar, G.; Hofmann, H.; Zimmermann, G.; Fedor, F. Comparison of stimulation techniques in a geothermal injection well in a sedimentary aquifer in Szentes, Hungary. *Geotherm. Energy* **2025**, *13*, 20. [CrossRef]
75. Buñing, B.C.; Sarmiento, Z.F.; Aleman, E.T.; Saw, V.S. Casing Inspection Caliper Surveys: Results and Implications to Operations in Leyte Geothermal Production Field. In Proceedings of the World Geothermal Congress, Antalya, Turkey, 24–29 April 2005; p. 2038. Available online: https://www.worldgeothermal.org/pdf/IGAstandard/WGC/2005/2038.pdf (accessed on 7 October 2026).

76. Jiang, H.; Jiang, A.; Zhang, F.; Feng, Y. Study on creep characteristics and fractional creep damage constitutive model of limestone under thermal-mechanical action. *Int. J. Damage Mech.* **2023**, *32*, 262–288. [CrossRef]
77. Zhang, Y.; Zhao, G.-F. A multiphysics method for long-term deformation analysis of reservoir rock considering thermal damage in deep geothermal engineering. *Renew. Energy* **2023**, *204*, 432–448. [CrossRef]
78. Jain, G.; Singh, A. An elastoplastic semi-analytical solution for enhanced geothermal wellbore stability considering temperature-sensitive failure criterion. *Geothermics* **2024**, *121*, 103046. [CrossRef]
79. Jiang, G.; Barbour, A.J.; Skoumal, R.J.; Materna, K.; Taron, J.; Crandall-Bear, A. Relatively stable pressure effects and time-increasing thermal contraction control Heber geothermal field deformation. *Nat. Commun.* **2024**, *15*, 5159. [CrossRef] [PubMed]
80. Weijermars, R.; Jackson, M.P.A.; van Harmelen, A. Closure of open wellbores in creeping salt sheets. *Geophys. J. Int.* **2014**, *196*, 279–290. [CrossRef]
81. Zhang, F.; Zhao, J.; Hu, D.; Sheng, Q.; Shao, J. Creep Strain and Permeability Evolution in Cracked Granite Subjected to Triaxial Stress and Reactive Flow. *Geofluids* **2018**, *2018*, 4653937. [CrossRef]
82. Zimmermann, G.; Reinicke, A. Hydraulic stimulation of a deep sandstone reservoir to develop an Enhanced Geothermal System: Laboratory and field experiments. *Geothermics* **2010**, *39*, 70–77. [CrossRef]
83. Zheng, T.; Axelsson, G.; Kang, F.; Zhao, J.; Zhang, P.; Bai, T.; Yang, X.; Zhou, Q.; Luo, Y. Sustainable geothermal yield through effective reinjection: Reservoir engineering case study of the Dezhou sandstone geothermal reservoir, North Shandong Plain, China. *Geotherm. Energy* **2025**, *13*, 36. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.